

UNIVERSITAT DE BARCELONA

Anàlisi genètica i molecular de la síndrome de Down. Aïllament de gens a la regió cromosòmica 21q22.2-q22.3: identificació i caracterització del gen Minibrain (MNBH)

Jordi Guimerà i Vilaró

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Centre de Genètica Mèdica i Molecular Institut de Recerca Oncològica (IRO) Hospital Duran i Reynals

Anàlisi genètica i molecular de la síndrome de Down: aïllament de gens a la regió cromosòmica 21q22.2-q22.3, identificació i caracterització del gen *Minibrain (MNBH*)

Jordi Guimerà i Vilaró

Barcelona, setembre 1998

A *María Sosé* i al nostre fill *Felip*

•*

2

•

-

A la meva mare Mercè i als meus germans Francesc i Olga Als meus avis Pere i Andrea Al meu tiet Josep

•*

•

CACACACACACACACACACACACAGCTAGCGCTATATCGTAGCAGCCGCGCAATAAG CTTTTTTTTTTTTTTTTTTCATCGCGCGCGCGCGCATCTAGCGATTCGGCGCGCGGTAAATCGACA GCAGCGAGCGAGCATGCTGCTAGCTALUGATGCTAGCTAGCTAGCGCTATATCGTAGCA GCCGCGCGCGATATGCTAGCTGATGCCACACACGATGCTAGCTGACGATALUCGGCATG CAGCCGACGAGCATTGCAGTTCGAGCATGCTGTATCGAGCAGCTGAGCGACATCCGATG ${\tt CTGATCGATGCGATCGAGCTATCTGALUAGCGCGATATATGTATGAGCCGGTAGCGCAT}$ CCGGCATATTCGATGCGCGCGCTAGCTALUACGTATGGCGCATCTAGCGATTCGGCGCGCGGT CTATATCGTAGCAGCCGCGCAAAAGCGATATGCTAGCTGATGCCGCTGACGATCGGCAT GCTAALUTGCATTGAGCAGCGAGCACGACGAGCAGCAGCCGACGAGCATTGCAGTTCGA GCATGCTGACGTATCGAGCAGCTGAGCGACATCCGATGCTGATCGATGCGATCGACTAT GCTGATGCTAGCTAGCTAGCGCGCTAALUTATCGTAGCAGCCGCGCGCGATATGCTAGCTG ATGCCACACATATATATATATACGATCAACCTTTCGCGGGTCTCTTTCTCCCCCTTCCTG CTAGCTCAGCAGCPCISLANDCGAGCACGAMNBHCGAGCAGCGACGTGCAGCCGACGAG CATTGCAGTTCGAGCATGCTGACGTATCGAGCAGCTGAGCGACATCCGATGCTGATCGA TGCGATCGAGCTATCTGAGCGCGATATCCCCGATATAGCCGGTAGCGCATTAAAACGGTT ACGTATGGGCGCATCTAGCGATTCGGCGCGCGALUGTAAATCGACAGCAGCGAGCGAGCAT AGCGCTATATCGTAGCAGCCGCCGCAAAAGCGATATGCTAGCTGATGCCGCTGACGATCG CGAGCATGCTGACGTATCGAGCAGCGCCGCGCGCGATATGCTAGCTGATGCCACACACG ATGCTAGCTGACGATCGGCATGCTATGCATGAGCAGCGACGACGACGACGACGACGTG CAGCCGACGAGCATTGCAGTTCGAGCATGCTGACGTATCGAGCAGCTGAGCGACATCCG ATGCTTAAAAAAAAAAAGATCGATGALUCGATCGAGCTATCTGAGCGCGATATATATAGC TTCGATGCGCGCTATGCTACGTATGGCGCATCTAGCGATTCGTTGGTGTGGGCAGCGGC

...i és tan gran la nostra ignorància,

que en voler explicar quelcom tan complicat d'una manera senzilla

ens ha calgut vint-i-set lletres

per expressar allò

pel què la natura tan sols en necessita quatre.

I. AGRAÏMENTS

L'home només té l'opció de fer grans avenços quan suma voluntats. Fruït d'aquest esperit, aquesta tesi s'ha d'entendre com el resultat de la suma sinèrgica de les voluntats de qui, durant aquests sis anys han estat al meu costat i m'han ajudat abrivadament i que mai han decaigut en els moments difícils, i que quan miro enrera, me n'adono que són molts.

En especial voldria agrair:

Al Dr. Xavier Estivill per donar-me l'oportunitat de realitzar una tesi doctoral sobre la síndrome de Down, de treballar en el seu grup i per confiar en mi des del primer moment sense deixar de fer-ho en cap ocasió. Per la seva passió per la ciència i per la seva capacitat de treball, i per oferir-nos a tots unes infrastructures envejables.

A la Dra. Melanie Pritchard per haver-me dirigit el meu treball durant aquest anys d'una manera minuciosa, per la seva gran saviesa i experiència. Per intentar domar el meu anglès salvatge, per ajudar-me en les comunicacions orals dels congressos, per la seva gran capacitat de treball que li permet de trobar sempre un moment per a tothom, per les llargues discussions científiques. Per haverme heretat una determinada manera d'entendre la ciència. Per haver dirigit amb la seva batuta de manera encomiable el grup del cromosoma 21. Només em resta el dubte de saber si Mozart hagués dirigit el seu rèquiem amb tanta elegància i precisió.

Al Dr. Manuel Palacín, pel seu pragmatisme, per la seva inacabable curiositat, perquè no sempre es fàcil de trobar algú de qui tots els seus estudiants de doctorat et parlin bé, i de qui m'hagués agradat ser alumne seu.

Al Dr. Michael Lovett, per haver-me acollit en el seu laboratori de Dallas i haver-me mostrat la llum del final del túnel. A la Teresa Gallardo i en Andrew Simmons per fer-me sentir com a casa durant l'aventura americana. A ells i al Richard Del Maestro, la Jennifer Ashley, la Carol Venter, en Gregory Clines per ensenyar-me tots els ets i uts, més els secrets de la *cDNA selection*, i de la construcció d'una bona genoteca de cDNA.

Al Dr. Salvador Martínez, amb qui he pogut compartir força hores en el laboratori, a qui li tinc una estima i una admiració especial, per la seva capacitat i imaginació, per tot allò que m'ha ensenyat i ajudat. Al Dr. Luís Puelles, per acollir-me en el seu Departament d'anatomia de Múrcia, i a tota la gent del laboratori per la seva hospitalitat, i per haver-me ensenyat l'exquisida cuina de Múrcia, que plau d'una manera singular per la seva delicadesa, finor i excel·lència.

Al Dr. Francisco Tejedor, per la seva amistat, per totes les discussions científiques sobre mnb i MNBH i per ésser l'autor de l'article que més centenars de vegades he llegit.

A la Dra. Anna Planas, a qui també li tinc un afecte especial, pel seu acolliment, ajut, simpatia, capacitat de treball, decisió i valentia. I que espero que ja m'hagi perdonat els plats trencats.

Al Dr. Bob Williamson, per haver-me deixat treballar en el seu Departament de Genètica a Londres, pels seus consells i ajut... A tot el seu grup d'investigació per l'afecte i dedicació personal que em van dedicar, i en especial a la Sahar i a en Jaime.

A la Dra. Montse Milà, per la seva simpatia i ganes de fer ciència.

A la Dra. Ana Palacio i al Dr. Manuel Tamparillas, per llur valentia, i en especial per la seva hospitalitat quant vaig ser a Saragossa.

A la Caty Casas, per la seva amistat sincera, per saber escoltar, per la feina compartida, pel seu esmerç científic, per compartir hores intempestives al laboratori, per la seva sensibilitat. I amb qui sempre estaré en deute...

A la Mara Dierssen, per irradiar la melodia del foc de l'entusiasme en totes direccions, perquè li volem agrair tot el seu ajut, consells personals i afecte que ens ha mostrat, en especial a en Felip. I per ser algú a qui, personalment i científicament, m'agradaria semblar-me.

A l'Assumpció Bosch, per haver-me ensenyat l'abc de la genètica, i perquè el grup de recerca del cromosoma 21 és en gran part herència de la seva vàlua, valentia i gran esforç, en iniciar la singladura d'aquest grup.

A en Miguel Chillón, pels seus incommensurables consells sobre genètica, i infinites hores dedicades a ensenyar-me genètica, per compartir les llargues i intempestives hores de nit, dissabtes i diumenges sencers al laboratori. Pel seu bon humor, capacitat i honestedat. Per les seves bones idees i l'esperit inquiet. A la Julia Calonge per la seva amistat i recolzament, per tots i cadascun dels bons moments viscuts, pel seu esperit tant inquiet com crític, per totes les llargues hores de nit i caps de setmana fent ciència. Per la seva herència i per tot el que m'ha ensenyat i he après d'ella.

A en Toni Barrientos per ser un treballador infatigable, perquè no és fàcil compartir la feina amb algú de la seva talla científica i cultural. Per ser un amic a qui respecto i admiro profundament.

A en Juanjo Fuentes, pels incommensurables consells i favors, pel seu ajut incondicional, per ser una extraordinària persona, per la seva amistat sincera i lleial, i per l'interès cultural que contagia.

A en Mick Linch, per saber fer bona ciència, per les seves brillants idees. Pel seu ajut amb el meu anglès, per les seves traduccions spanglish-anglès. Per la seva feina silenciosa i pels seus ingents esforços malaurats de domar-nos. Per ensenyar-nos a tots, que en ciència hi ha dos camins: el ràpid i el bo.

A la Belén Peral, per ser una excel·lent persona, per la seva amistat i estima sincera, i perquè malgrat totes les roses que he regalat a les noies del laboratori, ha estat una noia de Madrid qui millor ha entès l'essència de la diada de Sant Jordi.

A la Mercè Pérez, per ser una bona amiga a qui li reconec moltes virtuts, per saber donar sense esperar propina, pel seu gran interès per la ciència, i per damunt de tot, per la gran lliçó que ens ha donat a tots... Voldria dedicar aquesta tesi també a la seva filla Laia.

A la Consol Bancells, per la seva extrema eficàcia, sensibilitat i ajuda amb la confecció d'aquesta memòria. Per la gran tendror que ens ha dedicat. I perquè darrera un gran home sempre hi ha una dona extraordinària.

A l'Esther Villén, por esos ojos negros, i a qui li desitjo molta sortI

A en Javi Jiménez, per la seva lleialtat, honestedat i amistat sincera, i per haver-se guanyat l'estima d'en Felip. Talment a la Marta Oset.

A en Carles Pucharcós i en Marià Tarragó, pels seus ajuts amb la informàtica.

A la Vassiliki Fotaki, per la seva vàlua, per la seva àmplia cultura, pel seu esperit rebel i entremaliat, i per ser una amiga a qui li reso també una especial admiració.

A la Nònia Feliubadaló, per la seva espontaneïtat i el seu somriure amable, i perquè hi ha quelcom de meravellós i sublim en la seva manera d'expressar-se, que va més enllà de la bellesa i de l'originalitat.

A en Jesús Purroy, per la seva inquietud cultural i a qui també li reconec moltes virtuts.

A tota la tripulació del grup del cromosoma 21: la Caty Casas, en Carles Pucharcos, l'Anna Domènech, l'Assun Solans, la Vassiliki Fotaki, en Juanjo Fuentes, la Marga Nadal, l'Anna Puig, la Lali Genescà i en Xavi Altafaj per les xerrades de sobretaula i per la seva complicitat, per haver solcat junts mars de treball, oceans d'esforços. Especialment als primers quatre, per haver respirat sang, ferro i pols en la ingent construcció del mapa de transcripció; per compartir moments bons i no tan bons.

A la Cris Ramos, per la seva extrema eficàcia i interès que m'ha dedicat.

A Rosa, per la burocràcia i els temes administratius, i per la seva confiança. Talment a la Gloria

Als meus amics per creure en mi, per fer-me més agradable els dies, i per fer que els senti sempre al meu costat.

A en Víctor Volpini, pels seus aclariments teòric-científics, i per ensenyar-nos que el saber no ocupa lloc.

A la Teresa Casals i a la Núria Sala pel seu afecte. A la Virginia Nunes per defensar la figura del becari.

A l'Anna Ruiz, la Loli Ramos, la Sara Larriba i l'Anna Aviñó per llur estima i amistat.

A la Creu Roja, per haver-me deixat gaudir de l'afecte i de la companyia de persones amb la síndrome de Down, i que ha significat, entre totes les experiències com a voluntari, una de les més entranyables.

I a tot un seguit de companys de viatge, en especial a la l'Anna Puig, l'Antonia Gaona, en Benjamin Rodríguez, la Conxi Lázaro, la Cristina Ramos, en David Otero, l'Eduard Serra, l'Elisabet Ars, l'Encarna Ulloa, l'Helena Kruyer, la Isabel Banchs, en Jordi Corral, en Jose Manuel Soria, la Judith García, la Laura Carim, la Lidia Carmona, la Margarida Nadal, la Mariona Font, la Marta Morell, la Mercè Miranda, la Meritxell Carrió, en Miguel Angel Pujana, la Mònica Gratacòs, la Montse Gómez, la Núria Lopez, la Núria Morral, en Rafael De Cid, la Raquel Rabionet, la Roser Llevadot, la Ruth Forner, la Susana Puig, en Toni Matilla, en Xavier Muñoz i la Yolanda Espinosa.

A la Mercè, la Mara, en Xavier E, la Belén i la Mariona per l'exhaustiva lectura crítica i constructiva d'aquesta tesi. Talment a la Caty, en Jesus, l'Anna A, la Consol i la Nònia.

A la Cristina Fillat, la Mara Dierssen, la Mercè Pèrez i la Mariona Arbonès pel tarannà de postdoc resat i per la nova sàvia insuflada al Centre Mèdic de Genètica Molecular; pels ajuts i consells personals, i per ensenyar-nos que quan la intel·ligència és discreta, és doblement brillant. I a les incorporacions darreres de la Mònica Escarceller, en Lauro Sumoy, en Josep M Arán i la Susana De La Luna.

Al Dr. Jaume Bertranpetit, per les seves magistrals classes de genètica humana, per fer-nos veure l'extraordinària complexitat d'allò que creiem que era nimi, per explicar-nos de manera senzilla, allò que realment és complex.

A la Dra. Elvira Joan, per les seves magistrals classes de genètica i per esperonar-nos un agut sentit crític (prou que ens va fer suar). A tots dos, Jaume i Elvira, perquè amb la perspectiva que atorga el temps, són les seves magistrals classes els records acadèmics més grats que hom té del pas per la Facultat de Biologia.

A tota la família Coch, pels seus ajuts i inquietuds despertats, i en especial a la Matilde, per la seva estima.

A tota la família Cuevas pel present, especialment a la Pepita, pel seu afecte i per les seves incommensurables hores amb l'incombustible Felip.

A l'Institut de Recerca Oncològica, pel seu suport econòmic per dur a terme la meva tesi. Talment a la Fundació Catalana Síndrome de Down/Marató de TV3-1993, Institut Català de la Salut/Generalitat de Catalunya, Fondo de Investigaciones Sanitarias de la Seguridad Social (FIS 94/1905-E), Dirección General de Enseñanza Superior (DGES/PM95-0106-CO2-01), Unió Europea. CEC/BIOMED1-Biomed2. Human Genome Analysis (GENE CT93-0015, GENE CT93 0037,GENE CT93-0107, GENE CT96-0554).

I per últim, no menys important, a la meva família....

A la meva mare Mercè... és potser, la bogeria que afecta les mares durant el període de criança i que ens fa veure allò que ningú no veu, i que ens fa sentir allò que ningú no sent, que ens les fa veure sempre sorpreses i admirades. Voldria dedicar-li aquesta tesi per tot el que en mi ha representat, perquè sempre és dins meu, i perquè continua constituint per a mi, de manera renovada, un exemple a seguir.

Als meus avis Pere i Andrea, pel seu amor i ajut. Al meu tiet Josep, per la seva confiança i amistat, i perquè sense el seu sacrifici personal no hagués tingut l'oportunitat d'escriure aquesta tesi. I encara que sàpiga que mai els hi podré agrair prou tot el gran esforç que han fet per mi, la seva presència la sento sempre molt a prop.

...i a en Felip, que sense saber què carai és un gen, i malgrat ha fet tot el possible perquè no escrivís aquesta tesi, sense el seu somriure, tot plegat hagués estat molt diferent.

...sí, sí, també tindré paraules d'agraïment per a tú, que nu de tota ànima, desproveït de tota essència i despullat d'esperit, et vesteixes amb llàgrimes alienes.

II. PRESENTACIÓ

Aquest treball està emmarcat dins la línia d'investigació sobre el cromosoma 21 humà i la síndrome de Down que el Centre de Genètica Mèdica i Molecular duu a terme i de la qual aquesta tesi en forma part.

La síndrome de Down s'ha d'estudiar com una patologia genètica complexa, potser la més complexa de les que es coneixen en humans. Unes paraules de Burton L. Shapiro són una bona mostra de la dimensió genètica de la síndrome de Down.

Chromosome 21 loci may have more or less of an effect on development of different traits and may have pleiotropic effects.

Without exception, traits that characterize Down syndrome are complex; they should be viewed and analyzed accordingly. I would urge those interested in DS to recognize the complexity of development of the signs that permit its diagnosis. To do otherwise would be a return to "beanbag genetics", the misleading tendency of early Mendelians to equate genes and characters, as if there were a one-to-one relation with no interaction among them.

Burton L. Shapiro Whither Down syndrome critical regions? 1997. Hum Genet., 99: 421-423.

Aquesta tesi està estructurada en set parts ben diferenciades. En la primera tracto els aspectes característics i generals de la síndrome de Down. En la segona presento els objectius que van motivar el treball d'aquesta tesi. En la tercera part descric detalladament els materials i mètodes emprats, a fi i efecte de permetre la realització de cadascuna de les tècniques presentades i la reproducció dels resultats mostrats. En la quarta recopilo els resultats dels treballs presentats en forma de publicacions. En la cinquena part presento un resum dels resultats i la seva discussió i es dóna pas a l'apartat de les conclusions del treball científic. La setena i darrera part de la memòria presentada conté la informació bibliogràfica referenciada.

He intentat, de manera més o menys encertada, trobar aquella traducció al català més correcta de cadascun dels mots tècnics propis del llenguatge científic, ja que en el procés d'incorporació de terminologia al català s'han produït força vacil·lacions. Tanmateix, s'ha conservat la denominació original en cursiva. L'única finalitat d'aquest manlleu lèxic (bàsicament llatinismes i anglicismes) ha estat la de facilitar la comprensió del text.

Tenint com a horitzó aquest mateix objectiu i per tal d'establir una regularitat i una coherència entre la presentació de les sigles acronímiques que apareixen en els articles publicats en anglès i aquelles que apareixen en els diferents apartats de la tesi, he preferit conservar l'ordre de les inicials de les sigles emmanllevades. D'aquesta manera, referint-se a la síndrome de Down trobarem DS i no SD.

Finalment, he volgut no duplicar la descripció dels resultats en els diferents apartats d'aquesta memòria, de manera que en l'apartat de resum i discussió he posat l'accent en aquells resultats que no s'han escaigut en l'apartat de resultats dels articles publicats, però que per la seva importància, per facilitar i ampliar la comprensió dels resultats presentats, he cregut adequat presentar-los.

III. LLISTA D'ABREVIATURES I SIGLES

A	Adenina
A600 i A800	Absorbància a una longitud d'ona de 600 i 800 nm respectivament
Ab	Oligopètid beta-amiloide
AD	Malaltia d'Alzheimer (<u>Alzheimer D</u> isease)
AFP	Alfa-fetoproteïna
AMP	Adenosina monofosfat
APOE	Apolipoproteïna E
APP	Gen precursor de la proteïna beta-amiloide (Amyloid Precursor Protein)
APS	Persulfat d'amoni
ATP	Adenosina trifosfat
BAC	Cromosoma artificial de bacteri (Bacterial Artificial Chromosome)
BLAST	Programa informàtic per a la recerca de similituts de seqüència (Basic
	Local Alignment Search Tool)
BSA	Albúmina de sèrum de vedella (<i>Bovine Serum Albumin</i>)
C	Citosina
cDNA	Àcid desoxiribonucleic complementari
СЕРН	Centre de les famílies de referència per al genotipagte dels microsatèl lits
	(<u>C</u> entre d' <u>É</u> tude du <u>P</u> olymorphism <u>H</u> umain)
Ci	Curie
сM	Centimorgan
$C_0 t_1/2$	És el producte de la concentració dels àcids nucleics (mols de nucleòtid per
	litre) i el temps (segons)
cpm	Comptes per minut
cR	Centiray
CVS	Vellositats corials
Da	Dàlton
dATP	Deoxiadenina trifosfat
DEPC	Dietil pirocarbonat
DMSO	Dimetil sulfòxid
DNA	Àcid desoxiribonucleic
DO	Densitat òptica
DO ₂₆₀ i DO ₂₈₀	Densitat òptica llegida a una longitud d'ona 260 i 280 nm respectivament
DS	Síndrome de Down
dscDNA	Doble cadena de cDNA (<u>D</u> ouble <u>S</u> trand <u>c</u> DNA)
DSCR	Regió crítica de la síndrome de Down (<u>D</u> own <u>S</u> yndrome <u>C</u> ritical <u>R</u> egion)
DSCR1	Gen humà "Down Syndrome Candidate Region 1"
DTT	Ditiotreitol
Dyrk	Gen de rata " <u>D</u> ual <u>Y</u> east <u>R</u> elated <u>K</u> inase"
EDTA	Àcid etilendiamina tetra acètic

VIII

EST	Fragment d'una sequència codificant (<u>Expressed Sequence Tag</u>)	
FAD	l'orma famillar de la malaltia d'Alzheimer (Familiar Alzheimer Disease)	
FISH	Hibridació in situ fluorescent (<u>l'luorescence In S</u> itu <u>Hybridisation</u>)	
S	Força de la gravetat	
a	Guanina	
GDB	Bane de dades del genoma (<u>G</u> enome <u>D</u> ata <u>B</u> ase)	
11C21	Cromosoma 21 humà (<u>IIuman Chromosome 21</u>)	
hnRNA	RNA nuclear inmadur (<u>IIctcrogeneous Nuclear RNA</u>)	
HSA21	Cromosoma 21 humà	
HTF	Illes IITF (<u>H</u> pall <u>T</u> iny <u>Fragments islands</u>) o illes CpG	
IFTG	Isopropil beta-D-tiogalactopiranòsid	
KDa	Kilodàlton	
LINE	Sequències de DNA repetitiu (Long Interspersed Nuleotide Elements)	
Mb	Megabases	
MCS	Lloc múltiple de clonatge (<u>Multiple Cloning Site</u>)	
MED-1	Seqüència de DNA que promociona múltiples inicis de transcripció	
	(<u>M</u> ultiple start site <u>Element D</u> ownstream)	
MMU16	Cromosoma 16 de ratolí	
mnb	Gen minibrain de Drosophila	
Mnbh	Gen <i>minibrain</i> múrid	
MNBH	Gen <i>minibrain</i> humà	
mRNA	Àcid ribonucleic missatger	
MS-AFP	Alfa-fetoproteïna del sèrum matern	
nt	Nucleòtid	
ORF	Pauta de lectura oberta (<u>Open Reading Frame</u>)	
PAC	Cromosoma artificial del bacteriòfag P1 (P1 Artificial Chromosome)	
pb	Parell de bases	
PBA	Proteïna beta-amiloide	
PCR	Reacció en cadena de la polimerasa (Polimerase Chain Reaction)	
PFGE	Electroforesi en gel de camps polsants (Pulsed Field Gel Electrophoresis)	
PIC	Contingut informatiu del polimorfisme (Polymorphism Information	
	Content)	
RACE	Tècnica per amplificar els extrems dels cDNAs (Rapid Amplification of	
	<u>c</u> DNA <u>E</u> nds)	
RFLP	Fragment de restricció de longitud polimòrfica (Restriction Fragment	
	Length Polymorphisms)	
RNA	Àcid ribonucleic	
rpm	Revolucions per minut	
RT-PCR	Reacció de la PCR a partir de cDNA generat mitjançant la reacció de la	
	transcripció inversa	

SDS	Dodecilsulfat de sodi	
SINE	Sequència de DNA repetitiu (Short Interspersed Repetitive DNA Element)	
sscDNA	cDNA de cadena senzilla (<u>Single Strand cDNA</u>)	
STR	Repetició curta en tàndem (Small Tandem Repeat)	
STS	Qualsevol seqüència anònima de DNA curta i única en el genoma que pot ésser amplificada específicament per la PCR (<u>Sequence Tagged Site</u>)	
Т	Timina	
Taq	DNA polimerasa de <u>Termophilus aq</u> uaticus	
TBE	Tampó de tris, borat i EDTA	
TE	Tampó de tris-EDTA	
TEMED	Tetrametil etilendiamina	
Ts16	Ratolí amb trisomia del cromosoma 16	
Ts65Dn	Ratolí amb trisomia parcial del cromosoma 16	
U	Unitats	
UDG	Uracil DNA glucosilasa	
ufc	Unitat formadora de colònia	
ufp	Unitat formadora de placa	
VNTR Repeticions en tàndem de nombre variable entre 9 i 64 nt (Var		
	Numer of Tamdem Repeats)	
YAC	Cromosoma artificial de llevat (Yeast <u>A</u> rtificial <u>C</u> hromosome)	
3'-UTR	3' de la regió no traduïda del cDNA (<u>UnTranslated R</u> egion)	
5'-UTR	5' de la regió no traduïda del cDNA (<u>UnTranslated Region</u>)	

IV.	ÍNDEX
	III D LA

1.	INTRODUCCIÓ	1
	1.1. SÍNDROME DE DOWN	3
	1.1.1. Antecedents històrics	3
	1.1.2. Trets fenotípics de la síndrome de Down	4
	1.1.2.1. Característiques generals	4
	1.1.2.2. Alteracions neurològiques	6
	1.1.2.2.1. Alteracions morfològiques	6
	1.1.2.2.2. Alteracions neuroquímiques	7
	1.1.2.2.3. Alteracions neurofisiològiques	7
	1.1.2.2.4. Similitud amb la patologia d'Alzheimer	7
	1.1.3. Classificació citogenètica de la síndrome de Down	9
	1.1.4. Síndrome de Down i edat materna	10
	1.1.5. Origen del cromosoma extra en la trisomia 21	11
	1.1.5.1. Trisomies lliures	11
	1.1.5.2. Trisomia per translocació del cromosoma 21	13
	1.2. CROMOSOMA 21 HUMÀ	14
	1.2.1. Cromosoma 21	14
	1.2.2. Marcadors polimòrfics del tipus microsatèl·lit	15
	1.2.3. Mapes genòmics	16
	1.2.3.1. Mapa de punts de trencament del cromosoma	16
	1.2.3.2. Mapa de fragments llargs de restricció	17
	1.2.3.3. Mapa de lligament genètic	17
	1.2.3.4. Mapa d'híbrids de radiació	18
	1.2.3.5. Mapa de clons solapats	19
	1.2.3.6. Mapa de transcripció	19
	1.2.4. Gens i malalties localitzades en el cromosoma 21	20
	1.2.5. Regió crítica de la síndrome de Down	22
	1.2.6. Models múrids de la síndrome de Down	24
	1.2.6.1. Ratolins trisòmic	25
	1.2.6.1.1. Ratolí amb trisomia completa del cromosoma MMU16	25
	1.2.6.1.2. Ratolí amb trisomia parcial del cromosoma MMU16	25
	1.2.6.1.3. Ratolí amb trisomía parcial del cromosoma MMU16 (Ts1Cje)	26

	1.2.6.2. Ratolins transgènics	27
	1.2.6.2.1. Ratolins transgènics amb YACs	27
	1.2.6.2.2. Ratolins transgènics d'un determinat gen	28
	1.2.6.2.2.1. Ratolí transgènic del gen APP	28
	1.2.6.2.2.2. Ratolí transgènic del gen CuZnSOD	28
	1.2.6.2.2.3. Ratolí transgènic del gen Ets-2	29
	1.2.6.2.2.4. Ratolí transgènic del gen PFKL	29
	1.2.6.2.2.5. Ratolí transgènic del gen HMG14	29
	1.2.6.2.2.6. Ratolí transgènic del gen S100B	29
	1.3. DIAGNÒSTIC PRENATAL I CONSELL GENÈTIC	30
	1.3.1. Diagnòstic prenatal	30
	1.3.2. Risc de recurrència de la síndrome de Down	32
	1.4. AÏLLAMENT DE SEQÜÈNCIES TRANSCRIPCIONALS	32
	1.4.1. Consideracions generals	32
	1.4.2. Tècnica de la captura d'exons o exon trapping	34
	1.4.3. Identificació d'illes CpG o CpG islands	36
	1.4.4. Tècnica de la PCR d'Alu-Splice	37
	1.4.5. Biocomputació de regions seqüenciades	38
	1.4.6. Tècnica de la selecció de cDNA o cDNA selection	39
	1.5. GEN MINIBRAIN	40
	1.5.1. Fenotip minibrain en Drosophila	41
	1.5.2. El gen Dyrk de rata (Mnbh)	43
	1.6. GENS, FUNCIÓ I FENOTIP	44
2.	OBJECTIUS	45
3.	MATERIAL I MÈTODES	49
	3.1. OBTENCIÓ D'ÀCIDS NUCLEICS	51
	3.1.1. Precipitació d'àcids nucleics	51
	3.1.1.1. Precipitació amb etanol	51
	3.1.1.2. Precipitació amb isopropanol	51
	3.1.2. Preparació de DNA	52
	3.1.2.1. Extracció de DNA genòmic a partir de cèl·lules sanguínies	52
	3.1.2.2. Extracció de DNA a partir de gels	53

3.1.2.2.1. Gels d'agarosa	53
3.1.2.2.2. Gels d'acrilamida	55
3.1.2.3. Obtenció de DNA plasmídic a partir de bactèris	55
3.1.2.3.1. Minipreparació	55
3.1.2.3.2. Maxipreparació	57
3.1.2.4. Minipreparació del DNA de llevat	58
3.1.3. Preparació de RNA	59
3.1.3.1. Extracció de RNA total a partir de teixit	59
3.1.3.2. Determinació de la quantitat de RNA	60
3.1.3.3. Extracció del DNA contaminant de la preparació de RNA	61
3.2. PURIFICACIÓ DELS ÀCIDS NUCLEICS	62
3.2.1. Ultracentrifugació en gradient de CsCl	62
3.2.1.1. Gradient de CsCl/bromur d'etidi	62
3.2.1.2. Extracció del bromur d'etidi	62
3.2.2. Purificació d'àcids nucleics mitjançant fenol	63
3.3. MANIPULACIÓ DELS ÀCIDS NUCLEICS	63
3.3.1. Transcripció inversa seguida de PCR (RT-PCR)	63
3.3.2. Hibridació in situ fluorescent de metafases	64
3.3.3. Introducció de DNA plasmidi a cèl·lules	65
3.3.3.1. Mètode del CaCl ₂	65
3.3.3.2. Mètode del PEG 3500	66
3.3.4. Reacció de desfosforilació del DNA	67
3.3.5. Transferència de DNA a una membrana de niló	67
3.3.5.1. Dot blot	68
3.3.5.2. Membrana de cosmidis d'alta densitat	68
3.3.5.3. Transferència de Southern	68
3.3.6. Hibridació de membranes	69
3.3.6.1. Hibridació de membranes	69
3.3.6.2. Extracció de la sonda de la membrana després de la hibridació	70
3.3.7. Marcatge del DNA	70
3.3.7.1. Marcatge del DNA mitjançant oligonucleòtids a l'atzar	70
3.3.7.2. Marcatge de DNA per PCR	72
3.3.7.3. Sondes d'oligonucleòtids	72

3.3.7.4. Marcatge terminal del DNA 72

3.3.7.5. Oligonucleòtids com a encebadors per al genotipatge	73
3.3.8. Purificació d'una sonda de DNA	74
3.3.8.1. Per a sondes >50 pb	74
3.3.8.2. Per a sondes <50 pb	74
3.3.9. Determinació de l'activitat específica i la quantitat	de cpm
de la sonda	75
3.3.10. Fosforilació de l'extrem 5' d'àcids nucleics	75
3.3.11. Reacció de lligament del DNA	75
3.3.12. Anàlisi de RNA per transferència de Northern	77
3.3.12.1. Electroforesi en gel d'agarosa/formaldehid	77
3.3.12.2. Transferència de l'RNA del gel a la membrana	77
3.3.13. Construcció del T-vector	78
3.3.14. DNA glicosilasa d'uracil	79
3.3.15. Bioinformàtica	80
3.3.15.1. Editors de seqüència	80
3.3.15.2. Prediccions de l'estructura del DNA i dels dominis d	te les
proteïnes	80
3.3.15.3. Disseny d'oligonucleòtids	80
3.3.15.4. Identificació de regions codificants de proteïnes	80
3.3.15.5. Recerca de similituds i identitats	81
3.3.15.5.1. BLAST	81
3.3.15.5.2. FASTA	81
3.3.15.6. Recerca d'informació	81
3.3.15.6.1. Entrez	81
3.3.15.6.2. MEDLINE	81
3.3.16. Disseny manual d'encebadors	81
3.4. MAPA DE COSMIDIS	82
3.4.1. Construcció de genoteques de cosmidis	82
3.4.1.1. Anàlisi dels YAC	82
3.4.1.1.1. Anàlisi de la mida i estabilitat	83
3.4.1.1.2. Anàlisi de quimerisme	83
3.4.1.1.2.1. Generació de productes ALU-PCR a partir dels YA	AC per a
estudis de FISH	83
3.4.1.1.2.2. PCR de la bombolla	84

3.4.1.2. Obtenció de DNA encapsulat en boletes d'agarosa	85
3.4.1.3. Recuperació del DNA de les boletes d'agarosa	86
3.4.1.4. Fraccionament parcial del DNA de YAC	87
3.4.1.5. Desfosforilació de la digestió parcial del DNA del	
llevat	88
3.4.1.6. LLigament del DNA fraccionat del YAC als cosmidis	89
3.4.1.6.1. Digestió del cosmidi amb XbaI	89
3.4.1.6.2. Desfosforilació del cosmidi digerit amb XbaI	89
3.4.1.6.3. Segona digestió amb BamHI	89
3.4.1.6.4. Lligament del DNA fraccionat al cosmidi Supercos	
doblement digerit	90
3.4.1.7. Empaquetament del cDNA fraccionat del YAC	91
3.4.1.7.1. Preparació de les cèl·lules hostes	91
3.4.1.7.2. Empaquetament dels cosmidis	91
3.4.1.7.3. Infecció de les cèl·lules hostes	91
3.4.1.7.4. Primera selecció d'aquells clons amb DNA del YAC	92
3.4.1.7.5. Segona selecció d'aquells clons amb DNA del YAC	92
3.4.2. Construcció d'un mapa contigu de cosmidis	93
3.4.2.1. Localització pel contingut de STS	93
3.4.2.2. Mapa de restricció	93
3.4.2.3. Ribosondes	94
3.4.2.4. Sondes per PCR lineal	95
3.5. TÈCNICA DE LA SELECCIÓ DE cDNA	95
3.5.1. Preparació de RNA poli(A) ⁺	95
3.5.2. Preparació de cDNA a partir de RNA poli(A) ⁺	96
3.5.2.1. Conversió de mRNA a dscDNA d'extrems roms	97
3.5.2.1.1. Síntesi de cDNA	97
3.5.2.1.2. Conversió del sscDNA a dscDNA	98
3.5.3. Obtenció de dscDNA amb adaptadors	99
3.5.3.1. Creació de l'adaptador de doble cadena	99
3.5.3.2. Unió de l'adaptador de cadena doble o linker al dscDNA	100
3.5.4. Marcatge dels cosmidis amb biotina	101
3.5.5. Tècnica de la selecció de cDNA	102
3.5.5.1. Amplificació de la genoteca de cDNA	102

	3.5.5.2. Supressió de les seqüències repetitives del cDNA primari	102
	3.5.5.3. Primer cicle de selecció de cDNA	103
	3.5.5.4. Amplificació de la primera selecció de cDNA	104
	3.5.5.5. Segon cicle de selecció	105
	3.5.5.6. Avaluació de l'enriquiment	105
	3.5.5.7. Clonatge i caracterització dels cDNA seleccionats	106
	3.6. MAPA TRANSCRIPCIONAL	106
	3.6.1. Determinació de l'origen de cada EST als YAC de l'HC21	106
	3.6.2. Determinació de l'origen de cada EST als cosmidis	107
	3.7. CARACTERITZACIÓ D'UN NOU GEN HUMÀ: (MNBH)	108
	3.7.1. Aïllament de la sequència codificant	108
	3.7.1.1. Identificació de l'EST de MNBH	108
	3.7.1.2. Clivellatge de genoteques recombinants de cDNA	108
	3.7.1.2.1. Primer clivellatge	108
	3.7.1.2.2. Segon clivellatge	110
	3.7.1.3. Genoteques de cDNA	110
	3.7.1.4. Seqüenciació del DNA dels nous clons de MNBH	111
	3.7.2. Aïllament de la seqüència sencera del cDNA	112
	3.7.2.1. Aïllament dels extrems 5' i 3'-UTR	112
	3.7.2.1.1. RACE de cDNA	112
	3.7.2.1.2. RACE Marathon	114
	3.7.2.1.3. Genoteques de cDNA	114
	3.7.3. Hibridació in situ	115
	3.8. ANÀLISI DELS MICROSATÈL·LITS	116
	3.8.1. Anàlisi dels microsatèl·lits del tipus (CA)n	116
	3.8.7.1. Determinació de l'heterozigositat dels microsatèl·lits	116
	3.8.7.2. Genotipatge o determinació al·lèlica	117
	3.8.2. Construcció de mapes genètics	119
4.	RESULTATS	121
	4.1. RESULTATS DIRECTES DE LA TESI	123
	4.1.1. Cosmid contig and transcriptional map of three regions of human	
	chromosome 21q22: identification of 37 novel transcripts by direct selection	127
	4.1.2. A human homologue of Drosophila minibrain (MNB) is expressed	

-

	in the neuronal regions affected in Down syndrome and maps to the critical	
	region	139
	4.1.3. Minibrain (MNBH) is a single copy gene mapping to human	
	chromosome 21q22.2	147
	4.1.4. Human Minibrain homologue (MNBH): characterization alternative,	
	splicing, differential tissue expression, and overexpression in Down syndrome	153
	4.2. TREBALLS COL·LABORATIUS	185
	4.2.1. Two dinucleotide repeat polymorphisms at 21q22.3 (D21S416	
	and <i>D21S1235</i>)	187
	4.2.2. Identification of two highly polymorphic CA-repeats (D21S1224	
	and D21S1261) on human chromosome 21q22.3	191
	4.2.3. Five new microsatellite polymorphisms at the q21 region of human	
	chromosome 21	197
	4.2.4. Characterisation of three microsatellite polymorphisms (D21S1262,	
	D21S1419 and D21S1421) from band 21q22.1	203
	4.2.5. New alleles at microsatellite loci in CEPH families mainly arise from	
	somatic mutations in the lymphoblastoid cell lines	209
	4.2.6. The EUROGEM map of human chromosome 21	219
	4.2.7. Integration of 30 CA-repeat markers into the cytogenetic, genetic	
	and YAC maps of human chromosome 21	223
5.	RESUM I DISCUSSIÓ	233
	5.1. CONSTRUCCIÓ DE LA GENOTECA DE COSMIDIS	241
	5.2. SELECCIÓ DIRECTA DE cDNA DE L'HC21	245
	5.3. CONSTRUCCIÓ D'UN MAPA TRANSCRIPCIONAL	250
	5.4. CARACTERITZACIÓ D'UN NOU GEN HUMÀ: MNBH	253
6.	CONCLUSIONS	265

7. REFERÈNCIES

269

1. INTRODUCCIÓ

1

....

1.1. SÍNDROME DE DOWN

La síndrome de Down (DS) és l'alteració genètica humana més comuna: amb una incidència d'1/700 nounats (Epstein, 1995). És també la causa genètica coneguda més freqüent de discapacitat intel·lectual i d'alteracions cardíaques, i és considerada com un model per a l'estudi de les aneuploïdies humanes. Tenint en compte les interrupcions desitjades d'embarassos, i que no totes les concepcions amb trisomia 21 sobreviuen fins al naixement, el nombre estimat de fetus que presenten la DS respecte al total d'embarassos és d'1 de cada 150 (Freeman *et al.*, 1991).

1.1.1. ANTECEDENTS HISTÒRICS

L'evolució exponencial del coneixement que tenim de la DS va des de les primeres representacions de l'antiga cultura Olmeca d'individus amb característiques d'aquesta síndrome (1.000 a. C.), o bé les pintures de l'any 1505 trobades al tríptic de l'altar de la passió d'Aachen, en què es reprodueix el perfil d'un infant amb el fenotip característic de la DS, fins al coneixement molecular del nostre segle. Des de les teories etiopatogèniques precromosòmiques de la regressió a l'orangutan i d'altres com el retorn a un tipus de raça primitiva (mongols) o l'esgotament matern a causa de múltiples embarassos, fins a les modernes teories de la triple dosi gènica i del desequilibri en el nivell d'expressió d'aquests gens.

El primer esment mèdic de la DS, encara que poc acurat, correspon a J.E.D. Esquirol (1838). No fou fins a l'any 1866, que J. Langdon Down descriví acuradament els trets clínics d'aquesta síndrome i la considerà com una entitat amb personalitat pròpia i separada (Down, 1866). En honor seu, el seu cognom esdevindria posteriorment el nom de la síndrome. L'any 1956, Tjio i Levan començaren l'etapa cromosòmica de la DS, descrivint 46 cromosomes per als humans. L'aplicació de la citogenètica permeté a Lejeune identificar un cromosoma extra del grup G (l'autosoma més menut, anomenat 21) en cariotips de persones amb la DS (Lejeune, 1959; Jacobs *et al.*, 1959). D'aleshores ençà, la DS es coneix com a trisomia 21. L'etapa postcromosòmica començà amb l'aparició de les noves tècniques i coneixements de biologia molecular, com l'estudi dels polimorfismes de DNA que ha permès estudiar l'origen parental de l'error meiòtic responsable de la no disjunció cromosòmica (Warren *et al.*, 1987; Sherman *et al.*, 1994). El 1989, Warren *et al.* publicaren el primer mapa de lligament de marcadors de DNA del cromosoma 21, que més tard possibilitaria la construcció del primer mapa físic solapat de cromosomes artificials de llevat (YAC) (Chumakov *et al.*, 1992a; 1992b).

Actualment, la recerca biomèdica de la DS ha esdevingut multidisciplinària: la neurologia, la immunologia, la psicopedagogia, la clínica i la recerca genètica. Aquesta

darrera ha de permetre aïllar i identificar tots els gens que es localitzen al cromosoma 21, i així aprofundir en el coneixement molecular de la DS. La identificació dels diferents gens i de la funció de llurs proteïnes permetrà relacionar-los amb els trets fenotípics de la DS. És d'esperar que aquesta informació permeti pal·liar o curar alguns dels quadres clínics típics de la trisomia 21.

1.1.2. TRETS FENOTÍPICS DE LA SÍNDROME DE DOWN 1.1.2.1. CARACTERÍSTIQUES GENERALS

La primera observació destacable és l'alta incidència de mortalitat espontània de fetus amb la DS a l'úter, que és de l'ordre d'un 25-35% (Hook, 1978; Hook *et al*, 1995). Aquesta xifra de mortalitat és heterogènia segons l'estadi de gestació. Així, la taxa de mortalitat més gran que s'observa correspon als fetus de fins 17 setmanes, doblant la taxa de mortalitat d'aquells fetus que en tenen més.

El protocol per a la determinació clínica de la DS (Epstein *et al.*, 1991) depèn de la presència d'un nombre de trets que afecten diferents òrgans i teixits (Epstein, 1986) tot i que llur expressió i penetració és variable (Taula 1). Els nounats i infants tenen invariablement hipotonia i una discapacitat intel·lectual (Epstein, 1991; Korenberg *et al.*, 1994) que augmenta en la segona dècada de vida (Carr, 1994; Morgan *et al.*, 1979). Si la discapacitat intel·lectual pogués ésser quantificable en termes del coeficient d'intel·ligència, aquest decauria durant l'adolescència (Fig. 1). A més a més, s'observa una minvada capacitat de memòria, de llenguatge, de processament àudio-verbal i una progressiva pèrdua de les funcions cognitives (Mazzoni *et al.*, 1994; Wang i Bellugi, 1994). Els defectes cardíacs congènits estan presents en el 60% dels individus amb la DS i representen el 70% de tots els casos amb aquest tipus de defecte (Korenberg *et al.*, 1992). Durant la tercera dècada de vida s'observen canvis patològics i neuroquímics en el cervell semblants als de la malaltia d'Alzheimer (Williams i Matthyse, 1986; Nadel i Epstein, 1992).

Les persones amb la DS tenen un risc entre deu i vint vegades més elevat de desenvolupar leucèmia aguda respecte a les persones sense la DS (Evans, 1972). Les malformacions cardíaques congènites, juntament amb les infeccions i la leucèmia, són les causes de mort més comunes, i tot això condiciona que l'esperança de vida sigui reduïda (Thase, 1982; Fryers, 1986), tot i que ha millorat espectacularment de la meitat de segle ençà, que era de nou anys el 1929. Aquest increment és el resultat de la millora dels mitjans mèdics i socials a totes les edats de les persones amb la DS. Un factor clau en la determinació de la longevitat, sobretot pel que fa als primers anys de vida, és la presència o absència de problemes cardíacs. Es calcula que la població amb la DS sense aquests

Trets característics	Casos amb exploració positiva (%)
Caracteristiques fisiques	
Plecs palpebrals oblics	80
Plec cutani en coll i clatell	80
Paladar estret	- 75
Braquicefàlia	75
Hiperlaxitud	75
Absència de pont nasal	70
Adducció 1r dit del peu	70
Mans petites (curtes i amples)	65
Coll curt	60
Anomalies en la dentició	60
Epicant	60
Occipit pla, cara ampla, perfil camœs	90
Boca oberta	60
Problemes oculars (miopia, estrabisme,	55
cataractes, taques de Brushfield o iris a	<u>)</u>
clapes)	
Solc simiesc al palmell	55
Pavellons auditius dismòrfics	50
Protusió lingual	45
Dermatoglifs anormals	85
Anomalies esquelètiques	
Clinodactília (5è dit de la mà curt i corbat)	55
Displàsia de la pelvis en els nounats	70
Malformacions atles occipital	20
Discanacitat intel·lectual	100
(Profund CI<20, grey CI 20-50, moderat CI>50)	(10-70-20)
Hinotinoïdisme	7-17
Trota nounalàgica	1-17
Hinotonia	100
Alpoionia Lonta desenarició dels primers reflexes	80
Detala desaparicio dels primers renexos	100 dels indiv > de 35 anve
Patologia similar a la malaltia d'Alzheimer	100 dels maiv. > de 55 anys
Anomalies congenites	22.42
Detectes al cor (HD)	30-40
Persistència del ductus arterios	40 del total de HD
Comunicació interventricular	31 "
Comunicació interauricular	9 "
Tetralogia de Fallot	6 "
Anomalies gastrointestinals	
Atrèsia duodenal	2,5
Imperforació de l'anus	1
Megacòlon	0,6
Infertilitat en homes	100 (amb l'excepció d'un sol cas)
Leucèmia	1.1
Megacariocítica aguda	200 a 400 vegades augmentat
Linfoide-Mieloide	10 a 20 vegades augmentat
Reacció leucemoide (nounats)	0.1
Reactio reactionate (nounais)	0,1

Taula 1. Trets fenotípics dels individus amb la DS

6 Introducció

darrers problemes té una esperança de vida d'uns 10-20 anys menys que la població sense la DS, essent més gran la diferència per a les dones.



Figura 1. Mitjana i rang observat del coeficient d'intel·ligència (CI) en nens amb la DS en funció de l'edat. Dades de Morgan, 1979.

1.1.2.2. ALTERACIONS NEUROLÒGIQUES

Tal i com s'extreu de la taula 1, són molts els òrgans implicats en la DS, però el sistema nerviós central (SNC) pren una importància cabdal perquè és un dels organs més afectats, i perquè en la seva alteració rauen dues patologies clíniques de primera importància, no solament científica sinó també social com són: 1) la discapacitat intel·lectual observada en les persones amb la DS i 2) les alteracions neuropatològiques similars a les observades en la malaltia d'Alzheimer, present quasi en el 100% dels individus amb la DS més grans de 35 anys (Kesslak *et al.*, 1994; Korenberg, 1995), i una demència similar a l'Alzheimer a partir dels 40 anys, en el 30% dels individus amb la DS (Dalton *et al.*, 1993).

1.1.2.2.1. Alteracions morfològiques

El pes del cervell en la majoria dels nens amb la DS comença a minvar per sota l'interval normal als 3-6 mesos de néixer. S'observa també una reducció del diàmetre frontooccipital i una disminució de les mides de les circumvolucions temporals superiors. El cerebel és desproporcionadament més petit, així com el tronc cerebral, l'hipocamp i el cos callós (Schapiro *et al.*, 1989; Weis, 1991; Raz *et al.*, 1995). Durant la infantesa, també s'observa retard en els processos de mielinització de les fibres d'associació intercortical dels lòbuls fronto-temporals i alteracions de les propietats electrofisiològiques de la membrana de les neurones. S'observa també, de forma constant i generalitzada, una disminució (del 20 al 50%) de la densitat neuronal en totes les làmines de l'escorça cerebral, sobretot en les làmines II i IV (Wisniewski, 1990), especialment en les capes riques en cèl·lules granulares de l'hipocamp, del cerebel i del tàlem (Ross *et al.*, 1984).

Altres alteracions observades després de néixer són la pèrdua d'arboritzacions i d'espines dendrítiques, així com modificacions en l'estructura de les espines, que arriben a ésser relativament atròfiques als 4 mesos d'edat (Marín-Padilla, 1976; Suetsugu i Mehraein, 1980).

1.1.2.2.2. Alteracions neuroquímiques

La densitat i els patrons de distribució dels receptors colinèrgics del tipus muscarínic de l'escorça cerebral, l'hipocamp, i nuclis basals del prosencèfal en nounats amb la DS és normal, però està marcadament disminuïda en certs nuclis del tronc de l'encèfal, nuclis del col·licle i la substància negra del mesencèfal (Flórez *et al.*, 1990).

1.1.2.2.3. Alteracions neurofisiològiques

S'observen alteracions en la transmissió nerviosa interneuronal degudes a una reducció notable de la densitat sinàptica, sobretot en cadascuna de les unitats columnars de l'escorça cerebral, de l'hipocamp i del cerebel, que redueixen l'eficiència de la transmissió sinàptica (Wisniewski, 1990). Això motiva una pertorbació ens els processos d'entrada i sortida d'informació d'aquestes regions, que repercuteix greument sobre els processos d'aprenentatge com la memòria, l'atenció i l'activitat motora, entre d'altres.

1.1.2.2.4. Similitud amb la patologia de la malaltia d'Alzheimer

A partir dels 35-40 anys, el cervell de tota persona amb la DS mostra lesions patològiques similars a les de la malaltia degenerativa d'Alzheimer (Kesslak *et al.*, 1994; Korenberg *et al.*, 1994). Aquestes lesions poden comportar, en el pitjor dels casos, en un tipus de demència específica de la trisomia 21 (pèrdua progressiva de la funció intel·lectual, alteració greu de la memòria, reducció de la capacitat lingüística i de la comprensió) i similar a l'observada en la malaltia d'Alzheimer (Dalton *et al.*, 1993). En l'àmbit neuropatològic, la demència observada en els individus amb la DS es caracteritza per la degeneració preferent, però no selectiva, de les neurones colinèrgiques dels nuclis prosencefàlics basals, per la presència de dipòsits extracel·lulars de proteïna amiloide que formen les plaques neurítiques i la presència intracel·lular de filaments del citoesquelet,

8 Introducció

que s'organitzen en cabdells neurofibril·lars. Cap d'aquestes característiques és per se suficient per originar aquesta patologia; les seves bases moleculars, però, no són prou conegudes. En el braç llarg del cromosoma 21 s'ha localitzat el gen *APP* que codifica per a una proteïna anomenada proteïna precursora de beta-amiloide (B-APP). En el cas de la malaltia d'Alzheimer (AD), la proteïna B-APP és la responsable de la forma familiar d'Alzheimer pre-senil autosòmica dominant (FAD) en alguns casos molt puntuals de famílies portadores de mutacions en el gen *APP*. La fragmentació proteolítica diferencial d'aquesta proteïna precursora dóna lloc a isoformes d'una altra proteïna lliure i soluble d'uns 40-42 aminoàcids, anomenada proteïna beta-amiloide (PBA o Aß) (Glenner i Wong, 1984). Aquesta darrera proteïna és la component principal d'uns nuclis, entorn dels quals es formen les plaques neurítiques extracel·lulars que són masses de restes degenerats de prolongacions neuronals (dendrites i cilindreixos).

Encara que s'ha observat que els nivells de la proteïna B-APP al plasma i al cervell de les persones amb la DS són molt superiors respecte a les persones sense la síndrome (Rumble et al., 1989), és necessari que s'origini la PBA, que és l'element neurotòxic, per a desenvolupar la malaltia. Tot i així, la sola formació de les plaques neurítiques associades a PBA no és suficient per a desenvolupar la demència en la majoria dels casos, i hi cal també l'aparició dels cabdells neurofibril·lars (masses degeneratives intracel·lulars de filaments del citoesquelet que apareixen en certes neurones). D'aquesta manera, encara que la sobreexpressió del gen APP en les persones amb la DS pot intervenir en el desenvolupament de la demència, podria ésser necessària també l'activitat d'altres proteïnes, com les involucrades en el desenvolupament de la demència en pacients amb l'AD com: 1) la presenilina 1 (PS1 o S182, en el cromosoma 14) com a proteïna responsable de la majoria dels casos de la FAD (Sherrington et al., 1995), 2) la presenilina 2 (PS2 o STM2, en el cromosoma 1) com a responsable d'alguns casos de FAD, i 3) altres proteïnes que també podrien estar involucrades com l'apolipoproteïna E4 (APO E4) o la proteïna tau. Totes aquestes proteïnes poden interaccionar entre si, de manera que el gen APP de l'HC21 podria afavorir el desenvolupament de les plaques neurítiques i dels cabdells neurofibril·lars. Aquestes plaques i cabdells, juntament amb una pèrdua de neurones especialment important en l'hipocamp, en l'escorça cerebral i en els nuclis prosencefàlics basals (que és una font d'aferències colinèrgiques a l'hipocamp i escorca cerebral) constitueixen les modificacions neurològiques més importants en el desenvolupament progressiu de la forma de demència d'Alzheimer, i que també estan presents en els cervells de les persones més grans de 35 anys amb la DS.

Darrerament s'ha suggerit que tres còpies del gen APP podrien incrementar els nivells d'una de les isoformes neurotòxiques de la PBA (AB-42) i causar una patologia similar a l'AD en aquells individus amb la DS (Teller *et al.*, 1996). Dades a favor d'aquesta hipòtesi són aquells individus amb la DS que tenen una trisomia parcial del 21 [46, XX, rec(21)dup q, inv(21) (p12q22.1)] que exclou el gen *APP* i, malgrat que, atípicament, han arribat a una edat avançada, no tenen evidències de demència o de patologia similar a l'AD (Prasher *et al.*, 1998). Cal fer esment que aquest tipus de demència associada a la trisomia 21 es presenta fins a vint anys després de l'aparició de les plaques amiloides. Els mecanismes exactes de la degeneració neuronal en la DS encara no es coneixen bé, però s'ha trobat que les neurones corticals en cultiu de persones amb la DS pateixen apoptosi i acaben degenerant, mentre que les cèl·lules d'individus sense DS romanen viables (Busciglio i Yankner, 1995).

Tot plegat, i malgrat que el 100% dels individus amb la DS presenten alteracions neuropatològiques similars a les de la malaltia d'Alzheimer, no es pot dir que totes les persones amb la DS desenvoluparan indefectiblement una demència, ja que un percentatge variable de persones amb la DS atenyen la cinquena i sisena dècada de la seva vida sense presentar signes de demència (Mann, 1988). Nogensmenys, com a grup representen una població amb un risc més alt, ja que totes aquestes alteracions neurològiques els fan ésser més susceptibles, en la majoria dels casos, al desenvolupament progressiu d'aquesta patologia.

1.1.3. CITOGENÈTICA DE LA SÍNDROME DE DOWN

1) Trisomia regular (93-96%)

47, +21

2) Trisomia amb 46 cromosomes i dup(21q) (<5%).

Translocació robertsoniana del:

46, -14, + rb(14;21) 2,5% 46, -13, + rb(13;21) 0,8% 46, -15, + rb(15;21) 0,4% 46, -22, + rb(22;21) 0,2% 46, -21, +rb(21;21) 0,01%Translocació no balancejada:

p.e. 46, -14, +tr(14;21)

Isocromosoma i(21q):

dicèntric 50% dels casos i(21q)

monocèntric 50% dels casos i(21q)

3) Mosaics 47, +21/46 (<5%)

En la majoria dels casos, la trisomia 21 està causada per una no disjunció del cromosoma 21 en la línia germinal d'un dels dos pares, la qual cosa motiva l'existència d'un cromosoma 21 extra i lliure (93-96%). Les translocacions robertsonianes representen menys del 5%, i els casos de mosaïcisme entre línies cel·lulars amb trisomia 21 completa i línies cel·lulars euploides representen també menys del 5% aproximadament (Epstein, 1995). Els casos de mosaïcisme poden esdevenir-se tant d'un error meiòtic com d'un error mitòtic, i en tots dos casos es generen dues poblacions de cèl·lules 47,+21 i 45,-21. Aquestes darreres cèl·lules monosòmiques són inviables i es perden durant la vida fetal.

1.1.4. SÍNDROME DE DOWN I EDAT MATERNA

Hi ha una reconeguda associació entre la incidència de la trisomia lliure del cromosoma 21 humà (HC21) i l'edat de la mare (Hassold i Jacobs, 1984). L'edat materna és un factor determinant en la freqüència d'avortaments espontanis de fetus amb aneuploïdies i en la incidència de la DS en nounats, ja que a partir dels 40 anys molts dels oòcits de la mare són aneuploides (Hassold i Chiu, 1985). La Fig. 2 mostra l'increment de la incidència de la trisomia 21 en mares d'edats diferents. Fins a l'edat de 30 anys, el risc és d'1 per cada 1.000-1.500 nounats, a partir dels trenta anys el risc s'incrementa de manera ostensible, i les mares de 45 anys tenen un risc d'1 nounat amb la DS per cada 30 naixements.



Figura 2. Nombre de nounats amb la DS per cada 1.000 neixements, en relació a l'edat materna (Hook, 1982).

Cal esmentar que l'edat materna influeix en els errors de tipus meiòtic, però no en els mitòtics, que representen el 4,5% de les trisomies 21 (Antonarakis *et al.*, 1993). Un dels factor postulats per a explicar aquest fenomen és la reducció de la recombinació genètica en el braç 21q. Estudis de lligament genètic amb marcadors de DNA de l'HC21 en aquelles trisomies que s'esdevenen d'un error matern de la meiosi I (MMI), mostren que la taxa de recombinació sembla minvar en els cromosomes que no presenten disjunció (33 cM davant els 72 cM de la població control), i marcadament reduïda en la regió proximal (Sherman *et al.*, 1994). A més a més, les dades d'altres estudis indiquen també que el nombre de casos en què no es detecta recombinació en les trisomies d'origen MMI és excessiu, fins arribar a un 30% més (Yoon *et al.*, 1996). Aquesta manca de quiasmes es fa més palès en la regió proximal, suggerint que la presència només de la recombinació distal augmenta la probabilitat de la separació prematura dels cromosomes homòlegs en la metafase, perdent la seva unió al fus acromàtic i migrant a l'atzar (Hawley *et al.*, 1994). Cal esmentar que aquesta teoria no és universalment acceptada (Hamers *et al.*, 1990).

En el període prenatal, entre els mesos quart i sisè de vida fetal, els oòcits comencen el procés de la meiosi I (MI) i es detenen en l'estat de profase, en el període de diplotè. En aquest període, els quiasmes semblen ésser la força d'unió més important que mantenen units els bivalents fins la metafase. El procés de recombinació ja iniciat entre les cromàtides germanes de cromosomes homòlegs no es completa fins l'ovulaciò, i no és fins a la fecundació de l'oòcit, que el procés de la meiosi continua fins al final. D'aquesta manera, la trisomia 21 relacionada amb l'edat de la mare deguda a errors MMI pot resultar de la freqüència minvada de quiasmes en oòcits vells, ja que la possibilitat de separació prematura de les cromàtides bivalents en la metafase I és més gran com més temps trigui aquesta a completar-se.

1.1.5. ORIGEN DEL CROMOSOMA EXTRA EN LA TRISOMIA 21 1.1.5.1. TRISOMIES LLIURES

L'alta informativitat dels marcadors de DNA polimòrfics de l'HC21 ha permès determinar l'origen parental de l'HC21 extra en les trisomies lliures. A més a més, les anàlisis de segregació al·lèlica dels marcadors de DNA al llarg de l'HC21, incloent-hi els marcadors centromèrics i pericentromèrics (per a identificar els esdeveniments de recombinació) poden indicar en quin estadi de la meiosi s'ha produït l'error de segregació. La reducció a homozigositat d'un marcador pericentromèric és interpretada com a error de meiosi II (MII), la retenció d'heterozigositat per a un determinat marcador és interpretada com a resultat d'un error de meiosi I (MI). Per exemple, si la

mare té els al·lels AB per a un *locus*, els al·lels del pare són CD i la descendència amb trisomia 21 té els al·lels ABC per a un *locus*, l'error és assignat a la MI, mentre que si la descendència tingués els al·lels AAC per aquest marcador, l'origen de l'error és determinat a la MII. Finalment, l'homozigositat de tots els marcadors del braç 21q, incloent-hi els pericentromèrics, és interpretat com un error postzigòtic o mitòtic, ja que la seva assignació està basada en l'assumpció que la possibilitat de recombinació en la mitosi és negligible, ja que els entrecreuaments o *crossovers* mitòtics són molt rars.

Aquests estudis indiquen que els errors mitòtics són responsables del 4,5-5,3% dels casos de trisomia 21, que no estan associats amb una edat avançada dels pares, i que no hi ha una preferència en l'origen parental de l'HC21 en aquests tipus d'errors (Antonarakis *et al.*, 1993; Yoon *et al.*, 1996). Cal tenir en compte que els esdeveniments de doble recombinació entre distàncies genètiques petites poden no ésser detectats, i es pot assignar erròniament un error mitòtic com a meiòtic II. És important esmentar que, si l'error mitòtic va tenir lloc en etapes molt primerenques del zigot o de l'embrió amb poques cèl·lules, el resultat dels marcadors de DNA serà indistingible del cas que hagués estat un error MII.

Les dades d'aquests estudis mostren també que la resta de trisomies 21 provenen d'un error meiòtic, i que el seu origen és paternal en el 8,8% dels casos, mentre que en la majoria de trisomies degudes a un error meiòtic, l'origen és maternal (91,2%) (Taula 2) (Yoon *et al.*, 1996).

Aquestes proporcions són molt similars a les obtingudes per altres autors (Petersen et al., 1993). Aquest darrer estudi corrobora les dades dels errors meiòtics de procedència materna i dels mitòtics exposades anteriorment, però s'observa que els errors paterns de meiosi II (PMII) tenen una probabilitat més gran d'esdevenir que no pas els errors paterns de meiosi I (PMI). Aquesta discordança en les dades de l'origen meiòtic paternal podria ésser deguda al fet que els casos d'origen patern són minoritaris i la mostra de la població no és prou gran.

Les conclusions d'aquests estudis es mostren a continuació:

1) En la majoria dels casos, els errors meiòtics que originen la trisomia 21 són d'origen matern. Només el 8,8% dels casos corresponen a errors en l'espermatogènesi.

2) La majoria dels errors meiòtics materns tenen lloc durant la meiosi I (75,3%).

3) Els errors de meiosi II comprenen el 24,7% dels errors d'origen matern.

4) Del 8,8% dels errors que són d'origen patern, el 50% són de procedència de la meiosi I, i l'altre 50% ho són de la meiosi II.

5) Els errors mitòtics són responsables del 5,3% dels casos de trisomia 21 i no estan associats amb una edat avançada de la mare.

6) L'edat avançada de la mare suposa un factor de risc tant pels errors de procedència MMI com pels de MMII.

Origen Parental de l'error	Freqüència	Proporció
Errors meiòtics		
Maternal (M):		
Meiosi I (MMI)	67	MMI/(MMI + MMII) = 75,3%
Meiosi II (MMII)	22	MMII/(MMI +MMII) = 24,7%
Origen desconegut	8	
Subtotal	97	Maternals/Total = 85,8%
Paternal (P):		
Meiosi I (MPI)	4	MPI/(MPI + MPII) = 50,0%
Meiosi II (MPII)	4	MPII/(MPI + MPII) = 50,0%
Origen desconegut	2	
Subtotal	10	Paternals/Total = 8,8%
Errors mitòtics (MT)	6	Mitòtic/Total = 5,3%
Total casos informatius	113	

Taula 2. Origen del cromosoma 21 extra en les trisomies lliures

Dades de Yoon et al, 1996.

1.1.5.2. TRISOMIA PER TRANSLOCACIÓ DEL CROMOSOMA 21

En el 5% dels individus amb la DS, la trisomia és deguda a una reordenació cromosòmica del cromosoma 21 a un altre cromosoma (p.e. el 14, 21, 13 ó 15).

En tots els casos examinats de trisomia 21q causats per una translocació robertsoniana *de novo* t(p.e. 14q;21q), el cromosoma extra és de procedència materna. En el 83% d'aquests casos s'observa una retenció a l'heterozigositat per a un marcador proximal, suggerint que el mecanisme més possible és que la translocació de cromàtides tingui lloc abans de l'entrecreuament de la meiosi I, seguit per una segregació normal en la meiosi I i II (Petersen *et al.*, 1991).

Mitjançant les anàlisis de polimorfismes de DNA s'ha observat que les reordenacions 21q;21q, en contra del que fins ara s'havia pensat, no sempre són translocacions robertsonianes rob(21q;21q). En la majoria dels casos estudiats (82%), les duplicacions dup(21q) són isocromosomes i(21q). La procedència parental d'aquests

errors és materna i paterna al 50%, i probablement s'originen per la manca de separació tant de les cromàtides en la meiosi II, com de les cromàtides germanes en la divisió mitòtica. La resta de casos (18%) són translocacions robertsonianes rob(21q;21q), causades per la fusió entre cromàtides heteròlogues, i deguts sempre a errors originats en les cèl·lules germinals maternes, a l'igual que s'observa per a les translocacions t(14q;21q) (Antonarakis *et al.*, 1990; Petersen *et al.*, 1991; Shaffer *et al.*, 1991; Shaffer *et al.*, 1992).

1.2. CROMOSOMA 21 HUMÀ

1.2.1. CROMOSOMA 21

L'HC21 ha estat objecte d'un gran esforç internacional sota els auspicis del Projecte Genoma Humà, hi ha esdevingut paradigma dels estudis de mapatge i subclonatge cromosòmic. Ha estat a l'avantguarda de l'estudi genètic i molecular dels cromosomes humans. És el primer cromosoma del qual s'ha disposat d'un mapa d'híbrids somàtics, d'un mapa de marcadors de DNA altament polimòrfics i d'un mapa de YAC. Aquesta vocació pionera i modèlica és deguda bàsicament a dues raons: 1) el cromosoma 21 és el cromosoma humà més petit i pot representar un bon model per a establir projectes experimentals i expectatives de recerca i 2) té una gran rellevància mèdica, ja que entre d'altres malalties genètiques localitzades al cromosoma 21 (veure apartat 1.2.4), la seva trisomia està associada a la DS.

El braç curt (21p) té una mida de 10-15 Mb (Doll, 1992) i acaba en una regió satèl·lit. En la part proximal d'aquest satèl·lit existeix una segona constricció que conté la regió organitzadora nucleolar, així com múltiples còpies de gens ribosomals. La regió adjacent al centròmer 21p conté nombroses seqüències de DNA repetitiu del tipus alfasatèl·lit i beta-satèl·lit. En aquest braç 21p no s'hi ha trobat gens, excepte en la banda p12 que és molt rica en gens ribosomals. Aquest braç curt és molt similar als dels altres quatre cromosomes acrocèntrics (13, 14, 15 i 22) compartint, doncs, nombroses còpies de diverses famílies de DNA repetit en tàndem. El braç 21p no es considera important per al fenotip de la DS, puix que els fenotips d'aquelles trisomies 21 o isocromosomes monocèntrics, en què aquest braç petit està delecionat (21pter-proximal q21), no són diferents dels fenotips de les trisomies 21 lliures, i que la total absència d'aquest braç 21p no està associada a cap tret clínic (Daniel, 1979).

El seu centròmer és ric en sequències específiques de DNA satèl·lit (Willard, 1990), que contenen llocs de reconeixement per a proteïnes que s'uneixen al centròmer durant la divisió cel·lular (Masumoto *et al.*, 1989). La composició de bases des del centròmer fins la part proximal 21q22.1 és uniformement rica en nucleòtids A-T. La composició de la part distal 21q22.1 fins a 21q22.3 és variable, però generalment és moderadament rica en G-C, i la part telomèrica 21q22.3 és altament rica en G-C (Gardiner *et al.*, EMBO 1990a). Aquestes variacions en la composició de bases està d'acord amb les densitats pobres, moderadament altes, i bastant altes, pel que respecta a cadascuna d'aquestes tres regions cromosòmiques respectivament, quant a nombre de gens, d'illes CpG i de dianes per a l'enzim de restricció *Not*I que s'han trobat fins ara.

El seu braç llarg (21q) mesura unes 37 megabases (Mb) aproximadament, o sia uns 37 milions de nucleòtids (Ichikawa *et al.*, 1993), i representa poc més de l'1,2% respecte els 3.200 Mb que es calcula que hi ha en un genoma humà haploide (Morton, 1991). S'estima que el braç 21q conté uns 600-1.000 gens (Antonarakis, 1993).

1.2.2. MARCADORS POLIMÒRFICS DEL TIPUS MICROSATÈL·LIT

Les repeticions curtes en tàndem (STR) o microsatèl·lits es defineixen com a sequències de DNA relativament curtes (<100 pb) amb un motiu (<7 pb) repetit en tàndem. Es troben dispersos pel genoma de tots els eucariotes, incloent els llevats, peixos, amfibis, insectes i mamífers. La seva distribució en el genoma humà sembla ésser a l'atzar i la seva mitjana és d'un cada 6-10 kb (Beckman i Weber, 1992). Els motius de repetició A, AC, AG, AAN i AAAN comprenen el 76% de tots els microsatèl·lits, on N pot ésser una C, G o T. Els microsatèl·lits més abundants són els del tipus (A)n amb un 34%, i els (CA)n amb un 19% del total de microsatèl·lits del genoma humà. El principal avantatge dels dinucleòtids del tipus (CA)_n polimòrfics és la seva alta heterozigositat (Weber i May, 1989; Litt i Luty, 1989) i la seva regular distribució: un cada 20-30 kb (Stallings et al., 1991). La descoberta dels STRs ha revolucionat el genotipatge per les següents raons: 1) només es requereix una poca quantitat de DNA a diferência de les anàlisis dels fragments de longitud polimòrfics de restricció (RFLP), 2) són en general fàcils de genotipar mitjançant la tècnica de la reacció en cadena de la polimerasa (PCR), a més a més poden ésser utilitzats en anàlisis de múltiples microsatèl·lits, ensems per la PCR o PCR multiplex, 3) són distribuïts ubíquament per tot el genoma humà a diferència dels minisatèl·lits (VNTR), amb una distància mitjana entre microsatèl·lits altament informatius d'un cada 100 kb, 4) exhibeixen una heterozigositat de fins el 70-80% a diferència dels RFLPs, on l'heterozigositat és rarament més gran del 40%, i 5) són altament polimòrfics, a diferència també dels RFLPs.

El nombre de repeticions d'aquests motius $(CA)_n$ pot variar d'un individu a l'altre, i és aquesta variació al·lèlica que els confereix la seva utilitat com a marcadors genètics. El nombre d'al·lels per a una determinada repetició curta en tàndem del tipus $(CA)_n$ pot ésser molt elevat, i, per tant, la seva heterozigositat pot ésser superior al 70%, amb un contingut informatiu del polimorfisme (PIC) sovint superior al 0,60. El PIC està en funció del nombre de repeticions, el nombre d'al·lels i de llur freqüència al·lèlica. D'aquesta manera, la informativitat d'un microsatèl·lit generalment incrementa amb el nombre de repeticions, especialment en el rang entre 11-17. La taxa de mutació *de novo* per als microsatèl·lits del tipus (CA)_n va des d'1x10⁻³ (Weissenbach *et al.*, 1992) fins a 4,5x10⁻⁴ per *locus* i per gàmeta (Kwiatkowski *et al.*, 1992, Banchs *et al.*, 1994). Aquesta taxa és proporcional a la freqüència d'heterozigots.

La recerca de les repeticions (CA)_n polimòrfics en el genoma humà està més que justificada. El seu ús s'ha estès: 1) per a la localització de malalties multifactorials, o bé per a identificar el gens implicats en determinades malalties monogèniques (Rosen *et al.*, 1993; Calonge *et al.*, 1994), 2) per a la construcció de mapes genètics (Bosch *et al.*, 1994; 1996), 3) per a la detecció d'inestabilitat genètica en malalties canceroses (Parsons *et al.*, 1993), 4) per a les anàlisis forenses i de paternitat (Fuentes *et al.*, 1993), i 5) per a l'anàlisi evolutiva d'una mutació o l'estudi geogràfic d'una població (Morral *et al.*, 1994).

1.2.3. MAPES GENÒMICS

Un mapa genòmic és l'ordenació dels diferents *loci* al llarg del cromosoma. Cada *locus* és un punt de referència, com per exemple un fragment de restricció, una repetició de DNA, un punt de trencament cromosòmic, una seqüència d'un marcador de DNA, un gen, un fragment d'una seqüència codificant (EST) o qualsevol altre tipus de seqüència anònima de DNA (STS). Tots aquests *loci* són emmagatzemats a la base de dades del GDB (http://gdbwww.gdb.org). Sovint els mapes del cromosoma 21 fan referència només al braç llarg 21q perquè la construcció de mapes del braç 21p és molt més difícil, atès que els braços curts de tots els cromosomes acrocèntrics comparteixen una gran similitud en la seqüència i contenen moltes seqüències de DNA repetitiu, incloent-hi gens ribosomals (Arnheim *et al.*, 1980) i de DNA satèl·lit, sobretot els anomenats tipus I, II i III (Choo *et al.*, 1988).

1.2.3.1. MAPA DE PUNTS DE TRENCAMENT DEL CROMOSOMA

Aquests tipus de mapes són també anomenats mapes d'híbrids somàtics o SCH maps. Els punts de trencament o chromosomal breakpoint que tenen lloc d'una forma natural al cromosoma 21 generen cromosomes anormals, que són introduïts individualment en una cèl·lula somàtica híbrida humana-rosegadora. Entre totes aquestes cèl·lules híbrides proporcionen una col·lecció de fragments de DNA definits i limitats a una determinada regió cromosòmica. El pannell de línies cel·lulars híbrides del cromosoma 21 delimita prop de 30 intervals (Gardiner et al, 1988; 1990; Patterson, 1991; 1993), on es poden

localitzar fragments de DNA, ja sigui pel mètode de la PCR o bé pel de la transferència de *Southern*. El principal desavantatge d'aquest tipus de mapa és la seva baixa resolució.

1.2.3.2. MAPA DE FRAGMENTS LLARGS DE RESTRICCIÓ

A partir d'una línia cel·lular híbrida humana-rosegadora (WAV-17) que conté el cromosoma 21 com a únic component humà, es realitza una digestió del DNA amb una endonucleasa de restricció de tall infreqüent, generalment *Nof*l. Les mides dels fragments de restricció generats són discriminades mitjançant la tècnica d'electroforesi en camps polsants (PFGE). El seu principal avantatge és que proporcionen una estimació força acurada de les distàncies físiques en kilobases. Hi ha publicats dos mapes de restricció amb l'enzim *Nof*l per al cromosoma 21q, que donen com a resultat una mida de 38 Mb per al primer (Ichikawa *et al.*, 1993) i de 43 Mb per al segon (Wang i Smith, 1994). Tot i que aquests dos mapes concorden en l'ordre de quasi tots els *loci*, divergeixen en la mida total del cromosoma. Aquestes diferències poden ésser degudes tant a les variables experimentals com a l'estat de metilació de les seqüències CpG; a la digestió parcial de les endonucleases o a la baixa resolució del DNA de baix pes molecular en les anàlisis de PFGE.

També disposem de mapes locals amb altres enzims de restricció: Delabar et al., 1993; Gingrich et al., 1993; Dufresne-Zacharia et al., 1994; Patil et al., 1994; Osoegawa et al., 1996; Burmeister et al., 1991; Dutriaux et al., 1994; Eki et al., 1996; Guimerà et al., 1997.

1.2.3.3. MAPA DE LLIGAMENT GENÈTIC

Es tracta d'un mapa de marcadors de DNA ordenats per anàlisi de lligament genètic o *linkage*. Les distàncies entre *loci* polimòrfics són estimades per les freqüències de recombinació meiòtica entre uns i altres al llarg del tot el cromosoma. La unitat de distància per als mapes genètics és el centimòrgan (cM), on 1 cM equival a una freqüència de recombinació de 0,01 i a una distància física mitjana de 300-500 kb en el cas del cromosoma 21q (McInnis *et al.*, 1993), tot i que les distàncies físiques i les genètiques no són equivalents. Aquesta correlació és molt variable segons les regions: així en el telòmer 1 cM equivaldria a una distància menor a l'HC21 (150 kb). Els primers mapes foren construïts mitjançant l'anàlisi per transferència de *Southern* de fragments de restricció de longitud polimòrfica com a marcadors de DNA (RFLPs) (Tanzi *et al.*, 1988; Tanzi *et al.*, 1989). Més recentment, els mapes de segona generació s'han construït utilitzant la tècnica de la PCR, genotipant marcadors polimòrfics CEPH. Mentre
la distància física és la mateixa entre dones i homes, existeix una enorme diferència pel que fa a la distància genètica. El mapa de lligament masculí es més petit que el mapa femení, ja que els cromosomes femenins solen recombinar més que els masculins, malgrat que en la part més distal de la regió 21q la situació és la inversa i en els cromosomes masculins hi ha un excés de recombinació. Les longituds totals són de 48 cM per al mapa masculí i de 84 cM per al mapa femení. La mida mitjana de sexes del mapa és de 66 cM (McInnis *et al.*, 1993; Matise *et al.*, 1994). No es coneixen, però, ni el significat biològic ni les bases moleculars d'aquestes diferències en la recombinació.

Actualment, la distància genètica mitjana entre marcadors adjacents és inferior a 1,5 cM (Antonarakis, 1996) (cal recordar que per identificar i localitzar el gen responsable d'una malaltia hereditària amb segregació mendeliana mitjançant estudis de lligament genètic, n'hi ha prou amb mapes de 5-10 cM de resolució). L'any 1996, Bosch *et al.*, publicaren un nou mapa de l'HC21 que integra 30 nous microsatèl·lits al mapa ja existent mitjançant llur genotipatge a les famílies de referència CEPH., obtenint una mida mitjana de sexes de 64,4 cM, amb una mida per al mapa femení i masculí de 79,2 i 49,4 cM respectivament.

1.2.3.4. MAPA D'HÍBRIDS DE RADIACIÓ

Per a la construcció d'aquests mapes s'irradien les cèl·lules somàtiques híbrides humanorosegadores amb raigs X, les quals contenen el cromosoma 21 com a únic component genètic humà. D'aquesta manera es generen fragments entre 8 i 10 Mb de mida que es tornen a fusionar amb la línia cel·lular del rosegador. Utilitzant la freqüència de trencament entre els diferents marcadors en combinacions de dos en dos es pot estimar la distància entre els *loci* i construir així un mapa de lligament, talment com si fos un mapa de lligament genètic basat en les freqüències de recombinació meiòtica, però sense la necessitat de genotipar marcadors de DNA polimòrfics (Goss i Harris, 1975).

Aquest mapa pot complementar el mapa de lligament genètic, sobretot en aquells punts on aquest darrer presenti una manca de recombinacions meiòtiques. Les distàncies entre *loci* s'expressen en centirays (cR) i depenen de la dosi d'irradiació emprada per a la creació dels híbrids. El primer mapa d'híbrids de radiació del cromosoma 21 va tenir una mida de 650 cR, incloent-hi 28 *loci* (Cox *et al.*, 1990). En aquest mapa, 1 cR correspon a 62 kb de mitjana. Més recentment s'han localitzat 161 STSs en els fragments genòmics generats per radiació, originant un mapa de 1.562 cR, on 1 cR equival a 24 kb de mitja (Stewart *et al.*, 1997). Els dos tipus de mapes coincideixen en l'ordre de la majoria de *loci*, malgrat que la distància està determinada en base a probabilitats, igual que en el mapa de lligament genètic de marcadors de DNA

1.2.3.5. MAPA DE CLONS SOLAPATS

Es tracta de tenir representat l'HC21 en petits fragments ordenats. Així, fragments del cromosoma 21 foren subclonats en diferents vectors. Un tipus de vector emprat fou el cromosoma artificial de llevat (YAC) que pot contenir des de 100 kb fins a 2.000 kb de DNA. Utilitzant seqüències polimòrfiques de DNA o altres STSs com a marcadors, aquests YAC poden ésser ordenats i localitzats mitjançant la PCR fins a construir un mapa contigu de clons solapats que representin tot l'HC21.

El braç llarg del cromosoma 21 fou el primer cromosoma humà a partir del qual es va construir aquest tipus de mapa (Chumakov *et al.*, 1992; Nizetic *et al.*, 1994). Aquesta col·lecció de YAC s'estén des del 21cen fins el 21qter i comprèn uns 810 YAC ordenats al llarg de 191 *loci*. El nombre mínim de YAC solapats que representen el braç llarg de l'HC21 és de 55 YAC, resultant una distància física de 33 Mb. Hi ha 9 regions del cromosoma 21 anomenades forats o *gaps* que no estan cobertes per YAC; tanmateix, la localització cromosòmica alguns d'aquests YAC es errònia. Els dos principals desavantatges dels YAC són l'alta freqüència de clons quimèrics (clons on els inserts són fragments de DNA no contigus en el genoma) i que poden arribar a suposar fins el 46% en algunes genoteques de YAC, i les delecions o reordenacions dels inserts. Aquest tipus de mapa es troba, doncs, en continua millora.

A part dels YAC, s'utilitzen també altres vectors que destaquen per la seva estabilitat i per tenir un inferior grau de quimerisme. Aquests vectors són: els cromosomes artificials del bacteriòfag P1 (PAC) que accepten fragments de DNA fins a 150 kb, els cromosomes artificials de bacteri (BAC) que accepten fragments de DNA fins a 130 kb i els cosmidis (unes 40 kb). Malgrat que no s'han completat encara els mapes cromosòmics de tots ells per a l'HC21, sí que podem disposar d'ells per a determinades regions, incloent-hi la regió crítica de la síndrome de Down (Patil *et al.*, 1994; Nizetic *et al.*, 1994; Dufresne *et al.*, 1994; Gosset *et al.*, 1995; Osoegawa *et al.*, 1996; Ohira *et al.*, 1996; Hubert 1997; Guimerà *et al.*, 1997; Lapenta *et al.*, 1998).

1.2.3.6. MAPA DE TRANSCRIPCIÓ

El mapa de transcripció representa la localització en el mapa físic dels gens i unitats de transcripció o EST. La recerca d'aquests EST de l'HC21 es duu a terme mitjançant diversos mètodes, encara que els dos més importants són la selecció de cDNA o cDNA selection (Cheng et al., 1994; Peterson et al., 1994; Tassone et al., 1995; Yaspo et al., 1995, Ohira et al., 1997; Guimerà et al., 1997) i la tècnica de captura d'exons o exontrapping (Lucente et al., 1995; Yaspo et al., 1995; Chen et al., 1996; Guimerà et al., 1997; Ohira et al., 1997; Dahmane et al., 1998). És remarcable el gran esforç

internacional que s'ha dedicat a aquesta tasca, dins la qual s'emmarca el treball aquí presentat. S'han identificat més de 1.500 EST de l'HC21, on possiblement estiguin representats més del 40% dels gens que es troben en el cromosoma 21, però només d'un petit nombre s'en coneix la seqüència sencera del cDNA o *full-length*.

1.2.4. GENS I MALALTIES LOCALITZADES EN EL CROMOSOMA 21

La densitat de gens varia segons quina sigui la banda cromosòmica. A la banda 21q hi ha regions riques i regions pobres en gens. Les bandes que es tenyeixen amb Giemsa o bandes G són riques en nucleòtids A-T i el seu contingut de gens és bastant baix. Contràriament, les bandes negatives amb Giemsa o bandes R solen ésser riques en dinucleòtids C-G i també en gens. Així, les bandes T, que són una subclasse de les bandes R, són especialment riques en gens. Les bandes cromosòmiques de la meitat telomèrica 21q22.3 i 21q22.1, les quals no es tenyeixen amb Giemsa, probablement contenen la majoria de gens d'aquest cromosoma (Gardiner *et al.*, 1990). Contràriament, la llarga banda positiva amb Giemsa 21q21 és aparentment una zona pobre en gens com ho confirmen els resultats dels experiments de recerca de seqüències transcripcionals en aquesta zona (Xu *et al.*, 1995).

S'han caracteritzat 101 gens a l'HC21 (accessibles al GDB), els quals només representen el 10-16% del nombre total de gens estimats de l'HC21. La funció proteica d'alguns d'aquests gens és coneguda i estan implicats en procesos cel·lulars tan importants i diferents com la divisió cel·lular, la senyalització i comunicació cel·lular, el desenvolupament cel·lular, el sistema immunitari, l'estructura, la mobilitat, el metabolisme, i factors de transcripció. Alguns d'aquests gens s'han associat específicament a una determinada malaltia:

- El gen APP (<u>Amyloid Precursor Protein</u>) que és el precursor de la proteïna ßamiloide. S'ha associat aquest gen amb una rara forma familiar autosòmica dominant de la malaltia d'Alzheimer (FAD), ja que mutacions puntuals en aquest gen se segreguen amb la malaltia en algunes famílies (Goate *et al.*, 1991; Murell *et al.*, 1991; Chartier *et al.*, 1991). Tanmateix, no és el gen principal involucrat en aquesta malaltia, ja que la majoria dels casos de la FAD estan lligats al braç llarg de cromosoma 14 (Schellenberg *et al.*, 1992).

- El gen CBS (Cystathionine <u>B</u>-Synthase) està involucrat en l'homocistinúria (Skovby et al., 1984; Avramopoulos et al., 1993), una malaltia metabòlica autosòmica recessiva caracteritzada per retard mental i alteracions oculars, musculars i esquelètiques. - El gen ITG\$2 (integrin \$2) codifica, per a un marcador de la superfície cel·lular, l'antigen CD18. S'ha associat amb la síndrome de la deficiència d'adhesió de leucòcits (Arnaout *et al.*, 1990). Aquells pacients que tenen mutacions puntuals en aquest gen es caracteritzen per tenir infeccions bacterianes recurrents.

- El gen AML1 (<u>Acute Myeloid Leukemia 1</u>) està sovint associat a translocacions t(8;21) involucrades en la leucèmia mieloide aguda (Miyoshi *et al.*, 1991).

- El gen PFKL (<u>Phosphofructokinase of Liver</u>) està relacionat amb una forma d'anèmia hemolítica causada per una deficiència de la fosfofructoquinasa hepàtica (Levanon *et al.*, 1986).

- El gen SOD1 (<u>Superoxid Dismutase 1</u>) transforma l'oxigen en peròxid d'hidrogen i està relacionat amb una forma familiar d'esclerosi lateral amiotròfica (FALS) (Rosen *et al.*, 1993), que és una malaltia degenerativa de les neurones motores que causa paràlisi progressiva i un desenllaç fatal, normalment al cap d'uns cinc anys després del seu inici.

- El gen HLCS, relacionat amb la deficiència de la biotina carboxilasa.

- El gen de l'epilèpsia mioclònica progressiva del tipus d'Unverricht-Lundborg (EPM1) (Lehesjoki *et al.*, 1991). Recentment s'han descrit dues mutacions en el gen *cistatina* B situat en la banda cromosòmica 21q22.3 en tres famílies no relacionades que presentaven EPM1 (Pennacchio *et al.*, 1996). Posteriorment s'ha trobat que la mutació majoritària és una expansió d'una seqüència repetitiva en tàndem de 12 nucleòtids (CCCCGCCCCGCG) en la zona promotora del gen. Aquesta expansió varia de mida i pot arribar a tenir més de 60 repeticions, mentre que en la població normal oscil·la entre 2 i 3 repeticions (Lafrenière *et al.*, 1997; Virtaneva *et al.*, 1997).

- El gen de la síndrome autoimmune poliglandular de tipus 1 (APECED) que es caracteritza per una resposta immune anormal (Aaltonen *et al.*, 1994).

D'altra banda, en el cromosoma 21 s'han trobat loci per a diverses malalties hereditàries monogèniques:

- La microcefàlia recessiva i retard del creixement intrauterí estan lligats a la monosomia parcial del cromosoma 21, entre els marcadors D21S11 i D21S55 (Matsumoto et al., 1997).

- Una forma de la malaltia maníaco-depressiva (Straub, 1994).

- Diversos loci per a una forma de sordesa hereditària (Veske et al., 1996).

- Una forma familiar d'esclerosi lateral amiotròfica (ALS1), una malaltia autosòmica dominant on hi ha una pèrdua progressiva de la funció motora (Siddique *et al.*, 1991).

- Una severa forma de la síndrome d'Usher (USH1E) caracteritzada per retinitis pigmentària combinada amb una pèrdua congènita de la capacitat auditiva (Chaib *et al.*, 1997).

- La síndrome Knobloch (Sertie et al., 1996).

- L'holoprosencefàlia (Muenke et al., 1995).

- Alguns casos de sarcoma d'Ewings i altres tumors neuroectodèrmics primitius associats a translocacions cromosòmiques (Zucman, 1993).

- L'amiloïdosi cerebrovascular de tipus holandès (HCHWA-D), una malaltia autosòmica dominant amb hemorràgies cerebrals recurrents i deposició de plaques amiloides en el cervell, provocada per mutacions en el residu 692 o en el 693 de la proteïna APP (Hendriks *et al.*, 1992).

1.2.5. REGIÓ CRÍTICA DE LA SÍNDROME DE DOWN

No tots els individus caracteritzats clínicament com a DS tenen una trisomia total del cromosoma 21 (veure http://www.infobiogen.fr/services/aneu21). Els mecanismes que causen els diferents fenotips que exhibeixen algunes persones amb la DS no són encara ben entesos, entenent com a fenotip tot paràmetre mesurable, com ara els trets clínics, físics, cel·lulars, morfològics, de comportament, etc. Una de les maneres d'abordar aquest estudi fou estudiar el casos dels individus amb trisomies parcials de l'HC21. Això ha permès establir una correlació entre un determinat fenotip i la regió trisòmica involucrada en aquest fenotip. La construcció d'aquests mapes fenotípics mitjançant correlacions fenotip-genotip ha permès delimitar molecularment determinades regions cromosòmiques de l'HC21 a un particular tret de la DS (Fig. 3) (Korenberg, 1991; Delabar et al., 1993). La delimitació d'aquestes regions són objecte de continus acotaments. L'any 1990, els treballs de Williams et al. van determinar que el fragment del cromosoma 21 entre els loci D21S13 i D21S58, involucrant les bandes cromosòmiques 21q11 i 21q22.1 i possiblement part de la regió 21q22.2, no estava implicat en la típica presentació clínica de la DS, tot i que descrivien la implicació d'aquesta zona en la formació d'altres caràcters com una discapacitat intel·lectual moderada, una estatura minvada i petites diferències en el patró de dermatoglifs dels palmells. Aquests estudis suggereixen que regions distals al marcador D21S58 podien contribuir a les manifestacions fenotípiques més característiques de la DS. Altres estudis també corroboren aquestes observacions (Niebuhr, 1974; Williams et al., 1975).

Actualment, s'ha suggerit que una dosi multigènica no balancejada de la regió que envolta el marcador de DNA D21S55 en la banda 21q22.2, entre els marcadors D21S17 i ETS2, proximal i distal al marcador D21S55 respectivament, pot contribuir significativament a la DS. La mida d'aquesta regió ha estat estimada entre unes 400 i

3.000 kb (Rahmani *et al.*, 1989; Rahmani *et al.*, 1990). Quant aquesta regió es troba en triple dosi està associada a 13 dels 24 trets fenotípics observats a la DS, incloent-hi hipotonia muscular, talla petita, alguns trets facials, anomalies de les mans i dels peus i una greu discapacitat intel·lectual (Delabar *et al.*, 1993). Aquesta regió, anomenada regió crítica de la síndrome de Down (DSCR) o també regió cromosòmica de la DS (DCR) (Rahmani *et al.*, 1989), representa aproximadament l'1-6% del braç 21q i s'estima que conté uns 10-100 gens.



Figura 3. Mapa fenotípic de la trisomia 21 basat en anàlisis citogenètiques de trisomia parcial del cromosoma 21. Mapa actualitzat de l'original (Korenberg, 1991).

Aquests mateixos estudis no exclouen la contribució d'altres regions del cromosoma 21 responsables d'altres fenotips característics de la DS. Així, els defectes cardíacs (relatius al canal àtrio-ventricular) s'han localitzat en una regió més distal, d'unes 5-6 Mb, entre els marcadors *D21S267* i *MX1* (Korenberg *et al.*, 1992). Alguns

24 Introducció-

trets facials, alçada curta i una discapacitat intel·lectual moderada són trets fenotípics que també s'han localitzat a regions més proximals de la DSCR.

Les associacions genotip-fenotip amb un determinant fragment de l'HC21 són extremadament difícils i controvertits, donada la penetració incompleta de la majoria dels trets de la DS, així com el fet que possiblement molts gens del cromoma 21 i dels altres cromosomes condicionen l'expressió final de cada tret. L'assignació d'una regió crítica a la regió de solapament entre diferents trisomies parcials per a un determinat tret fenotípic pot ésser incorrecta en cas que aquest tret estigui causat per més d'un gen. Si aquest fos el cas, el gen candidat d'aquest fenotip podria ésser erròniament localitzat a la zona de solapament. D'aquesta manera, és important saber si un tret està determinat per més d'un locus. En cas contrari, hauríem de saber quina variabilitat del tret correspon a una determinada regió cromosòmica (Korenberg et al., 1994). L'abordatge d'aquest darrer punt és molt difícil atesa la penetració (probabilitat d'expressar un tret determinat causat per un gen) i l'expressivitat (variabilitat en l'expressió fenotípica del tret) d'alguns trets clínics de la DS. Només quan la penetració i l'expressivitat d'un tret és igual entre la població amb trisomia total del cromosoma 21 i la població amb trisomia parcial per a una determinada regió de l'HC21 es pot pensar que existeix un sol locus responsable de la variabilitat d'aquest tret (Korenberg et al., 1994).

Oposadament a aquesta afirmació, altres autors creuen que aquesta pèrdua de balanç genètic provoca una interacció entre tots els productes gènics del genoma humà que no ha estat seleccionada per l'evolució, i, per tant, dóna als individus amb la DS una susceptibilitat a la genètica a l'entorn i als errors estocàstics, que els individus sense la DS poden contrarestar (Shapiro, 1994). D'aquesta manera, aquests darrers autors neguen l'existència de la regió crítica argumentant que el fenotip de la DS és més probable que sigui degut a les múltiples interaccions entre els múltiples al·lels de cada *locus* del genoma humà que no pas a l'existència d'un *locus* únic com a responsable d'un conjunt de trets (Shapiro, 1997). Aquestes discrepàncies entre els diferents autors, pel que fa a l'existència de regions crítiques per a determinats trets fenotípics ha motivat el desús dels termes DSCR o DCR (6th International Workshop on Chromosome 21, Cold Spring Harbor, 1996). Sigui com sigui, no hi ha dubte que l'etiologia multifactorial de la DS és extraordinàriament complexa.

1.2.6. MODELS MÚRIDS DE LA SÍNDROME DE DOWN

Els models murins permeten estudiar la funció biològica d'un gen *in vivo*, així com el resultat del desequilibri de la seva expressió quan la seva dosi gènica està alterada i la repercussió d'aquest sobre l'expressió d'altres gens, que també poden estar alterats. Això

permet establir correlacions entre determinats fenotips bioquímics, morfològics i conductuals dels ratolins trisòmics i dels transgènics amb algunes de les característiques observades en la DS.

1.2.6.1. RATOLÍNS TRISÒMICS

1.2.6.1.1. Ratolí amb trisomia completa del cromosoma de MMU16

Part de l'HC21 és sintènica a determinats segments de tres cromosomes de ratolí, el 10, 16 i 17. Una part distal del cromosoma 16 de ratolí (MMU16) comparteix un gran nivell de conservació amb una part del braç llarg de l'HC21 (anomenat també HSA21) (Epstein *et al.*, 1985). Aquest fragment sintènic del genoma múrid inclou la DSCR de l'HC21 i conté almenys 18 *loci*, tots ells situats entre els gens *APP* i *MX1* de l'HC21, amb el mateix ordre seqüencial. Tanmateix, no es pot dir que MMU16 i HSA21 siguin equivalents, ja que el primer, a banda d'ésser més llarg, conté gens que es troben localitzats en altres cromosomes humans com el 3, 16 i 22.

Per tot això, el valor com a model del ratolí trisòmic del cromosoma MMU16 (Ts16) és limitat. A més a més, els fetus amb trisomia simple del MMU16 no són viables i normalment moren *in utero* entre el 15è i el 18è dia de l'embaràs a causa de greus cardiopaties congènites i de profundes alteracions en el desenvolupament del sistema nerviós (Miyabara *et al.*, 1982). Encara que el ratolí Ts16 no serveixi com a model per a estudiar el desenvolupament postnatal, ha permès analitzar les alteracions originades en el desenvolupament fetal. D'aquesta manera el ratolí Ts16 presenta alteracions en el sistema nerviós, hipoplàsia cerebral, dèficit en la neurogènesi de les neurones colinèrgiques, degeneració d'aquestes neurones en l'hipocamp, deficiències en el sistema immunitari, malformacions en les vàlvules cardíaques (100% dels embrions de ratolí) i alteracions en l'aparell urinari.

1.2.6.1.2. Ratolí amb trisomia parcial del cromosoma MMU16 (Ts65Dn)

Donades les limitacions del ratolí Ts16 com a model animal de la DS, es va desenvolupar un ratolí amb trisomia parcial del cromosoma 16 (Ts65Dn). La trisomia d'aquests ratolins queda restringida a un determinat segment del genoma múrid (9-12 Mb, representant el 15% del cromosoma 16) que correspon a aquella regió de l'HC21 localitzada entre el marcador D21S16 i el final del cromosoma 21 (21q21-22.3), incloent la DSCR (Reeves *et al.*, 1995). Aquest model animal té també un valor limitat, ja que el ratolí trisòmic Ts65Dn conté també una trisomia per a una regió pericentromèrica del cromosoma 17 de ratolí (<10%), la qual no és homòloga a l'HC21. A diferència del ratolí Ts16, el ratolí Ts65Dn és viable i es desenvolupa, creix i envelleix. Això ha permès observar que la trisomia del cromosoma MMU16 en ratolí està associada

26 Introducció-

a un seguit d'alteracions que inclouen un endarreriment primerenc del desenvolupament, una reducció del pes en el naixement, una esterilitat en el mascle i unes alteracions en els trets facials i alteracions conductuals. Els estudis mostren que l'activitat locomotora dels ratolins trisòmics és superior respecte als controls, demostrant una hiperactivitat en diferents proves experimentals en la fase obscura del cicle. Així mateix s'ha referit també una major activitat exploratòria. Els resultats, però, suggereixen que aquesta conducta és en certa mesura estereotipada i poc atenta, indicativa d'un probable dèficit d'atenció (Escorihuela *et al.*, 1995; Reeves *et al.*, 1995).

Els ratolins Ts65Dn mostren alteracions en la memòria, tant la de curt termini com la memòria de treball, i un dèficit en el procés d'aprenentatge, concretament en la tasca d'aprenentatge espaial o test del laberint aquàtic de Morris (Escorihuela et al., 1995; Reeves et al., 1995), així com un procés degeneratiu més ràpid en alguns aspectes sensòrio-motors com el reflex prènsil i la coordinació motora, si bé després s'iguala en edats molt avançades. En relació a la dinàmica sinàptica, s'ha demostrat alteracions neuroquímiques en alguns sistemes de neurotransmissió. Així, la resposta del sistema enzimàtic acoblat al receptor B-adrenèrgic en la transducció de senyals cel·lulars està profundament alterada, tant la seva activitat basal com l'estimulada. Les seves consequències són una comunicació interneuronal inadequada que repercuteix crucialment en la instauració dels processos de memòria, d'aprenentatge, d'atenció i d'activitat psicomotora (Dierssen et al., 1997). Tanmateix, els nivells basals d'AMP cíclic es troben disminuïts en totes les estructures cerebrals, sobretot en l'hipocamp i en l'escorça cerebral, però no en el cerebel (Dierssen et al., 1996). S'observen resultats molt similars quan s'estimula l'activitat adenilat-ciclasa directament amb forskolina que actua sobre la subunitat catalílita de l'enzim. L'alteració del sistema de segons missatgers constitueix un factor de gran importància en la dinàmica de la comunicació intercel·lular i podria explicar parcialment les alteracions conductuals descrites. A nivell morfològic, la comparació entre els ratolins trisòmics Ts65Dn i els controls mostra que la triplicació dels gens distals del cromosoma 16 no representa una alteració profunda en l'organització general del SNC. Aquestes dades són consistents amb les observades en la DS. En el ratolí Ts65Dn no s'observen trets fenotípics com els defectes cardíacs ni esquelètics. L'absència d'aquesta neuropatologia en els ratolins Ts65Dn fa pensar que aquests trets poden ésser deguts en part a una dosi no balancejada d'alguns gens de l'HC21 no representats en el fragment del MMU16 duplicat en el ratolí Ts65Dn.

1.2.6.1.3. Ratolí amb trisomía parcial del cromosoma MMU16 (Ts1Cje)

La regió trisòmica del ratolí Ts1Cje és més petita que la del ratolí Ts65Dn i comprèn la regió cromosòmica entre Sod1 i Mx1, que correspon a la banda sintènica 21q22.1-22.3 de l'HC21, icloent la DSCR. A diferència del Ts65Dn, el ratolí Ts1Cje no és trisòmic per la part distal del cromosoma MMU16, la qual conté gens que no es localitzen en l'HC21. Si bé el ratolí Ts1Cje tampoc mostra alteracions morfològiques macroscòpiques, s'observen alteracions conductuals i d'aprenentatge espaial més lleus que les mostrades pel ratolí trisòmic Ts65Dn, però més greus que les mostrades pel ratolí transgènic del YAC 152F7 (com s'explicarà en l'apartat 1.2.6.2.1) (Sago *et al.*, 1998).

1.2.6.2. RATOLINS TRANSGÈNICS

1.2.6.2.1. Ratolins transgènics amb YAC

Els ratolins transgènics que esdevenen de la microinjecció en oòcits de ratolí d'un determinat YAC que conté material genètic humà han permès estudiar els efectes fenotípics de la trisomia de només uns pocs gens humans, amb l'avantatge que tots ells s'expressen amb el seu promotor natural. A més a més, permeten estudiar la interacció entre els diferents productes transgènics. Amb aquestes pretensions es construeixen les genoteques *in vivo* de ratolins que representen col·leccions de ratolins transgènics per a un determinat YAC o per a un clon P1 de la regió 21q22.2 (Smith *et al.*, 1995). Aquestes genoteques permeten identificar la localització d'un gen responsable d'un fenotip concret a determinades regions del cromosoma, basant-se en els efectes de sensibilitat de dosi. Aquests ratolins poden ésser utilitzats posteriorment en experiments genètics d'encreuament amb altres ratolins transgènics o nul·lisòmics per a un determinat gen i estudiar la interrelació entre diferents gens.

Les anàlisis funcionals d'una genoteca *in vivo* de 2 Mb de la regió 21q22.2 mostren que solament els ratolins transgènics amb el YAC 152F7 (amb una mida de 570 kb) exhibeixen problemes d'aprenentatge i de memòria, comparats amb els animals controls no transgènics o amb altres ratolins transgènics amb altres YAC que no contenen aquesta regió cromosòmica (Smith i Rubin, 1997). És més, aquells ratolins que són transgènics només pel fragment telomèric de 180 kb del YAC 152F7 mostraven el mateix fenotip que aquells ratolins que eren transgènics per tot el YAC. D'aquesta manera es localitzà la causa genètica d'aquests problemes d'aprenentatge i de memòria a una determinada regió de l'HC21 entre els marcadors D21S394 i D21S337. L'únic gen caracteritzat fins ara en aquesta regió és el gen *MNBH* (Guimerà *et al.*, 1996). Aquests resultats suggereixen que una dosi alterada del gen *MNBH* pot estar associada al desenvolupament neuronal anormal i a les alteracions de l'aprenentatge en el test del laberint aquàtic de Morris que presenten aquests ratolins transgènics (Smith *et al.*, 1997). D'aquesta manera s'ha suggerit que el gen *MNBH* podria estar implicat en la

patologia molecular de la discapacitat intel·lectual de la DS (Smith *et al.*, 1997). La estratègia científica de generar genoteques *in vivo* de ratolins transgènics amb YAC de l'HC21 ha esdevingut una poderosa eina metodològica, ja que l'estudi gènètic de la DS podria centrar-se en aquells gens localitzats en els YAC pels quals els ratolins transgènics mostrin alteracions reminiscents a les observades en la DS.

1.2.6.2.2. Els ratolins transgènics d'un determinat gen

Els ratolins transgènics d'un determinat gen de l'HC21 ens ajuden a entendre millor quina és la possible contribució puntual d'un gen al fenotip de la DS. En aquest sentit són remarcables els ratolins transgènics del gen *APP*, del gen *SOD1* i del gen *Ets*-2.

1.2.6.2.2.1. Ratolí transgènic del gen APP

Els ratolins transgènics per a una de les isoformes del gen humà *APP* provoca una alteració en la relació d'abundància entre les seves isoformes, en favor de la ß-APP751 respecte a la isoforma ß-APP695. Aquests transgènics presenten dipòsits difusos extracel·lulars d'amiloide a nivell cerebral, especialment a l'escorça i a l'hipocamp (Quon *et al.*, 1991; Moran *et al.*, 1995). Aquests dipòsits s'assemblen molt a les estructures beta-amiloide característiques de la malaltia d'Alzheimer. Tanmateix, el ratolí transgènic amb la isoforma 695 del gen humà *APP* no presenta dipòsits al cervell, indicant que és la isoforma 751 la que està implicada en la gènesi de la proteïna amiloide PBA i per tant en la formació de les plaques amiloides. Cal fer esment que en pacients d'Alzheimer també es troben alteracions en l'expressió relativa d'aquestes diferentes isoformes.

Altres treballs indiquen que els animals transgènics que expressen altament una forma mutada del gen APP a partir d'un promotor heteròleg són capaços de formar plaques similars a les observades en l'AD (Games *et al.*, 1995). Cal fer esment que els ratolins transgènics amb el YAC que conté el gen APP humà i que sobreexpressen la proteïna completa ß-APP humana a partir del promotor natural no desenvolupen una patologia similar a la de l'Alzheimer (Lamb *et al.*, 1993).

1.2.6.2.2.2. Ratolí transgènic del gen CuZnSOD

Els ratolins transgènics que sobreexpressen el gen humà CuZn-superòxid dismutasa (CuZnSOD) no presenten alteracions significatives ni en la seva aparença ni en el seu comportament. Les seves plaquetes però, contenen concentracions més baixes de 5-hidroxitriptamina (5-HT) donada la seva baixa capacitat per captar-la des de la sang. Aquesta observació és idèntica a la que es detecta en les plaquetes de les persones amb la DS. També s'observen unions neuromusculars anormals amb signes de desenvolupament

pobre i canvis degeneratius prematurs, que recorden a les observades en els individus amb trisomia total de l'HC21 (Groner *et al.*, 1990; Avraham *et al.*, 1991). Existeix també una disminució de la biosíntesi i alliberament de les prostaglandines E_2 i D_2 , tal i com succeeix en fibroblasts obtinguts de persones amb la síndrome de Down (Minc-Golomb *et al.*, 1991).

1.2.6.2.2.3. Ratolí transgènic del gen Ets-2

Els ratolins transgènics que sobreexpressen el protooncogèn i factor de transcripció *Ets*-2 desenvolupen anomalies esquelètiques, especialment en la regió crànio-facial i viscerocranial i en l'esquelet cervical (Sumarsono *et al.*, 1996), similars a les observades en el ratolins trisòmics del cromosoma 16 i en els individus amb la DS, els quals sobreexpressen també el gen humà *Ets*-2.

1.2.6.2.2.4. Ratolí transgènic del gen PFKL

El gen *PFKL* codifica per a una de les subunitats de les quatre que formen la fosfofructoquinasa (PFK), un enzim de la via glucolítica que fosforila la fructosa-6-P donant lloc a la fructosa-1,6-difosfat. És fàcil trobar una activitat incrementada de l'enzim glucolític fosfofructoquinasa en eritròcits i fibroblasts d'individus amb la DS (Elson *et al*, 1994). Els animals transgènics amb aquest gen mostren un increment de la concentració d'aquest enzim en el fetge, en els hematies i en el cervell embrionari (Groner *et al.*, 1995). No se sap quina és l'alteració que pot comportar aquest gen en els individus amb la DS.

1.2.6.2.2.5. Ratolí transgènic del gen HMG-14

La proteïna HMG-14 és una proteïna nuclear que podria modular l'estructura de la cromatina transcripcionalment activa. Els ratolins transgènics que sobreexpressen el gen humà HMG-14 tenen una incidència més elevada de quists en l'epiteli del timus (Bustin et al., 1995).

1.2.6.2.2.6. Ratolí transgènic del gen S100ß

S100ß és una proteïna expressada en el cervell humà que fixa calci i està implicada en el desenvolupament del cervell, en la neurofisiologia de l'hipocamp, en la proliferació de neurites *in vitro*, i en l'estimulació de la resposta d'astrogliosi davant diverses lesions cerebrals, incloent la degeneració neuronal de la malaltia d'Alzheimer. Diverses línies de ratolins transgènics amb aquest gen humà mostren un ventall de fenotips, com disfuncions en les propietats electrofisiològiques de l'hipocamp i una hiperactivitat

específica de les femelles (Gerlai i Roder, 1995). A més a més, en aquests ratolins s'observa també que les interaccions entre la glia i les neurones es troben alterades, mostrant una elevada expressió de marcadors de ramificació axonal incloent els neurofilaments L, els neurofilaments M i H fosforilats i la beta-tubulina. A nivell immunocitoquímic exhibeixen alteracions en la morfologia dels astròcits i en la ramificació axonal, especialment en el gir dentat (Reeves *et al.*, 1994). Altres característiques d'aquests transgènics són l'astrogliosi focal i la proliferació neurítica, especialment en el cerebel i en l'hipocamp (Cabin *et al.*, 1995).

Cal remarcar que els fenotips neurològics, morfològics, bioquímics, conductuals i altres deficiències que mostren tots aquests ratolins transgènics no reflecteixen les possibles interaccions entre l'expressió del gen exògen del ratolí transgènic i les expressions dels altres gens del genoma humà. A més, els nivells de sobreexpressió d'aquests gens en els ratolins transgènics no sempre guarden una proporció comparable amb la que es troben en els individus amb la DS.

1.3. DIAGNÒSTIC PRENATAL I CONSELL GENÈTIC 1.3.1. DIAGNÒSTIC PRENATAL

Destacaré els tres mètodes principals de diagnòstic prenatal de la DS: 1) anàlisi de l'alfafetoproteïna en sang materna (AFP), 2) amniocentesi i cariotip fetal, i 3) obtenció de vellositats corials (CVS) i cariotip fetal. El clivellatge de la DS mitjançant l'estudi de la concentració de l'alfa-fetoproteïna en el sèrum matern (MS-AFP) es pot realitzar a totes les dones embarassades, mentre que els estudis de citogenètica dels altres dos mètodes s'empren només per a una determinada població de risc, com en aquells pares que són portadors d'una translocació, o en aquelles mares que tenen un valor baix del contingut de MS-AFP, o una edat avançada de la mare (>35 anys). Cal tenir en compte que en les dones joves, el risc de patologia congènita mendeliana supera el risc de patologia cromosòmica.

El primer mètode detecta també aquells fetus amb defectes en el tub neural. Aquests defectes produeixen nivells de MS-AFP més baixos respecte les mares amb fetus sense defectes en el tub neural. L'anàlisi de la concentració d'AFP es pot aplicar també als fluids amniòtics. En tots dos casos s'observen uns valors disminuïts de 0,64-0,72 (Cuckle *et al.*, 1984; Cuckle *et al.*, 1985). Tanmateix, un disminució en els valors de la MS-AFP pot resultar de vegades en un increment de l'AFP del líquid amniòtic, suggerint que totes dues concentracions no sempre estan directament relacionades (Macri *et al.*, 1986). L'amniocentesi és l'obtenció de líquid amniòtic, i normalment és duta a terme entre les setmanes 14a i 16a de gestació. Té un risc lleugerament més petit que la CVS en termes de mortalitat fetal provocada, per ésser aquesta darrera una pràctica invasiva, encara que el temps d'obtenció del diagnòstic és més llarg en el cas de l'amniocentesi. En tots dos casos la finalitat darrera és fer un cariotip. La CVS es practica entre les setmanes 6a i 10a de la gestació.

Altres marcadors bioquímics de la trisomia 21 són els valors elevats de l'hormona gonadotrofina coriònica humana (alfa/beta hCG) (Bogart MH, 1992) i/o la determinació de valors baixos de l'estriol no conjuminat (uE_3) (Haddow, Canick, 163 1990). Emprant una combinació de les determinacions d'aquests tres marcadors, l'índex de risc es calcula a partir dels riscs derivats de més d'un marcador, incrementant així la sensibilitat del clivellatge fetal de la DS (Taula 3).

Taula 3. Rendiment dels programes de clivellatge de prevenció de la DS segons el nombre i el tipus de marcadors utilitzats

Tipus d'anàlisi	Nombre de falsos positius (en percentatge)		
Edat materna	17		
Edat/AFP/hCG	9		
Edat/AFP/hCG/uE ₃	8		
Edat/AFP/uE3 /beta hCG/alfa hCG	7		

AFP: Alfa-fetoproteïna, hCG: hormona gonadotrofina coriònica, uE_3 : estriol no conjuminat. Dades de M. Lluch (1995).

Aquests darrers marcadors utilitzats no pretenen diagnosticar la DS, sinó seleccionar la població de risc a la qual oferir el diagnòstic pre-natal citogenètic, ja que des del punt de vista de la relació cost-benefici, el rendiment en el nombre de casos diagnosticats amb la DS respecte al nombre total de casos estudiats és molt baix en aquells casos en els quals no hi ha cap indicació fundada (Taula 4).

Indicacions	Nombre de casos estudiats	Percentatge del total dels casos diagnosticats	Nombre de casos amb la DS	Percentatge
Edat >35 anys	2.491	64,7	25	1
Antecedents	453	11,7	4	0,8
Ansietat	906	23,5	1	0,1

Taula 4. Incidència de DS en diagnòstic prenatal citogenètic

Dades obtingudes a partir de 3.850 amniocentesis (Institut de Bioquímica Clínica) (Lluch, 1995).

1.3.2. RISC DE RECURRÈNCIA DE LA DS

¿El naixement d'un infant amb trisomia 21 lliure predisposa al naixement d'un altre infant amb la DS?. Estudis basats en dades de nounats indiquen que aquelles mares per sota dels 30-35 anys d'edat que ja han tingut un nounat amb la trisomia 21, tenen un risc significativament incrementat de tenir-n'hi un altre, mentre que per a les mares més grans, tenen un risc molt poc incrementat, en addició al risc que ja tenen segons la seva edat (Warburton, 1985). Aquestes dades corroboren la creença que els casos de trisomia 21 en mares joves provenen principalment de causes independents a l'edat de la mare. Els parents de segon grau no semblen tenir un risc incrementat significatiu (Eunpu *et al.*, 1986; Abuelo *et al.*, 1986; Berr *et al.*, 1990).

Pel que fa a les translocacions robertsonianes *de novo* hi ha també un risc incrementat que depèn del sexe parental. Si el portador és el pare, el risc és de l'1-2%, però si la translocació s'hereta de la mare, el risc és del 10-15% (Stene, 1970; Hamerton, 1971). L'explicació a aquest fenomen és encara desconeguda. En aquells casos rars on un dels pares tingui una translocació 46,-21,+rb(21q;21q), el risc és del 100%, ja que l'altra possibilitat 45,-21 no és compatible amb la vida.

Cal fer esment que la capacitat reproductora està minvada en les dones amb trisomia 21, però a diferència dels homes, tenen sovint una certa capacitat de reproducció. Aquestes dones tenen un risc de recurrència teòric al voltant del 50%, però a la pràctica és inferior, ja que el fetus trisòmics tenen una probabilitat més gran de no néixer respecte aquells que no ho són.

1.4. AÏLLAMENT DE SEQÜÈNCIES TRANSCRIPCIONALS 1.4.1. CONSIDERACIONS GENERALS

En els estudis d'identificació del gen responsable d'una malaltia determinada o els gens responsables d'una símdrome, cal aïllar i caracteritzar els gens de la regió candidata a tenir el gen. Les tècniques d'aïllament de seqüències transcripcionals que es descriuen a continuació no són tècniques dirigides a trobar el gen d'una determinada família de gens, o gens homòlegs d'altres espècies, sinó que intenten aïllar fragments de gens a partir d'una regió genòmica, sense tenir cap informació prèvia del gen que cerquem. D'aquesta manera, no parlarem dels mètodes de clivellatge de genoteques genòmiques amb sondes de cDNA o *direct screening with cDNA*.

Hi ha bàsicament dos tipus de tècniques: aquelles que estan basades en els motius típics de l'estructura dels gens i seleccionen elements funcionals involucrats en l'expressió d'aquest a partir del DNA i les basades en l'anàlisi de RNA d'un teixit determinat. Els motius d'estructura dels gens poden ésser: 1) els exons, en les tècniques de captura d'exons o exon trapping, i en la tècnica de la PCR d'*Alu-splice*, 2) els promotors, com en la tècnica de la recerca de les illes CpG, i 3) un compendi de tots ells, com en l'anàlisi de la seqüència de DNA per computadora. En la segona classe de tècnica, sabem *a priori* que el gen s'expressa en determinats teixits, com és el cas de la tècnica de la selecció de cDNA o *cDNA selection*. La diferència més important entre els dos tipus de tècniques és que el primer tipus representa mètodes que no depenen ni de l'especificitat de teixit ni de l'expressió temporal dels gens, mentre que el segon tipus sí que depèn de la presència d'un transcrit en una determinada genoteca de cDNA.

Cadascun d'aquests mètodes està afectat en grau diferent per aquests tres problemes: verificació, redundància i totalitat. El problema de la verificació, comprovar que els clons positius siguin realment trànscrits i descartar aquelles següències que no són veritablement expressades, és més complex per a aquells mètodes que identifiquen gens de manera independent a llur expressió en un teixit (p.e. exon trapping, Grail, PCR Alu-Splice). Aquesta limitació vé donada perquè no tots els gens s'expressen en un mateix teixit ni en un moment determinat, de manera que és molt difícil saber on s'expressa el gen trobat i quan ho fa. De vegades, això impossibilita determinar si una següència es bona fide o no des del punt de vista transcripcional. El problema de la verificació no és tan greu quan s'utilitzen tècniques basades en material expressat d'un determinat teixit (p.e. cDNA selection), perquè els cDNA seleccionats han d'ésser en la genoteca de cDNA del teixit emprat en la mateixa selecció dels cDNA. La tècnica de la selecció de cDNA, a diferència dels altres mètodes mencionats, pateix del problema de la totalitat, és a dir, cal aplicar la tècnica a diferents genoteques de cDNA de diferents teixits i edats per trobar tots els gens d'una regió determinada. D'aquesta manera, els problemes de verificació i de totalitat són recíprocs. El problema de la redundància és més propi de les tècniques basades en el cDNA, ja que l'anàlisi de la col·lecció de clons resultants d'aquesta tècnica mostra que sovint està enriquit amb sequències ribosomals i repeticions genòmiques. Aquest darrer problema no representa realment un gran handicap, ja que aquests clons es descarten amb una sola hibridació emprant DNA ribosomal o repetitiu com a sondes, respectivament. Per tal de superar aquestes limitacions inherents a cada tècnica, s'imposa la utilització de diverses tècniques amb esbiaix diferents, amb l'esperança que la suma d'aquests esbiaixos s'emmascarin i la seva combinació ofereixi un mapa transcripcional complet de la regió. En la Taula 5 es resumeixen els diferents esbiaixos i limitacions de cada tècnica. Ja que cap d'aquests mètodes permet trobar tots els gens que hi ha en una

. 4

regió determinada i cadascun d'ells pateix un esbiaix diferent, es fa necessària la combinació de tots ells per a obtenir un mapa de transcripció exhaustiu.

Per a comprendre la fisiopatologia dels diferents trets de la trisomia 21 cal conèixer cadascun dels gens que es localitzen al cromosoma 21. La seva caracterització ens ha de permetre conèixer llur funció i els seus efectes de sensibilitat de dosi, individualment o en combinació, en cultius cel·lulars o en animals transgènics. D'aquesta manera, s'espera elucidar la contribució fenotípica de cada gen a la DS.

1.4.2. TÈCNICA DE LA CAPTURA D'EXONS O EXON TRAPPING

Aquesta tècnica es basa en les seqüències consensus del DNA dels senyals de processament dels introns o splicing sites. D'aquesta manera, el DNA genòmic de la regió d'interès, prèviament clonat en vectors (cosmidis, YAC, etc.) es digereix amb un enzim de restricció (p.e. BamH1) i es subclona en un vector dissenyat especialment per a la captura d'exons, com el plasmidi pSPL3 (Fig. 4). Aquest vector té dues seqüències consensus de processament dels introns, una acceptora (SA) i una altra donadora (SD) a banda i banda del lloc múltiple de clonatge (MCS) que es troba a l'intró de l'HIV tat del plasmidi pSPL3 (Fig. 5). Aquest vector té unes seqüències que li permeten replicar-se en cèl·lules hostes d'*E. coli* i en cèl·lules eucariotes COS-7.

Després de transfectar les cèl·lules COS-7 amb el DNA purificat dels plasmidis pSPL3 recombinants, s'inicia la replicació i la transcripció a alts nivells, gràcies a l'origen de replicació de l'SV40 i del seu promotor. Les cèl·lules COS-7 reconeixen i processen els introns de les molècules de RNA pre-missatgeres durant el processament dels trànscrits. D'aquesta manera, l'exó present en el fragment de DNA genòmic subclonat queda atrapat entre les seqüències de tall i unió del vector i les de l'exó genòmic, formant una sola unitat transcripcional (Fig. 5). Virtualment, tot exó present en aquest fragment de DNA subclonat que tingui els senyals *consensus* intactes de tall i unió podrà processar els dos introns que l'envolten, obtenint així un RNA quimèric a partir dels promotors del propi vector.

Després de la recuperació de l'RNA citoplasmàtic es procedeix a la síntesi de la primera cadena de cDNA. Aquests productes de cDNA de cadena senzilla (sscDNA) s'amplifiquen mitjançant la reacció de la PCR amb primers específics del vector. Sovint, aquests darrers productes se sotmeten a una segona reacció d'amplificació amb encebadors interiors o *nested*, i posteriorment se subclonen en un vector per a la seva seqüenciació. Els clons amb exons capturats es poden reconèixer respecte als clons que contenen vector sense insert, per la diferent mobilitat electroforètica en un gel d'agarosa

В В B B B В В В В В В DNA del I cosmidi BamHI SA SD **BamHI** SD SA **BstXI BstXI** + mitja diana mitja diana Lloc múltiple de clonatge o MCS pSPL3 : BamHI BamHI SD SD SA SA **Bam**HI BamHI DNA Ш 1111 Subclonat MCS MCS RNA després del processament dels introns Exó atrapat després de la PCR

dels productes de la PCR amplificats amb els encebadors específics del vector (Auch i Reth, 1990; Buckler et al., 1991; Church et al., 1994).

Figura 4. Esquema de la técnica de la captura d'exons.

L'inconvenient d'aquesta tècnica és que no es poden capturar aquells gens minoritaris que no tenen introns, o bé aquells que necessiten factors específics de teixit

36 Introducció-

per a ésser processats. A més a més, els exons capturats representen sovint fragments petits de DNA <110 pb que dificulten enormement les anàlisis posteriors.

Diversos laboratoris han aplicat aquesta tècnica per aïllar exons específics de l'HC21 (Lucente *et al.*, 1995; Yaspo *et al.*, 1995; Chen *et al.*, 1996; Guimerà *et al.*, 1997; Ohira *et al.*, 1997; Dahmane *et al.*, 1998), que juntament amb la tècnica de la selecció de cDNA són les dues tècniques mitjançant les quals s'han-trobat més del 95% dels EST cercats de manera dirigida a l'HC21.





Figura 5. Esquema del vector pSPL3.

1.4.3. IDENTIFICACIÓ D'ILLES CPG O CpG ISLANDS

Aquest mètode rau en l'observació que una proporció elevada de promotors de vertebrats corresponents als gens anomenats de manteniment o *housekeeping gens* (60%) estan associats a regions molt riques en concentracions de dinucleòtids CpG no metilats, anomenades illes CpG (Bird, 1986). Aquestes illes CpG, anomenades també illes HTF, sovint tenen una mida d'1-2 kb. Malgrat que la seva concentració en el genoma és molt baixa, estan extremadament enriquides en els primers exons de gens, i sovint s'estenen fins al primer intró del gen.

Aquesta peculiaritat s'ha explotat per identificar gens. D'aquesta manera, les illes poden detectar-se mitjançant la presència de dianes de restricció per a enzims de tall infreqüent sensibles a la metilació (p.e. *Hpa*II) que reconeixen dianes que contenen dinucleòtids CpG no metilats, però no tallen següències CpG metilades. Així mateix, la identificació de tres o més dianes de restricció per a les endonucleases mencionades anteriorment, concatenades en un mateix cosmidi que conté DNA genòmic subclonat (i per tant no metilat), denota també la presència d'illes HTF (Lindsay i Bird, 1987). Una altra estratègia és la hibridació de genoteques de clons genòmics, utilitzant com a sonda radioactiva oligonucleòtids de mida curta (8 nt) (Estivill i Williamson, 1987). La sequència del DNA d'aquestes sondes és rica en nt G i C, i a més a més inclou dianes de restricció d'endonucleases que tallen preferentment en aquestes illes (p.e. Not1, SacII, BssHII, Apal). Una altra estratègia emprada és la digestió del DNA genòmic amb aquests enzims de restricció mencionats, els extrems dels quals són lligats a un oligonucleòtid específic o linker. Aquest DNA modificat és amplificat mitjançant la PCR, emprant un encebador complementari al cassette d'amplificació i un altre encebador consensus de les sequències Alu. Aquests productes són posteriorment subclonats i sequenciats (Valdés et al., 1994). A partir d'aquí hi ha dues estratègies per a aïllar els gens d'aquesta regió: 1) aquests fragments aïllats es fan servir com a sonda en anàlisis d'hibridació de genoteques de cDNA. Aquí el problema principal torna a ésser el de la verificació, ja que no sabrem en quin teixit s'expressarà el gen que s'ha trobat. En aquesta tècnica, però, aquest problema no és tan greu com en la tècnica de la captura d'exons, ja que la majoria dels gens sota el control de l'expressió d'una illa CpG sovint són d'expressió ubíqua, i 2) a partir de la seva sequència de DNA es pot dur a terme anàlisis amb programes de computadora com BLAST per a identificar similituds amb EST descrits (Altschul et al., 1990), o programes de predicció d'exons com GRAIL (Uberbacher i Mural, 1991).

1.4.4. TÈCNICA DE LA PCR D'ALU-SPLICE

Les seqüències de DNA de les subfamílies Alu pertanyen a la família de seqüències de DNA repetitiu SINE, i es troben distribuïdes per tot el genoma dels primats. Es calcula que el nombre de còpies en el genoma humà és aproximadament d'un milió (Deininger *et al.*, 1981), tenen una longitud d'uns 300 pb, i se'n troben una cada 4 kb (Hwu *et al.*, 1986). Aquestes observacions han servit per a dissenyar una tècnica d'aïllament d'exons (Fig. 6a) basada en la reacció de la PCR a partir de material genòmic (Fuentes *et al.*, 1997). Conseqüentment, s'han dissenyat encebadors a partir de la seqüència consensus de cada família d'elements Alu, tant a prop del 5' i del 3' dels extrems d'aquestes repeticions com sigui possible per a minimitzar el contingut de seqüència Alu en els productes de la PCR (Tagle i Collins, 1992).

Aquests encebadors donen l'especificitat de la sequència exónica, mentre que els encebadors Alu donen l'especificitat humana als productes de l'amplificació, ja que

38 Introducció-

aquestes seqüències repetitives només es troben en els primats i per tant no es troben en el ratolí. Per facilitar el posterior subclonatge direccional d'aquests productes de PCR, els dos tipus d'encebadors contenen una seqüència que és reconeguda per un enzim de restricció diferent. El producte de la PCR *Alu-splice* és un reguitzell de bandes que es subclonen en un vector de seqüenciació, prèviament tallat amb els dos mateixos enzims de restricció, les dianes dels quals estan presents en la seqüència del DNA dels encebadors utilitzats en la PCR. Només aquells productes que continguin els dos tipus d'encebador alhora seran subclonats (Fig. 6b), assegurant així l'origen humà de l'exó subclonat.

a



Figura 6. Esquema de la tècnica de la PCR Alu-Splice.

1.4.5. BIOCOMPUTACIÓ DE REGIONS SEQÜENCIADES

Les regions genòmiques senceres de les quals ja es disposa la sequència de nucleòtids poden analitzar-se per a la recerca de gens mitjançant algoritmes i fórmules matemàtiques complexes amb l'ajut de la computadora. Existeix un gran nombre d'aquests programes bioinformàtics, dels quals el programa més utilitzat per a aquesta finalitat és GRAIL (Uberbacher i Mural 1991), i representa una xarxa neural de programes de computadora que identifica exons codificants putatius en el DNA genòmic, a partir de la recerca de patrons típics dels gens.

Donat que en un futur no molt llunyà, i sota els auspicis del Projecte Genoma Humà, la totalitat d'aquest genoma estarà seqüenciat, aquest mètode representa una eina molt bona de futur.

1.4.6. TÈCNICA DE LA SELECCIÓ DE cDNA

La tècnica de selecció de cDNA o cDNA selection està basada en la selecció de fragments de cDNA sintetitzats a partir de material expressat de la cèl·lula o RNA (Lovett et al., 1991; Morgan et al., 1992). Aquí rau la diferència més important respecte als mètodes mencionats anteriorment que seleccionaven sequències genòmiques susceptibles d'ésser transcrites. La primera estratègia d'aquest mètode és escollir els teixits escaients. L'RNA missatger del teixit escollit és convertit a doble cadena de cDNA (dscDNA), i aquests són modificats amb l'addició d'uns adaptadors de DNA o linkers per a la seva porterior amplificació per la PCR. Aquests productes de cDNA amplificats per la PCR s'hibriden amb fragments immobilitzats de DNA de la regió del genoma d'interès (cosmidis, BAC, PAC o YAC). Per tenir un enriquiment més eficient, els cDNA amplificats per la PCR són prehibridats amb DNA de sequències genòmiques repetitives, amb DNA ribosomal i amb DNA de vector per tal d'emmascarar-los. Virtualment, tot cDNA de la regió d'interès s'hibridarà amb aquell cosmidi biotinilat que conté el gen que expressa aquest cDNA. Posteriorment, aquest cDNA específic d'una regió de l'HC21 serà inmobilitzat, mitjancant un imant, a les esferes d'estreptavidina que contenen nuclis de ferro. Després de rentar el complex d'hibridació es recupera el cDNA específic, el qual es torna a sotmetre a un segon cicle d'amplificació per la PCR per a dur a terme un segon cicle de selecció. Finalment, el cDNA seleccionat es subclona en un vector de seqüenciació (Fig. 7).

La tècnica de la selecció de cDNA pot arribar a enriquiments des de pocs milers fins a més de 100.000 vegades (Lovett, 1994) i és relativament insensible a la mida del fragment de DNA genòmic del qual es parteix. La gran limitació és tenir una bona genoteca de cDNA en la qual estiguin representades totes les espècies possibles de RNA. L'èxit de l'aplicació del mètode de la selecció de cDNA queda ben palès, atesa la gran quantitat de EST de l'HC21 descoberts mitjançant aquest mètode (Cheng *et al.*, 1994; Peterson *et al.*, 1994; Tassone *et al.*, 1995; Yaspo *et al.*, 1995, Ohira *et al.*, 1997; Guimerà *et al.*, 1997).Aquesta tècnica no proporciona directament "el gen", sinó una col·lecció de clons enriquits en sequències transcripcionals. Ara caldrà separar la palla de l'agulla.



Figura 7. Esquema de la tècnica de la selecció de cDNA.

En la Taula 5 es detallen les principals diferències i similituds entre els diferents mètodes d'identificació de gens.

Taula 5. Diferents esbiaixos de les tècniques de selecció de cDNA, de la captura d'exons, de biocomputació i de la PCR Alu-Splice.

,	Tècnica de la	Tècnica de la	Biocomputació	Tècnica de la
	captura	selecció de		PCR Alu-Splice
	d'exons	cDNA		_
Depèn de l'expressió	No	Sí	No _	No
temporal i espaial				
del gen				
Depèn de senyals	Sí	No	Si	Sí
consens dels gens				
Mida dels EST	<120 pb	300-500 pb	100-800 pb	<120 pb
obtinguts				
Mida del fragment	limitada a	Il·limitada	Il·limitada	Il·limitada
genòmic diana	poques kb			
Nombre d'exons	3	independent	independent	2
mínims				
Requereix presència	Sí	No	No és	Si
dels introns			necessària	
Verificació	Molt	Рос	Molt	Molt
	problemàtica	problemàtica	problemàtica	problemàtica
Redundància	Petita	Elevada, però	No	Elevada,
		no representa		representa un
		un gran		gran problema
		problema		
Totalitat	No	No	No	No
Permet buscar un	No	Sí	No	No
gen específic d'un				
teixit determinat				

En el nostre cas, vàrem optar pel mètode de la selecció de cDNA perquè és l'unic dels quatre mètodes mencionats que parteix de cDNA. Aquesta particularitat, lluny d'ésser una dificultat, representava un avantatge, ja que ens va permetre cercar gens que podien estar relacionats amb la formació i desenvolupament del SNC, i més concretament, cercar els gens involucrats en la discapacitat intel·lectual de la DS.

1.5. GEN MINIBRAIN

1.5.1. FENOTIP MINIBRAIN EN DROSOPHILA

En Drosophila, la proliferació i diferenciació dels neuroblasts de l'embrió produeix la formació del sistema nerviós de la larva. Així doncs, la proliferació postembrionària dels

neuroblasts durant el desenvolupament larval inicia la formació del sistema nerviós adult. (Hartenstein i Campos-Ortega 1984).

Les mosques de la fruita amb mutacions en el gen *minibrain (mnb)* es caracteritzen per una marcada reducció del volum dels lòbuls òptics (50-70%) i dels hemisferis centrals del cervell (40-50%). El fenotip dels mutants *mnb* mostren que no tenen grans alteracions en l'arquitectura neuronal del cervell de *Drosophila*, i suggereixen que *mnb* juga un paper molt important en la regulació dels diferents tipus de cèl·lules neuronals, més que no pas en l'arquitectura neuronal. En el cervell de la larva, els neuroblasts estan organitzats en capes compactes de cèl·lules que inicien divisions cel·lulars asimètriques, a partir de les quals es generen les cèl·lules mares ganglionaries i eventualment neurones (White i Kankel, 1978). La pèrdua de neurones en els lòbuls òptics i en el cervell de les mosques *mnb* comporta canvis de comportament respecte a les mosques salvatges o *wild type*. Així, s'observa que les mosques *mnb* tenen el sentit de l'olfacte disminuït, una activitat locomotora alterada, una capacitat de resposta més lenta a determinats estímuls i una fixació visual més pobra (Tejedor *et al.*, 1995). Aquests mutants també requereixen molt més temps per al seu desenvolupament (Heidenreich, 1982).

Cal fer esment també que els centres proliferatius dels neuroblasts estan alterats en els cervells de les larves dels mutants mnb. Aquesta alteració està caracteritzada per una distribució espaial irregular dels neuroblasts, i afecta la descendència neuronal. D'aquesta manera, la reducció de la mida del cervell té la seva principal causa en la reducció del nombre de neurones a causa d'un dèficit de proliferació, més que no pas a una neurodegeneració (Tejedor et al., 1995), encara que aquest darrer procés pot tenir lloc de manera secundària per raó d'un contacte insuficient entre les cèl·lules. Aquesta darrera condició és un pre-requisit per al manteniment de moltes neurones. D'aquesta manera, no és pot descartar la possibilitat que el dèficit de neurones en l'SNC sigui causat també, indirectament, per la formació de neuroblasts anormals, ja que els mecanismes de control del desenvolupament del sistema nerviós asseguren que les cèl·lules amb un desenvolupament o una connectivitat inapropiada siguin ràpidament eliminades per fagocitosis sense donar lloc a cap progenie (Glücksmann, 1951: Saunders, 1956; Clarke, 1990). D'aquesta manera, l'expressió aberrant d'aquells gens que actuen en la proliferació neuronal, com és el cas del gen mnb, podrien causar, a banda del dèficit proliferatiu, una subsequent degeneració neuronal. En aquest sentit, es desconeix si aquesta manca de neurones és consequència directa o indirecta de mutacions en el gen mnb (Tejedor et al., 1995).

El gen *minibrain* de *Drosophila* codifica per a una proteïna quinasa de la família serina-treonina. L'estructura genòmica comprèn 25 kb de DNA i codifica per a tres

trànscrits alternatius (trànscrit A: 5.498 nt, trànscrit B: 4.238 nt i trànscrit C: 4.440 nt). Aquests trànscrits tenen seqüències de DNA divergents en la part 3' codificant, generant diferents isoformes de la proteïna (Fig. 8) (les masses moleculars calculades són iso-A: 89 kD, iso-B: 59,1 kD i iso-C: 59,5 kD). L'extrem amino-terminal conté dos senyals potencials de translocació nuclear. El domini catalític conté subdominis i elements de seqüència típics de la família de les quinases (Hanks *et al.*, 1988; Taylor *et al.*, 1983). En el seu extrem carboxi-terminal conté una seqüència rica en glicines, alanines i serines (regió GAS), típica de la família de teixits larvals i en la neurogènesi (Bourouis *et al.*, 1989, 1990). La quinasa de *minibrain* comparteix una gran similitud en la seqüència de DNA amb altres proteïnes quinasa conegudes que intervenen en la regulació del cicle cel·lular (Sherr, 1993). Totes aquestes observacions suggereixen que la quinasa de *mnb* en *Drosophila* té una funció molt important en la producció i correcta proliferació de la descendència dels neuroblasts i en el control de diferenciació d'aquests.



Figura 8. Estructura genòmica del gen mnb de Drosophila. Esquema de Tejedor et al., (1995).

1.5.2. GEN Dyrk DE RATA

Paral·lelament a la descoberta de l'homòleg humà del gen *mnb* de *Drosophila* (Guimera *et al.*, 1986), Kentrup *et al.* (1996) descobriren el gen de rata homòleg de *mnb*. Aquest gen múrid anomenat *Dyrk* (o també *Mnbh*), codifica per a una proteïna quinasa de la família serina-treonina de 763 aminoàcids. Dyrk catalitza tant l'autofosforilació dels seus residus de serina i treonina com dels seus residus de tirosina, així com la fosforilació de residus tirosina d'histones. La seva activitat quinasa depèn de la presència dels residus tirosina entre els subdominis VII i VIII que conformen el domini activador (Kentrup *et al.*, 1996).

1.6. GENS, FUNCIÓ I FENOTIP

La identificació de tots els gens del cromosoma 21, l'elucidació de llur funció i els seus efectes en condició trisòmica permetran la seva associació amb alteracions fenotípiques específiques de la DS. El coneixement de la fisiopatologia molecular de la DS ens ha de permetre:

1) Conèixer si la trisomia d'un determinat fragment és suficient per a motivar un determinat tret fenotípic o pel contrari, allò veritablement important és una combinació al·lèlica particular.

2) Avaluar el paper dels efectes ambientals en l'expressió fenotípica.

3) Elucidar si existeix efectes epistàtics d'un *locus* modificador, que pot estar en el cromosoma 21 o no.

4) Conèixer si existeix un efecte d'empremta parental o imprinting.

5) Conèixer si hi ha efectes ambientals que puguin suposar un risc per a l'aparició de la trisomia 21.

6) Introduir mètodes de diagnòstic millors, més ràpids i més fiables.

7) Cercar noves possibilitats terapèutiques per a pal·liar, curar o prevenir l'aparició de determinats trets clínics de la DS.

– 2. OBJECTIUS

•

.

L'objectiu general d'aquesta tesi ha estat el de contribuir al coneixement de les bases moleculars de la síndrome de Down, mitjançant la construcció d'un mapa transcripcional de tres regions del cromosoma 21 humà i la caracterització d'un dels gens d'aquest nou mapa. Aquest objectiu general ha quedat definint en els següents objectius específics:

- Construcció d'un mapa transcripcional de la regió 21q22.2-22.3

1. Construcció d'una genoteca de cosmidis a partir de 5 YAC (72H9, 336G11, 238B1, 221B9 i 552A3) específics de tres regions del cromosoma 21 humà: la regió crítica de la síndrome de Down, una regió distal i una altra proximal (2,2 Mb).

- 2. Construcció d'un mapa contigu de cosmidis solapats de les tres regions.
- 3. Construcció d'una genoteca de cDNA de cervell fetal humà.
- 4. Aplicació de la tècnica de selecció de cDNA o cDNA selection a les 3 regions esmentades.
- 5. Caracterització i anàlisi dels EST identificats.
- 6. Construcció d'un mapa transcripcional.

- Aïllament i caracterització d'un nou gen humà: Minibrain (MNBH)

- 1. Caracterització de la regió codificant del gen MNBH.
- 2. Localització cromosòmica i molecular del gen MNBH en el cromosoma 21 humà.
- 3. Caracterització dels extrems 5'-UTR i 3'-UTR o full lenght del cDNA de MNBH.
- 4. Identificació dels límits de les regions dels exons i dels introns.
- 5. Anàlisi detallat de l'estructura genòmica de MNBH.

6. Caracterització dels possibles inicis de transcripció del gen MNBH, llurs promotors putatius i llur expressió diferencial en teixits humans.

7. Caracterització de les possibles variants transcripcionals de la regió codificant i identificació de les possibles isoformes de la proteïna MNBH.

8. Anàlisi dels possibles dominis funcionals de la proteïna deduïda de MNBH, i comparació amb les seves homòlogues de Drosophila i de rata.

9. Determinació del patró d'expressió del gen Mnbh en cervell adult de ratolí, mitjançant la tècnica d'hibridació in situ.

10. Quantificar la sobreexpressió del gen *MNBH* en teixit nerviòs de persones amb la síndrome de Down i en el ratolí Ts65Dn.

– 3. MATERIAL I MÈTODES

...

•

49

3.1. OBTENCIÓ D'ÀCIDS NUCLEICS

3.1.1. PRECIPITACIÓ D'ÀCIDS NUCLEICS

3.1.1.1. PRECIPITACIÓ AMB ETANOL

En presència de concentracions de 0,1 M a 0,5 M de cations monovalents, l'etanol indueix una transició estructural en les molècules d'àcids nucleics, les quals s'agreguen i precipiten (Eickbush i Moudrianakis, 1978). Els àcids nucleics s'han precipitat amb etanol al 96% i s'han rentat amb etanol al 70% per a extreure les sals i les molècules orgàniques petites.

Procediment

1. Afegir 1/10 volums d'acetat de sodi 3M, pH 5,2, i 2-2,5 volums d'etanol al 96% a la solució aquosa que conté el DNA. Mesclar les solucions. La precipitació es pot dur a terme a 4°C o bé a -70°C. Per a quantitats de DNA >50 µg/ml, la precipitació és instantània a temperatura ambient.

2. Centrifugar >2500 x g, 15 min a 4°C. Descartar la solució aquosa.

3. Afegir 1 ml d'etanol al 70% al precipitat de DNA. Incubar 10 min a temperatura ambient per a rentar les sals de la precipitació. Si les molècules de DNA són de mida petita (<200 pb) es pot utilitzar etanol al 95%.

4. Centrifugar >2500 x g, 15 min a 4°C. Descartar la solució aquosa i assecar el precipitat de DNA.

5. Dissoldre el DNA en H₂O o en tampó TE pH 8,0. Per facilitar la dissolució del DNA cal guardar la concentració final d'aquest a <1 mg/ml.

3.1.1.2. PRECIPITACIÓ AMB ISOPROPANOL

En aquells casos en que el volum aquós amb el DNA a precipitar era molt gran s'ha utilitzat isopropanol en comptes d'etanol. El procediment de precipitació és similar a l'anterior, però el DNA es precipita amb 1 volum d'isopropanol. Atès que el volum d'isopropanol emprat és la meitat que el d'etanol, aquest mètode ha estat molt útil per precipitar el DNA de les maxipreparacions del DNA dels cosmidis abans de la seva purificació per gradient de CsCl. Cal esmentar que l'isopropanol és menys volàtil que l'etanol i triga més temps en assecar-se per evaporació, i que algunes sals són menys solubles en isopropanol i precipiten amb els àcids nucleics, de manera que és necessari efectuar diversos rentats amb etanol al 70% per eliminar les sals contaminants.

En altres casos s'ha emprat també 1 volum de sec-butanol (2-butanol), el qual redueix el volum de la solució aquosa, concentrant el solut (fase inferior) abans de procedir a la precipitació del DNA amb isopropanol.

3.1.2. PREPARACIÓ DEL DNA

3.1.2.1. EXTRACCIÓ DE DNA GENÒMIC A PARTIR CÈL·LULES SANGUÍNIES Hem utilitzat sang perifèrica per obtenir grans quantitats de DNA genòmic per als estudis de transferència de *Southern*. Mitjançant el mètode que es presenta a continuació, la qualitat del DNA genòmic obtingut ha estat d'una gran puresa amb una relació $DO_{260}/DO_{280} > 1,8$; això ens indica que la solució aquosa de DNA purificada estava lliure de contaminants (les preparacions amb un 50% DNA i 50% proteïna tenen una relació DO_{260}/DO_{280} de 1,5). El rendiment típic de la preparació de DNA obtingut ha estat de 2 mg de DNA a partir de cada 1 ml de sang perifèrica. El DNA purificat ha tingut sempre una mida superior a 100 kb.

Procediment

1. Partir de 15 ml de sang perifèrica. Afegir-la a un tub falcon de 50 ml.

2. Afegir sèrum fisiològic fins a 50 ml, mesclar per inversió i centrifugar a 800 x g durant 10 min a 4°C.

3. Descartar el sobrenedant. Afegir tampó de lisi dels eritròcits fins a 50 ml i agitar suaument. Centrifugar a 2.500 x g, 15 min a 4°C i descartar el sobrenedant.

4. Afegir 3 ml de la solució de digestió dels eritròcits, 200 μ l de SDS 10%, i 500 μ l de solució de proteïnasa K. Agitar vigorosament fins a obtenir un aspecte homogeni. Aquest pas es basa en l'activitat proteolítica de la proteïnasa K i l'habilitat desnaturalitzant del detergent iònic SDS. A aquesta digestió s'hi afegeix EDTA per inhibir les DNases.

5. Incubar les mostres a 37°C amb agitació durant 12-18 hores.

6. Incubar les mostres a 37°C, com a mínim 1 hora, amb 10 μl de RNasa A.

7. Afegir un volum igual de fenol tamponat a pH 8,0 a la solució de DNA a purificar. Agitar durant 10 min per a una completa desproteïnització. L'extracció amb fenol cal repetir-la fins que no apareixi cap precipitat en la interfície aquosa-orgànica.

8. Centrifugar a 2.700 x g, 15 min a temperatura ambient per a extreure el sobrenedant aquós que conté el DNA.

9. Afegir un volum igual de fenol/cloroform/isoamílic (25:24:1) al sobrenedant. Agitar durant 10 min.

10. Centrifugar a 2.700 x g, 5 min a temperatura ambient per a extreure el sobrenedant aquós.

11. Afegir un volum igual de cloroform/isoamílic (24:1) al sobrenedant. Agitar durant 10 min.

12. Centrifugar a 2.700 x g, 5 min a temperatura ambient.

13. Precipitar el DNA del sobrenedant amb etanol al 96%. La precipitació és instantània i si la concentració de DNA és superior a 50-100 μ g/ml sovint s'obté una medusa de DNA.

14. Centrifugar 2.500 x g. Dissoldre el precipitat de DNA en tampó TE.

Guardar el DNA a 4°C. El DNA no hauria d'ésser guardat⁻⁻a -20 °C perquè a aquesta temperatura hi ha un gran nombre de trencaments, tant de cadena senzilla com de cadena doble.

Material i solucions

- Tampó de digestió dels eritròcits:

5 mM MgCl₂,

20 mM Tris HCl, pH 7,5.

- Solució de lisi dels eritròcits:

400 mM NaCl,

2 mM EDTA,

10 mM Tris HCl, pH 8,2.

- Solució de proteïnasa K:

0,2 mg/ml proteïnasa K,

1% SDS,

2 mM EDTA. Ajustar el pH a 8,0.

La proteïnasa K és molt làbil i cal afegir-la immediatament abans d'utilitzar-la.

- TE 10/0,2: 10 mM Tris HCl, pH 7,5, 0,2 mM EDTA.

- RNasa A: 10 mg/ml.

3.1.2.2. EXTRACCIÓ DE DNA A PARTIR DE GELS

L'extracció de DNA a partir d'una banda de DNA d'un gel ens ha estat motl útil per recuperar una banda específica d'una determinada mida discriminada per electroforesi.

3.1.2.2.1. Gels d'agarosa

El mètode emprat majoritàriament en aquesta tesi per purificar una banda de DNA d'un gel d'agarosa ha estat una modificació del procediment descrit per Vogelstein i Gillespie (1979). Representa un mètode ràpid, no tòxic i que proporcina una molt bona qualitat de DNA. Aquest mètode s'ha emprat per a la purificació d'una banda de DNA tallada a partir d'un gel d'agarosa per al seu posterior clonatge, o bé per a la purificació dels

productes de la PCR de cada EST obtinguts per la tècnica de la selecció de cDNA, per al seu marcatge i posterior localització en el mapa de YAC i de cosmidis.

Procediment:

1. Fer una electroforesi en gel d'agarosa a la concentració escaient.

2. Córrer el gel a un voltatge baix (30 volts) per a una millor resolució.

3. Tenyir el gel amb bromur d'etidi. Visualitzar les bandes amb llum ultraviolada d'ona llarga. L'exposició del DNA a la llum ultraviolada cal ésser tan curta com sigui possible, a fi i efecte de no degradar el DNA.

4. Tallar la banda d'agarosa que ens interessa amb una ganiveta estèril.

5. Afegir 3 volums (v/p) de NaI a la solució de DNA. Incubar a 50°C durant 15 min.

6. Quan l'agarosa està ben dissolta, afegir 5 μ l de la suspensió de glass milk. Incubar 5 min a temperatura ambient.

7. Centrifugar el complex DNA/glass milk a 1.500 x g, 5 seg. Descartar el sobrenedant.

8. Rentar el sediment o pellet tres vegades amb 500 µl de la solució de rentat.

9. Dissoldre el darrer sediment en tampó TE, pH 8,0, per obtenir una concentració final de DNA de 0,5 μ g/ μ l. Incubar 3-5 min a 50°C. Centrifugar a 2.500 x g i recuperar el sobrenedant que conté el DNA.

Darrerament s'ha utilitzat el kit "GENENCLEAN II® Kit", obtenint els mateixos resultats.

Material i solucions

- Solució de NaI 6M:

0,75 g de NaSO₃ en 40 ml H₂O,

45 g de NaI.

Filtrar amb paper Whatman 3MM o de nitrocel·lulosa. Guardar a la foscor.

- Solució de rentat:

20 mM Tris HCl, pH 7,4,

1 mM EDTA,

100 mM NaCl,

Afegir 1 volum igual d'etanol al 96%. Guardar a <0°C.

- Suspensió de glass milk:

Glass milk: (BIO 101, Inc).

Transferir un volum de 200-300 μ l de 100 μ m de glass milk a un volum igual d'H₂O.

- Kit de purificació GENECLEAN II® Kit, (# 1001-400 de BIO 101).

Ocasionalment s'ha emprat el kit de purificació de bandes "Qiaquick Gel Extraction Kit", (# 28704 de Qiagen). La quantitat de DNA recuperada amb aquest kit ha estat menor i hem detectat que el DNA recuperat amb aquest kit disminueix l'eficiència d'algunes reaccions posteriors, com en les reaccions de lligament. Una alternativa a aquesta darrera aproximació ha estat la purificació amb fenol/cloroform:

Procediment

- 1. Dissoldre l'agarosa 10 min a 65°C.
- 2. Afegir un volum (v/p) de fenol, 15 min a 4° C.
- 3. Centrifugar 2.500 x g, 15 min.
- 4. Recuperar el sobrenedant que conté el DNA i repetir els passos 2-3.
- 5. Afegir un volum de fenol/cloroform.
- 6. Mesclar i centrifugar a 2.500 x g, 10 min.

7. Afegir un volum de cloroform al sobrenedant. Mesclar i tornar a centrifugar amb les mateixes condiciones que en el pas 3.

8. Precipitar el DNA del sobrenedant amb dos volums d'etanol al 96%. Rentar les sals amb etanol al 70%. Dissoldre el precipitat de DNA en tampó TE.

3.1.2.2.2. Gels d'acrilamida

S'ha utilitzat aquest tipus de gel per discriminar fragments de DNA de mida molt petita (<100 pb). La quantitat de DNA que hem recuperat mitjançant aquesta tècncia és bastant inferior que en el mètode mencionat anteriorment, però suficient com per a ésser clonat o utilitzat en la PCR. La qualitat del DNA purificat ha estat bastant pur.

Procediment

- 1. Tallar amb una ganiveta estèril la banda de DNA d'interès del gel.
- 2. Afegir 30 μ l de H₂O destil·lada. Incubar >12 hores a temperatura ambient.
- 3. Recuperar l'H₂O amb el DNA.

3.1.2.3. OBTENCIÓ DE DNA PLASMÍDIC A PARTIR DE BACTÈRIS.

3.1.2.3.1. Minipreparació

Aquest mètode està basat en el procediment de lisi alcalina (Bimboim i Doly, 1979). El rendiment típic que s'ha obtingut és de 2-5 μ g de DNA per 3 ml de cultiu de cèl·lules. La puresa del DNA purificat ha estat prou bona per a la majoria d'aplicacions que hem aplicat posteriorment (sequenciació, digestió, etc.) sense la necessitat de fer una extracció posterior amb fenol-cloroform.

Procediment

1. Inocular 5 ml de medi LB estèril amb una colònia senzilla, crescuda en placa d'LB amb l'antibiòtic escaient. Créixer fins a la seva saturació a 37°C, o sia, 1x10⁹

56 Material i mètodes

cèl·lules/ml \circ DO = 1 (normalment s \circ n 16 hores) amb una agitaci \circ de 200 rpm per procurar un bon airejament.

2. Centrifugar 4 ml de cèl·lules a 2.500 x g, 30 seg a 4°C.

3. Dissoldre el sediment en 100 μ l de GTE, afegir 500 mg/ml de lisozim. Incubar durant 5 min a temperatura ambient.

4. Afegir 200 μl de la solució de NaOH/SDS (l'NaOH desnaturalitza tant el DNA com les proteïnes; l'SDS desnaturalitza les proteïnes bacterianes i fa més soluble els fosfolípids). Incubar solament 5 min a temperatura ambient, ja que llargues exposicions en condiciones alcalines desnaturalitzen el plasmidi irreversiblement. Aquesta forma de plasmidi desnaturalitzat migra més ràpidament en un gel d'agarosa i és resistent al tall d'enzims de restricció.

5. Afegir 150 μ l d'acetat de potassi (reaparella ràpidament i correctament el DNA plasmidi circular i no precipita), 5 min en gel.

6. Centrifugar 3 min a 2.700 x g, 30 seg. Transferir el sobrenedant a un nou tub.

7. Precipitar el DNA amb etanol al 96%. Rentar posteriorment amb etanol al 70%. Material i solucions

- Solució GTE:

50 mM glucosa,

25 mM Tris HCl, pH 8,0,

10 mM EDTA.

Esterilitzar a l'autoclau. Afegir RNasa fins a tenir una concentració final de 100 μ g/ml.

- Solució de NaOH/SDS:

200 mM NaOH,

1% (p/v) dodecilsulfat de sodi (SDS).

Preparar immediatament abans de fer-ho servir.

- Solució d'acetat de potassi:

3,0 M d'acetat de potassi,

Afegir pastilles de KOH fins a obtenir un pH 5,5.

- Medi de Lúria Bertani (LB):

10 g triptona,

5 g extracte de llevat,

10 g NaCl.

Ajustar el pH a 7,5 amb NaOH 1N.

- Plaques d'agar LB:

Afegir 15 g/litre de bacto-agar al medi LB. Esterilitzar a l'autoclau.
Tampó TE 1/01:
10 mM Tris HCl, pH 8,0,
1 mM EDTA.
Additius
Ampicil·lina:
Ajustar amb H₂0 la concentració final a 50 μg/ml de medi.
Tetraciclina:
Ajustar amb etanol la concentració final a 12,5 μg/ml.
Kanamicina:
Ajustar amb H₂0 la concentració final a 50 μg/ml.

3.1.2.3.2. Maxipreparació

Per a obtenir grans quantitats de DNA de cosmidi o de plasmidi vàrem emprar també el mètode de la lisi alcalina. Aquest mètode ens ha permés obtenir grans quantitats de DNA cru de cadascun dels cosmidis de la genoteca generada a partir de les tres regions cromosòmiques esmentades en l'apartat 3.4.1., per a la seva posterior purificació mitjançant centrifugació en gradient de CsCl/bromur d'etidi. Darrerament s'ha emprat el kit "Qiagen plasmid kit" (# 12236 de Qiagen) per a la maxipreparació de DNA de plasmidi, i s'ha obtingut una puresa prou bona per a la majoria d'aplicacions, però insuficientment pur per a la seva aplicació posterior a la tècnica de la selecció de cDNA.

Procediment

1. Inocular 500 ml de medi LB estèril amb l'antibiòtic escaient en un matràs de 2 litres (per maximitzar l'airejament) amb una colònia senzilla crescuda en placa d'LB. Créixer fins a saturació ($DO_{600} = 4$) a 37°C, 16 hores a 250 rpm.

2. Centrifugar les cèl·lules a 6.000 x g, 10 min a 4°C.

3. Dissoldre el sediment amb 4 ml de GTE. Afegir 1 ml de 25 mg/ml de lisozim (digereix la paret cel·lular). Incubar 10 min a temperatura ambient.

4. Afegir 8 ml de la solució NaOH/SDS i incubar 10 min a temperatura ambient.

5. Afegir 7,5 ml d'acetat de potassi, incubar 10 min gel.

6. Centrifugar a >20.000 x g, 10 min a 4°C. Transferir el sobrenedant a un nou tub.

7. Precipitar el DNA amb 0,6 volums d'isopropanol, mesclar per inversió i incubar a temperatura ambient de 5 a 10 minuts perquè no precipitin les sals.

8. Centrifugar a 15.000 x g, 10 min a temperatura ambient.

9. Rentar el precipitat de DNA amb 2 ml d'etanol al 70%.

10. Centrifugar a 15.000 x g. Assecar el precipitat de DNA.

11. Dissoldre el DNA en 2 ml de tampó TE pH 8,0.

Aquesta purificació és prou bona per a la majoria de reaccions posteriors. A partir d'aquí, de vegades s'ha fet una purificació posterior amb fenol/cloroform/isoamílic (25:24:1), com és el cas de la preparació dels vectors de clonatge. Aquest DNA cru s'ha utilitzat també per a la seva purificació per ultracentrifugació en gradient de CsCl i amb bromur d'etidi. Aquest darrera opció és del tot imprescindible per purificar el DNA d'aquells cosmidis que seran emprats posteriorment en la tècnica de selecció de cDNA.

3.1.2.4. MINIPREPARACIÓ DEL DNA DE LLEVAT

Procediment

1. Créixer les cèl·lules de llevat en 40 ml de medi AHC fins a saturació, amb una agitació de 220 rpm a 30°C (3 dies aproximadament o fins obtenir una $DO_{600} = 1$, una $DO_{600} = 1$ equival a 3 x 10⁷ cèl·lules de llevat/ml).

- 2. Centrifugar a 1.000 x g durant 5 min.
- 3. Dissoldre el sediment en 3 ml de solució R.
- 4. Afegir 25 μ l de 10 mg/ml de liticasa. Incubar durant 60 min a 37°C.
- 5. Centrifugar a 400 x g durant 5 min.
- 6. Dissoldre el sediment amb 5 ml de Tris 50 mM pH 7,4, EDTA 20 mM.
- 7. Afegir 0,5 ml de SDS al 10%. Mesclar.
- 8. Incubar a 65°C durant 30 min.
- 9. Afegir 1,5 ml d'acetat de potassi 5M. Incubar durant 60 min a 4°C.
- 10. Centrifugar 15.000 x g durant 10 min.

11. Transferir el sobrenedant a un tub que conté 2 volums d'etanol a temperatura ambient. Mesclar i centrifugar a 1.000 x g durant 15 min.

- 12. Dissoldre el precipitat en 2 ml tampó TE.
- 13. Afegir 150 µl de RNasa 1 mg/ml. Incubar durant 30 min a 37°C.
- 14. Purificar el DNA amb fenol/cloroform. Repetir la purificació una vegada més. Precipitació amb etanol al 96%. Rentar amb etanol al 70%.
- 15. Dissoldre el precipitat de DNA en TE.

Material i solucions

- Solució R:
- 0,9 M Sorbitol,
- 0,1 M EDTA, pH 7,5 (el pH és molt important).
- Medi de llevat AHC:
- 6,7 g de base nitrogenada de llevat sense aminoàcids,
- 10 g àcid de caseïna hidrolitzada,
- 20 mg hemisulfat d'adenina,
- Ajustar el pH a 5,8 amb uns 70 µl de 12 N HCl,

- Afegir H_2O fins un volum final de 1 litre.
- Esterilitzar a l'autoclau.
- Afegir 50 ml de D-glucosa al 40% (p/v).
 - Liticasa:
 - 10 mg/ml en H2O bidestil·lada i autoclavada.

3.1.3. PREPARACIÓ DE RNA

Per tal de tenir una preparació de RNA prou pura com per a ésser posteriorment utilitzat en la tècnica de la selecció de cDNA, hem optat pel procediment descrit per Chomczynski i Sacchi, (1987). Atesa la resistència a la inactivació de les RNases, tot el material emprat ha estat tractat prèviament amb H₂O de DEPC (dietil-pirocarbonat), escalfat a 200°C i esterilitzat a l'autoclau. Les RNases endògenes s'han inactivat mitjançant la desproteïnització amb fenol i la inhibició amb tiocianat de guanidina. El càlcul de la concentració de RNA purificat a partir d'una solució aquosa s'ha realitzat segons s'explica en l'apartat 3.1.3.2.

Darrerament s'ha emprat el kit d'extracció de RNA "RNeasy Midi" (# 75144 de Qiagen) per fer les transferències de *Northern* amb uns resultats prou bons. Una altra forma amb la que s'ha obtingut l'RNA ha estat mitjançant el kit ULTRASPECTM RNA de Biotech Laboratories), però l'RNA obtingut fou d'una qualitat inferior (més degradat).

3.1.3.1. EXTRACCIÓ DE RNA TOTAL A PARTIR DE TEIXIT

Procediment

1. Afegir 1 ml de solució desnaturalitzant per cada 100 mg de teixit i homogeneïtzar en un homogeneïtzador de tefló. El teixit cal que sigui ràpidament congelat en nitrogen líquid després de la seva extracció.

2. Transferir l'homogeneïtzat en un tub de polipropilè de 5 ml. Afegir 1 ml de fenol saturat amb H₂O de DEPC. Mesclar i afegir 0,2 ml de cloroform/isoamílic (48:1). Mesclar i incubar la suspensió durant 15 min a $0-4^{\circ}$ C. També es pot utilitzar cloropropà (el cloropropà és menys tòxic que el cloroform i el seu ús per a la separació de fases decreix la possibilitat de contaminar la preparació de RNA amb DNA).

3. Centrifugar a 10.000 x g, 20 min a 4°C. Transferir la fase aquosa a un nou tub, o sia, la fase superior (que és la fase que conté l'RNA lliure de DNA i proteïnes).

4. Precipitar l'RNA afegint-hi 1 ml (1 vol) d'isopropanol 100%. Incubar les mostres durant 30 min a -20°C. Centrifugar a 10.000 x g, durant 10 min a 4°C. Descartar el sobrenedant.

5. Dissoldre el precipitat de RNA en 0,3 ml de solució desnaturalitzant.

6. Precipitar l'RNA amb 0,3 ml (1 vol) d'isopropanol 100% a -20°C durant 30 min. Centrifugar a 10.000 x g, 10 min a 4°C. Descartar el sobrenedant.

7. Rentar el precipitat de RNA amb etanol al 76%. Incubar 10-15 min a temperatura ambient per dissoldre petites quantitats residuals de guanidina en el precipitat.

8. Centrifugar a 10.000 x g, durant 5 min a 4° C, i descartar el sobrenedant. Assecar el precipitat de RNA.

9. Dissoldre el precipitat de RNA en 200 μ l d'H₂O tractada amb DEPC. Incubar 10-15 min a 55-60°C. Guardar l'RNA en H₂O a -70°C.

10. Quantificar l'RNA, en una dilució 1/200 amb H₂O alcalina pH >7,5. Els pH àcids afecten l'espectre d'absorció de llum ultraviolada de l'RNA i decreix considerablement la relació A_{260}/A_{280} (caldrà sinó, ajustar el pH de l'aigua). Comprovar la integritat física de l'RNA en un gel d'agarosa desnaturalitzant (veure apartat 3.3.1.2.1.).

Material i solucions

- H₂O de DEPC 0,1%:

El dietil-pirocarbonat (DEPC) és un potent inhibir de les RNases mitjançant modificacions covaments en la seva molècula.

1 ml de DEPC a un litre d' H_2O .

Incubar durant 30 min a 37°C.

Esterilitzar en l'autoclau.

3.1.3.2. DETERMINACIÓ DE LA QUANTITAT DE RNA

La relació A_{260}/A_{280} de l'RNA cal que sigui entre 1,9 i 2,0 en 10 mM Tris HCl, pH 7,5, ja que els contaminants proteics absorveixen a una longitud d'ona de 280 nm i fan baixar aquesta relació. Hem afegit tampó perquè la relació A_{260}/A_{280} pot variar segons el pH. Fins i tot, l'addició de l'RNA pot fer varia aquesta relació. Una A_{260} de 1,0 correspon a una concentració de 40 µg/ml de RNA de cadena senzilla.

La quantitat de RNA (μ g/ml) = 40 x A₂₆₀ x factor dilució.

L'RNA aïllat es guarda a -20°C o -70°C. En aquestes condicions no s'observa degradació destacable fins a 1 any després.

Material i solucions

- Solució desnaturalitzant:

293 ml d'H₂0,

17,6 ml de citrat de sodi 0,75M, pH 7,0,

26,4 ml de N-laurosil de sarcosina 10% (p/v).

Afegir 250 g tiocianat de guanidina i mesclar a 60-65°C fins a dissoldre. Abans de fer-la servir afegir 0,35 ml de 2-beta-mercaptoetanol per cada 50 ml de la solució anterior.

- Acetat de sodi 2 M:

Afegir 16,42 g d'acetat de sodi (deshidratat) a 40 ml H_2O i 35 ml d'àcid acètic. Ajustar la solució a pH 4,0 amb àcid acètic. Afegir H_2O fins a obtenir un volum final de 100 ml.

- H₂O saturada amb fenol (fenol àcid):

Dissoldre 100 g de cristalls de fenol en H_2O a 60-65°C. Aspirar la fase aquosa (fase superior) i guardar a 4°C.

- H₂O de DEPC: afegir DEPC al 0,01% en l'H₂O bidestil·lada. Aquesta H₂O s'incuba durant 16 hores a 37° C amb agitació.

Material i solucions

- Tractament de solucions amb DEPC:

Afegir 0,2 ml de DEPC a 100 ml de la solució a tractar. Incubar durant 12 hores a 37°C, i esterilitzar la solució en l'autoclau per a inactivar la resta de DEPC.

3.1.3.3. EXTRACCIÓ DEL DNA CONTAMINANT DE LA PREPARACIÓ DE RNA

Petites contaminacions de DNA en la preparació de RNA poden determinar posteriorment l'èxit o fracàs de la tècnica de la selecció de cDNA. Per aquesta raó s'ha procedit a la seva extracció de la preparació de RNA.

Procediment

1. Dissoldre l'RNA en 50 µl de tampó TE.

2. Afegir:

10 μ l de MgCl₂ 100 mM /DTT 10 mM (El DTT trenca els ponts disulfur tal com el ß-mercaptoetanol),

0,2 µl de DNasa de 2,5 mg/ml lliures de RNasa,

0,1 μ l de l'inhibidor de la ribonucleasa (25-50 U/ μ l),

39,7 µl de tampó TE.

3. Incubar 15 min a 3°C.

4. Aturar la reacció afegint 25 µl de la solució stop de la DNasa.

5. Purificar l'RNA amb una extracció amb fenol-cloroform-isoamílic (25:24:1). Precipitar amb etanol al 96% i rentar amb etanol al 70%.

Materials i solucions

Solució stop de la DNasa:
50 mM EDTA,
1,5 M d'acetat de sodi,
1% SDS.

3.2. PURIFICACIÓ DELS ÀCIDS NUCLEICS

3.2.1. ULTRACENTRIFUGACIÓ EN GRADIENT DE CsCl

Aquesta tècnica fou descrita primerament per Radloff *et al.*, (1967) i és especialment important per obtenir DNA de plasmidi/cosmidi d'una extraordinària puresa. Mitjançant aquest mètode varem purificar el DNA de cadascun dels cosmidis emprats en la tècnica de la selecció de cDNA.

3.2.1.1. GRADIENT DE CSCL/BROMUR D'ETIDI

Procediment

1. Créixer el cultiu bacterià fins a saturació a 37° C. Aturar el creixement a 4° C i centrifugar a 4.000 x g. Dissoldre el sediment bacterià i seguir els passos 1-10 del procediment de la maxipreparació de l'apartat 3.1.2.3.2.

2. Dissoldre el precipitat de DNA cru en 8 ml de tampó TE. Ajustar la concentració de DNA a 50-100 μ g/ml. Afegir 0,5 ml de 5 mg/ml de bromur d'etidi i 0,95 g de CsCl per cada ml de la solució del DNA en tampó TE.

3. Transferir la solució a tubs d'ultracentrifuga de 10 ml. Equilibrar el pes dels tubs amb solució de CsCl/TE. Un cop segellat el tub es centrifuga a 60.000 x g durant 24 hores a 15°C. És important no posar fre a la ultracentrifuga per a no trencar el gradient.

4. Amb l'ajut de la llum ultraviolada i mitjançant una agulla 20-G es perfora el tub i es recupera la banda que correspon al DNA covalentment circular del cosmidi (la banda inferior). D'aquesta manera el DNA és lliure de contaminants com ara DNA cromosòmic (banda superior), RNA (sediment) i proteïnes (banda flotant) (Fig. 10b de l'apartat 3.5.3.1.1.).

3.2.1.2. EXTRACCIÓ DEL BROMUR D'ETIDI

Procediment

1. Afegir un volum d'isopropanol saturat amb sals a la solució de DNA.

2. Mesclar bé i recuperar la fase inferior que és la fase que conté el DNA.

3. Repetir l'extracció (≈ 4 vegades) fins que desaparegui el color vermell del bromur d'etidi de la fase de l'isopropanol (fase superior) El bromur d'etidi restarà en la fase superior d'isopropanol.

4. Afegir 2 volums d'H₂O.

5. Afegir 6 volums d'etanol al 96% i precipitar el DNA a -20°C.

6. Extreure les sals amb etanol al 70%. Assecar el precipitat de DNA i dissoldre'l en TE.

Material i solucions

- Isopropanol saturat amb sals

Saturar un volum de 2-Butanol amb un volum de:

5 M NaCl,

10 mM Tris pH 8,5,

1 mM EDTA.

Es formen dues fases, la fase superior és amb la que hem de treballar (isopropanol saturat).

- Solució de CsCl/TE:

Afegir 0,95 g de CsCl per cada ml de TE pH 7,5.

3.2.2. PURIFICACIÓ D'ÀCIDS NUCLEICS MITJANÇANT FENOL

Procediment

1. Afegir un volum igual de fenol a la solució de DNA a purificar. Agitar durant 10 min per a una completa desproteïnització. L'extracció amb fenol cal repetir-la fins que no apareix cap precipitat en la interfície aquosa-orgànica.

2. Centrifugar 5 min a temperatura ambient per extreure el sobrenedant aquós que conté el DNA.

3. Afegir un volum igual de fenol/cloroform/isoamílic (25:24:1) al sobrenedant. Agitar durant 10 min.

4. Centrifugar 5 min a temperatura ambient per extreure el sobrenedant aquós.

5. Afegir un volum igual de cloroform/isoamílic (24:1) al sobrenedant. Agitar durant 10 min.

6. Precipitar el DNA del sobrenedant amb etanol al 96%. La precipitació és instantània i sovint, si la concentració de DNA és superior a 50-100 μ g/ml s'obté una medusa de DNA.

3.3. MANIPULACIÓ DELS ÀCIDS NUCLEICS

3.3.1. TRANSCRIPCIÓ INVERSA SEGUIDA DE PCR (RT-PCR)

Aquesta tècnica ens ha permès obtenir cDNA a partir d'mRNA mitjançant l'acció de l'enzim transcriptasa inversa. Aquesta reacció s'ha dut a terme amb el kit de transcripció inversa de RNA "GeneAmp®RNA", seguint les recomanacions de la casa comercial. Com a encebadors de la reacció s'han utilitzat tan hexanucleòtids degenerats (dN)₆ com oligonucleòtids de 27 timines (dT)₂₇ en experiments paral·lels. Darrerament s'ha emprat el kit "Ready-To-GoTM" obtenint millors resultats.

Procediment.

1. Utilitzar 1 μ g de RNA total, o bé 100 ng de RNA poli(A)⁺ en la reacció de la transcripció inversa. Seguir les recomanacions de la casa comercial.

2. Inactivar la transcriptasa inversa durant 5 min a 99°C.

3. Utilitzar 2 µl de cDNA d'aquesta transcripció inversa com a motlle en la reacció d'amplificació de la RT-PCR:

1,8 µl de tampó de PCR 10x,

1,6 µl de la mescla dels quatre nucleòtids a 1,25 mM cadascun,

15 pmols dels encebadors directe i revers,

0,5 U de Taq polimerasa,

 H_2O fins a 20 μ l.

Materials i solucions

- Tampó de PCR 10x:

```
500 mM KCl,
```

100 mM Tris pH 8,3,

15 mM MgCl₂,

0,1% gelatina.

- Kit de transcripció inversa de RNA "GeneAmp®RNA" (# N808-0017 de Perkin-Elmer).

- Kit de transcripció inversa de RNA "Ready-To-Go[™]" (# 27-9264-50 de Pharmacia Biotech).

3.3.2. HIBRIDACIÓ IN SITU FLUORESCENT DE METAFASES

Per tal de construir una genoteca de cosmidis de l'HC21 s'han aïllat aquells YAC que estaven localitzats en les regions d'interès de l'HC21. Aquestes foren tres:

1) La regió crítica de la síndrome de Down, al voltant del marcador D21S55.

- 2) Una regió proximal.
- 3) Una regió distal.

Per dur a terme els estudis de FISH (hibridació *in situ* amb fluorescència) de cadascun dels YAC d'aquestes regions s'ha fet una PCR inter-*Alu* per obtenir una sonda èspecífica de cada YAC. Mitjançant aquesta tècnica vàrem descartar tots aquells YAC que presentaven quimerisme o eren deleteris. Els experiments d'hibridació *in situ* foren duts a terme per la Marga Nadal de la Unitat de Citogenètica del nostre Departament.

3.3.3. INTRODUCCIÓ DE DNA PLASMÍDIC A CÈL·LULES

Per a la introducció de DNA de plasmidi a cèl·lules procariotes s'han emprat les dues tècniques que a continuació es detallen. En la primera és necessari fer un xoc tèrmic, mentre que en el segon mètode no ho és. Tanmateix, hem comprovat que l'aplicació d'un xoc tèrmic en aquesta darrera millora l'eficiència de la transformació.

3.3.3.1. MÈTODE DEL CaCL₂

Procediment

- 1. Inocular 25 ml de medi LB amb 100 µl d'un cultiu saturat d'E. coli i incubar-ho
- a 37°C fins que les cèl·lules arribin a tenir una $DO_{600} = 0,3-0,4$.
- 2. Aturar el creixement bacterià a 4ºC.
- 3. Centrifugar a 3.500 x g, 10 min.
- 4. Dissoldre el sediment de les cèl·lules en 25 ml de 50 mM CaCl2 fred.
- 5. Incubar 20-30 min a 4°C.
- 6. Centrifugar les cèl·lules a 3.500 x g, 5 min a 4°C.
- 7. Dissoldre el sediment de cèl·lules competents en 1 ml de 50 mM CaCl₂ fred.

8. Afegir 100 µl de cèl·lules competents a un tub prèviament refredat a 4°C amb 5 µl (0,1-100 ng DNA) del producte de la reacció de lligament. Incubar de 10 a 30 min a 4°C.

9. Transferir el tub amb les cèl·lules a un bany de 42°C, exactament durant 90 seg.

10. Incubar les cèl·lules transformades durant 10 min a 4°C.

11. Afegir 900 µl de medi TBS contenint 20 mM de glucosa i incubar 45-60 min a 37°C. Agitar molt suau perquè les cèl·lules puguin expressar resistència a l'antibiòtic.

12. Dispensar 450 μ l en una placa d'LB-agar amb 50 μ g/ml d'ampicil·lina. Si el plasmidi permet la selecció blava/blanca, afegir 35 μ l d'X-gal i 20 μ l d'IPTG per litre de medi de LB-agar. Incubar 12-16 hores a 37°C.

13. Comprovar l'eficiència de la transformació:

La transformació de 0,2 ng del plasmidi ha de rendir una eficiència teòrica de >4 $\times 10^7$ ufc μg^{-1} de plasmidi. Això equival a 400 colònies per placa de Petri petita.

66 Material i mètodes

L'eficiència de la nostra transformació serà:

Eficiència (ufc/ μ g) = núm. colònies x ng⁻¹ de plasmidi x 1000 x 20.

La freqüència de la transformació obtinguda mitjançant aquest mètode sovint ha estat entre 10^6 i 10^7 ufc/µg de DNA

Controls:

Control de la transformació: transformar 100 pg de plasmidi.

Control de l'antibiòtic: dispensar 100 µl del cultiu saturat d'*E. coli* en una placa d'LB amb antibiòtic.

Materials i solucions

- X-gal 50 mg/ml:

50 mg d'X-gal (5-bromo-4-cloro-3-indolil-B-galactosidasa) per ml de dimetilformamida.

- IPTG (isopropil B-D-tiogalactopiranòsid) 100 mM:

23,8 mg d'IPTG per ml d'H2O estèril.

3.3.3.2. MÈTODE DEL PEG 3500

Aquest procediment fou dissenyat primerament per Chung et al., (1989). Hem observat que l'eficiència d'aquest mètode és ostensiblement superior a l'eficiència del mètode mencionat anteriorment

Procediment

1. Inocular 25 ml de medi LB amb 100 μ l d'un cultiu saturat d'*E. coli*. Incubar a 37°C fins que les cèl·lules arribin a tenir una DO₆₀₀ = 0,3-0,4.

2. Aturar el creixement bacterià a 4ºC.

3. Centrifugar a 3.500 x g, 5 min a 4°C.

4. Dissoldre el sediment en 1 ml de TBS fred, dissoldre les cèl·lules molt suaument.

5. Incubar les cèl·lules transformades durant 10 min en gel.

6. Afegir 100 µl de cèl·lules en TBS a 5 µl (0,1-100 ng DNA) del producte de la reacció de lligament a un tub prèviament refredat a 4°C. Incubar de 10 a 30 min a 4°C.

7. Si bé no és necessari aplicar un xoc tèrmic, aquest augmenta l'eficiència. Incubar les cèl·lules a 42°C durant 60 seg.

8. Continuar amb el pas núm. 8 del protocol anterior.

Material i solucions

- TBS:

10 mM MgCl₂,

10 mM MgSO₄,

10% (p/v) polietilenglicol PEG 3.500 (1 g),

5% de DMSO (dimetilsulfòxid) (500 µl),

Completar fins a un volum de 10 ml amb medi LB.

3.3.4. REACCIÓ DE DESFOSFORILACIÓ DEL DNA

Aquesta tècnica s'ha emprat principalment per desfosforilar els extrems del vector i evitar d'aquesta manera el reaparellament d'aquest amb si mateix en la reacció de lligament amb l'insert. S'ha emprat l'enzim HK fosfatasa en detriment de l'ús d'altres fosfatases com la fosfatasa alcalina, perquè la HK fosfatasa és un enzim inactivable per calor.

Procediment

1. Preparar la reacció de desfosforilació com segueix:

2 µl de tampó de la fosfatasa 10x,

1 U de HK (Heat-Killed) fosfatasa/µg de DNA,

12 µl de la solució amb el DNA a desfosforilar,

Afegir ara CaCl₂ fins a assolir una concentració final de 5 mM,

 H_2O fins a un volum final de 20 μ l.

2. Incubar la reacció durant 1-2 hores a 30°C.

3. Inactivar l'enzim durant 30 min a 65°C.

Material i solucions

Tampó de la fosfatasa 10x:
333 mM Tris acetat pH 7,8,
666 mM acetat de potassi,
100 mM acetat de magnesi,
5 mM DTT,
0,1% Tritó X-100.

3.3.5. TRANSFERÈNCIA DE DNA A UNA MEMBRANA DE NILÓ

Representa una tècnica per immobilitzar el DNA a una membrana de niló. El mètode de la transferència ha estat diferent segons el tipus d'experiment:

1) El DNA purificat mitjançant una maxipreparació de DNA pot ésser aplicat directament a la membrana amb l'ajut d'un aparell de buit. Aquest fou el cas de la fixació del DNA de cada YAC a la membrana o *dot blot*, per a determinar l'origen de cada EST aïllat per la tècnica de la selecció de cDNA a un determinat YAC de l'HC21.

2) Hom pot fer créixer la colònia damunt la membrana, després es llisa la cèl·lula i posteriorment es fixa el DNA. Aquest fou el cas de la fixació del DNA de cadascun dels cosmidis de les genoteques dels YAC a la membrana, per a determinar l'origen de cada EST aïllat a un determinat cosmidi de l'HC21 i per a la construcció del mapa contigu de cosmidis.

3) Es fa una transferència de DNA del gel d'agarosa d'una electroforesi a la membrana per capil·laritat o Southern blot.

3.3.5.1. DOT BLOT

Procediment

1. Per a cada mostra, afegir el volum necessari de NaOH 1 M i EDTA 200 mM, pH 8,2 fins a tenir una concentració final de NaOH 0,4M/EDTA 10 mM. Bullir a 100°C durant 10 min.

2. Rentar la membrana de niló amb 500 μ l H₂O 10 min.

3. Dispensar la mostra de DNA purificat a la membrana i aplicar el buit a l'aparell. Deixar que la mostra sigui succionada i atrapada a la membrana.

4. Rentar cada pou amb 500 µl de NaOH 0,4M després d'afegir la mostra de DNA.

5. Rentar la membrana amb 2x SSC i assecar-la.

6. Incubar la membrana durant 1-2 hr a 80°C per fixar el DNA covalentment a la membrana.

3.3.5.2. MEMBRANA DE COSMIDIS D'ALTA DENSITAT

S'ha utilitzat un robot (biomek 1.000, Beckman) equipat amb una eina de transferència de 96 colònies. El procés s'ha dut a terme segons les recomanacions d'Olsen et al., (1993).

Procediment

1. Emplaçar amb el robot un inòcul d'un cultiu singular de cadascun dels cosmidis a una membrana de niló que es troba damunt una placa d'LB amb agar i amb l'antibiòtic escaient.

2. Fer créixer les 1.536 colònies (4x4x96) tota la nit a 37°C damunt la membrana.

3. Tractar la membrana segons els passos 8-10 descrits en l'apartat 3.7.1.2.1.

3.3.5.3. TRANSFERÈNCIA DE SOUTHERN

S'ha optat per la transferència de Southern de tipus alcalina.

Procediment

1. Incubar el gel d'agarosa en 0,25 M HCl durant 10-15 min o fins que el blau de bromofenol del tampó de càrrega canviï de color. Es recomana aquesta despurinització del DNA per a la transferència de fragments de DNA més grans de 5 kb. 2. Desnaturalitzar el DNA del gel d'agarosa en un bany amb 0,4 N NaOH. Agitar durant 30 min a temperatura ambient o fins que el blau de bromofenol torni a recuperar el seu color original.

3. Mullar la membrana de niló en H2O i incubar-la en 0,4N NaOH, 30 min.

4. Muntar l'aparell de capil·laritat, emprant 0,4 M NaOH.

5. Transferir el DNA del gel a la membrana durant 12-24 hores.

6. Rentar la membrana de niló amb 2x SSC després de la transferència.

7. Si la membrana de niló està carregada positivament no és necessari fer una posterior fixació covalent del DNA. En la resta de casos, el DNA es fixa definitivament a la membrana mitjançant l'exposició a un forn de 80°C durant 2 hores, o bé 3-5 min al trans-il·luminador amb una longitud d'ona de 312 nm.

3.3.6. HIBRIDACIÓ DE MEMBRANES

3.3.6.1. HIBRIDACIÓ DE MEMBRANES

Sigui quin sigui el tipus de membrana: membranes de genoteques de cDNA, *dot blots*, membrana de cosmidis, membrana de transferència de *Southern* i transferència de *Northern*, les hem hibridat amb una sonda de DNA de concentració 2-10 ng/ml i amb un nombre de cpm no inferior a $1-2 \times 10^6$ cpm/ml per aquelles sondes que corresponen a fragments de DNA, mentre que per a les sondes d'oligonucleòtids la concentració de DNA emprada ha estat de 20-50 ng/ml i un contatge de $1-2 \times 10^7$ cpm/ml. En tots dos casos l'activitat específica de les sondes no ha estat mai inferior a 5×10^8 cpm/µg. Per obtenir una sensibilitat màxima en membranes de transferència de *Northern* s'ha emprat una activitat específica de 10^9 cpm/µg.

Procediment

1. Rentar la membrana amb el DNA immobilitzat amb 2x SSC.

2. Col·locar la membrana a un tub d'hibridació, de manera que la cara de la membrana que conté el DNA quedi exposada cap a dintre.

3. Prehibridar la membrana amb la solució d'hibridació durant 1-2 hores a 65°C. Afegir DNA d'esperma de salmó a una concentració final de 20 μg per ml de tampó d'hibridació, per disminuir el soroll de fons o *background*.

4. Substituir el tampó d'hibridació per un de fresc que contingui la mateixa concentració d'esperma de salmó que el tampó de prehibridació.

5. Desnaturalitzar la sonda durant 5 min a 100°C. Incubar la membrana durant tota la nit a la temperatura escaient. Aquesta temperatura depèn de la longitud de cada fragment de DNA.

70 Material i mètodes

6. Després de la hibridació, rentar la membrana amb 0,2x SSC/0,1% SDS durant 15 min a temperatura ambient.

7. Aplicar un segon rentat, més astringent si s'escau, amb 0,1x SSC/0,1% SDS durant 15 min a 42°C.

8. Si cal augmentar l'astringència, incrementar la temperatura i/o disminuir la concentració d'SSC fins a 0,01x/0,1% SDS.

L'activitat específica cal que sigui superior a >10⁸ cpm/µg. Aquesta activitat és suficient per a detectar 0,5 pg de DNA diana.

Material i solucions

- Esperma de salmó desnaturalitzat:

Dissoldre 10 mg d'esperma de salmó (Sigma tipus 3) en 1 ml d'H₂O. Clivellar el DNA fent-ho passar per una agulla de 17-G varies vegades. Bullir la preparació a 100°C 15 min. Congelar.

- Solució de tampó fosfat 1M, pH 7,1:

Afegir una solució de 1 M de NaH₂PO₄·H₂O (138 gr/l) a una solució de 1 M Na₂HPO₄·2H₂O (178 gr/l) fins a obtenir un pH de 7,1.

- Solució d'hibridació:

Barrejar un volum de la solució de tampó fosfat 1M, pH 7,1 amb un volum de SDS 14%.

Afegir 100 μ g/ml de DNA d'esperma de salmó, desnaturalitzat just abans d'utilitzar-la.

3.3.6.2. EXTRACCIÓ DE LA SONDA DE LA MEMBRANA

Per a tornar a hibridar una membrana que ha estat hibridada anteriorment és necessari extreure la sonda radioactiva de la membrana.

Procediment

1. Afegir la membrana a 500 ml de solució de deshibridació bullint al 0,1%. Agitar constantment.

2. Deixar refredar fins a temperatura ambient.

Materials i solucions

- Solució de deshibridació:

1% (p/v) de SDS,

0,1 x SSC,

40 mM Tris HCl, pH 7,5 - 7,8.

3.3.7. MARCATGE DEL DNA

Material i mètodes 71

3.3.7.1. MARCATGE DEL DNA MITJANÇAT OLIGONUCLEÒTIDS A L'ATZAR Aquest mètode produeix DNA marcat radioactivament de manera uniforme amb una alta activitat específica (Feinberg i Vogelstein, 1983). Fàcilment s'ha obtingut unes activitats específiques de 10^8 - 10^9 cpm/µg sonda. S'ha emprat aquesta tècnica per marcar fragments de DNA >180 pb.

Procediment

1. Purificar el DNA (per electroforesi si s'escau).

2. Preparar la següent reacció de marcatge a 4°C :

10 µl de tampó OLB,

4 μ l de BSA (albúmina de sèrum boví), és un enzim estabilitzador no específic que s'uneix a certs inhibidors de polimerases.

3 μl [alfa-32P] dCTP 3.000 μCi,

1 μ l de l'enzim Klenow (3U).

3. Portar el DNA (10-30 ng) a un volum de 32 μ l amb H₂O. Bullir la solució de DNA durant 2-3 min a 100°C. Refredar fins a 4°C.

4. Combinar els 18 μ l de la reacció de marcatge amb els 32 μ l del DNA desnaturalitzat (el volum final és de 50 μ l). Incubar immediatament a 37°C durant 2-3 hores, o tota la nit a temperatura ambient.

5. Aturar la reacció amb 1 µl de 0,5 M EDTA i 100 µl de tampó TE.

6. Separar el DNA marcat dels dNTP no incorporats mitjançant una cromatografia en columna de Sephadex G-50.

7. Extreure 1 µl abans i després del marcatge per determinar la quantitat de ³²P dCTP incorporat.

Material i solucions

- Tampó OLB:

Obtenir una mescla de la solució A/solució B/solució C en proporció 2:5:3

- Solució O:

1,25 M Tris HCl, pH 8,0 (1,51 g),

0,123 M MgCl₂·6 H₂O (0,25 g),

H₂O fins a 10 ml (9,4 ml). Esterilitzar a l'autoclau.

- Solució A:

470 µl de solució O,

9 μl ß-mercaptoetanol,

12,5 μ l 20 mM de dATP,

12,5 µl 20 mM de dTTP,

12,5 µl 20 mM de dGTP.

Solució B:
6,51 g HEPES,
H₂O fins a 12,5 ml pH 6,6,
Filtrar.
Solució C:
Hexadesoxiribonucleòtids [d(N)₆] en TE. DO 90 unitats/ml.
Guardar a -20°C.
BSA: 5 mg/ml.

3.3.7.2. MARCATGE DEL DNA PER PCR

S'ha emprat aquest mètode alternatiu quan la mida del fragment de DNA era inferior a 150 pb, sempre i quan hem disposat d'encebadors específics per a la seva amplificació per la PCR.

Procediment

1. Preparar una PCR amb els següents reactius:

100 ng de DNA genòmic o 0,12 pg d'un fragment clonat en un vector,

10 pmols de cada primer,

3 µl de [alfa-P³²] dCTP 800 Ci/mmol,

2 µl de dCTP 6 µM,

200 µM de dGTP,

200 µM de dTTP,

200 μ M de dATP,

2,5 µl de 10x tampó de la Taq polimerasa,

2 U Taq polimerasa,

 H_2O fins a 25 μ l de volum final.

2. Vint-i-vuit cicles d'amplificació són suficients.

3. Purificar la sonda per cromatografia amb Sephadex G-50. Determinar l'activitat específica (apartat 3.3.9.).

3.3.7.3. SONDES D'OLIGONUCLEÒTIDS

Procediment

1. Preparar una reacció de volum final 20 µl:

2 μl de tampó de la quinasa 10x,

8 μl (80 pmols) d'oligonucleòtid,

2 μ l de [gamma-³²P] dATP amb una activitat específica >3.000 Ci/mmol, 2 μ l BSA,

1 µl (10 U) de l'enzim T4 polinucleòtid quinasa,

5 μl H2O.

2. Incubar la reacció durant 60 min a 37°C. Aturar la reacció afegint 1 μ l de 0,5 M EDTA, o escalfant 10 min a 75°C.

3. Purificar la sonda per cromatografia amb una columna de Sephadex G-25. Determinar l'activitat específica.

Material i solucions

```
- Tampó de la quinasa 10x:
```

```
500 mM Tris·HCl pH 7,5,
```

```
100 mM MgCl<sub>2</sub>,
```

50 mM ditiotreitol (DTT).

- BSA:

5 mg BSA/ml d'H₂O.

3.3.7.4. MARCATGE TERMINAL DEL DNA

Aquest mètode ens ha permès marcar els extrems 5' de fragments molt petits de DNA < 80pb, encara que l'eficiència ha estat decreixen si l'extrem 5' era protuberant, rom o interior, repectivament. En aquest darrer cas hem desnaturalitzat abans el DNA. El protocol és idèntic al descrit en l'apartat 3.3.7.3., però s'ha emprat 8 nmols de DNA.

3.3.7.5. OLIGONUCLEÒTIDS COM A ENCEBADORS PER AL GENOTIPATGE Procediment

1. Preparar la reacció de marcatge com segueix:

3 μl de tampó de la quinasa 10x,

3 µl de l'oligonucleòtid específic del microsatèl·lit (50 ng),

1 μl de [gamma-³²P ό ³³P] dATP (3.000 Ci/mmol),

0,5 µl (4,5 U) de l'enzim polinucleòtid quinasa del bacteriòfag T4,

 H_2O fins a 30 µl de volum final.

- 2. Incubar durant 45-60 min a 37°C.
- 3. Aturar la reacció durant 10 min a 95°C.
- 4. Purificar els oligonucleòtids marcats mitjançant una columna de Sephadex G-25 amb 10 ml de TE 1/0,1 pH 7,5 (apartat 3.3.8.2.).

5. Utilitzar 1,5 pmols d'aquest encebador marcat en cada reacció de PCR, en contrast amb els 10 pmols de l'encebador no marcat.

Material i solucions

Tampó de la quinasa 10x:
500 mM Tris·HCl, pH 7,6,
100 mM MgCl₂,
50 mM DTT,
1 mM espermidina,
1 mM EDTA, pH 8,0.

3.3.8. PURIFICACIÓ D'UNA SONDA DE DNA

S'ha realitzat una extracció de les proteïnes, sals i nucleòtids de la reacció de marcatge del fragment de DNA, puix que hem observat que la purificació del DNA marcat per columna de cromatografia disminueix el soroll de fons de l'hibridació. El seu procediment varia segons el tipus de sonda.

3.3.8.1. PURIFICACIÓ DE SONDES >50 pb

Procediment

1. Taponar una xeringa de 1 ml amb cotó o llana de vidre. Omplir la columna amb la solució en suspensió de la reïna Sephadex G-50 en TE.

2. Centrifugar la reïna a 3.000 x g, 5 min per a compactar-la.

3. Aplicar la mostra de DNA (ajustada a un volum final de 200 μ l amb TE) al centre de la columna.

4. Col·locar la columna dintre d'un tub *eppendorf* de 1,5 ml. Recuperar el DNA purificat centrifugant a 3.000 x g, 5 min.

5. Contar les cpm per determinar la quantitat de radioactivitat incorporada.

6. Determinar l'activitat específica (segons l'apartat 3.3.9.).

Material i solucions

- Sephadex G-50:

Afegir 500 g de Sephadex G-50 mediu (# 17-0045-02 de Pharmacia Biotech) a 1 litre de TE 1/0,1, pH 7,5. Rentar el Sephadex tres vegades amb TE. Esterilitzar a l'autoclau.

3.3.8.2. PURIFICACIÓ DE SONDES <50 pb

Procediment

1. És el mateix que l'indicat anteriorment, però s'utilitza la reïna de Sephadex G-25 en TE.

2. Determinar l'activitat específica.

Material i solucions

- Sephadex G-25:

Afegir 500 g de Sephadex G-25 superfine (# 17-0572-02 de Pharmacia Biotech) a 1 litre de TE 1/0,1, pH 7,5. Rentar el Sephadex tres vegades amb TE. Esterilitzar a l'autoclau.

3.3.9. DETERMINACIÓ DE L'ACTIVITAT ESPECÍFICA I DE LA QUANTITAT DE CPM DE LA SONDA

Com s'ha comentat en l'apartat 3.3.6.1., l'activitat específica (AE) i el nombre de comptes per minut (cpm) de la sonda són dos paràmetres molt importants alhora d'hibridar membranes

Procediment

1. Extreure alíquotes abans i després de la purificació del marcatge del DNA per cromatografia de Sephadex.

2. Emprar la següent fórmula de càlcul:

AE (cpm/µg) = cpm precromatografia x factor de dilució x volum final de la reacció μg^{-1} de DNA emprats en la reacció de marcatge.

3. El total de les cpm de la fracció de la sonda bé atesa per la fórmula:

cpm de la sonda = cpm postcromatografia $_x \mu l$ del volum de la fracció de la sonda $_x$ factor de dilució.

3.3.10. FOSFORILACIÓ DE L'EXTREM 5' D'ÀCIDS NUCLEICS

Mitjançant aquest procés hem fosforilat l'encebador OLIGO3 de l'adaptador de DNA per a la seva posterior reacció de lligament a la genoteca de cDNA (apartat 3.5.3.1.), o bé l'encebador Anchor1 per a la seva posterior utilització en la tècnica de RACE (apartat 3.7.2.1.1.). La reacció és molt similar a la descrita per al marcatge de sondes d'oligonucleòtids detallat en l'apartat 3.3.7.3., però s'hi afegeix ATP (1 mM final) en comptes de [gamma-³²P] dATP.

Procediment

1. Afegir l'oligonucleòtid (2 µg) o el fragment de DNA a fosforilar al tub de reacció.

2. Afegir 10 µl del tampó de la T4 polinucleòtid quinasa (PNK) 10x.

3. Posar 10 µl d'ATP de 10 mM. (1 mM final).

4. Afegir 6 µl de l'enzim T4 polinucleòtid quinasa (10.000 U/ml).

5. Afegir ddH₂O fins a 100 μ l de volum final.

6. Incubar la reacció durant 60 min a 37°C.

7. Inactivar la PNK durant 20 min a 65°C.

3.3.11. REACCIÓ DE LLIGAMENT DEL DNA

Els enzims que catalitzen aquesta reacció són la T4 DNA lligasa i la lligasa d'*E. coli*. La primera utilitza ATP com a font d'energia, mentre la segona utilitza NAD. La diferència més important és que solament la T4 lligasa és capaç de lligar extrems roms en condicions estàndards. La relació molar (nombre de molècules de DNA) entre el vector i l'insert ha estat des de 1:1 fins a 1:5. Aquesta relació depèn de si és tracta d'un clonatge d'una banda única o de bandes múltiples, respectivament. La quantitat de DNA de l'insert que hem posat a la reacció per tenir una relació 1:1 ha estat segons la fórmula següent:

ng de DNA d'insert = mida de l'insert (pb) x ng de DNA del vector x pb⁻¹ mida del vector

Per lligar extrems roms de DNA ens ha calgut afegir més enzim lligasa (0,25 U) i PEG al 15% (p/v).

Procediment

1. Preparar una reacció de volum final 10 µl com segueix:

1 µl de tampó 10x de la T4 lligasa,

25 ng de vector,

ng de DNA a clonar, segons la fórmula i relació molar mencionats anteriorment.

1 μ l ATP de 10 mM,

 H_2O fins a 10 μ l,

0,005 U de l'enzim T4 DNA lligasa.

2. Incubar la reacció 12-16 hores a 16°C, o bé a temperatura ambient durant 1-3 hores.

3. Aturar la reacció amb 2 μ l de 0,5 M EDTA o escalfant 10 min a 75°C.

4. Per transformar les cèl·lules competents amb el DNA de la reacció de lligament hem utilitzat la meitat del volum de la reacció (5 μ l).

Material i solucions

- Tampó de la T4 lligasa 10 x:

660 mM Tris HCl, pH 7,5,

50 mM MgCl₂,

10 mM DTT,

10 mM ATP.

- Controls:

Control de la lligament: 25 ng de vector digerit amb extrems cohesius no desfosforilats.

Control de la desfosforilació: 25 ng de vector digerit amb extrems cohesius desfosforilats.

Control de la digestió del vector: 25 ng de vector digerit, però sense l'enzim T4 DNA lligasa.

3.3.12. ANÀLISI DE RNA PER TRANSFERÈNCIA DE NORTHERN

L'anàlisi de l'RNA per transferència de Northern o Northern blot ens ha estat molt útil per estudiar el patró d'expressió d'un gen en cada teixit, o bé la seva possible sobreexpressió en un teixit determinat.

3.3.12.1. ELECTROFORESI EN GEL D'AGAROSA/FORMALDEHID

Procediment

1. Dissoldre 1,0 g d'agarosa en 72 ml d'H₂O i bullir-ho. Deixar refredar fins a 65° C.

2. Afegir 10 ml de tampó MOPS 10x i 18 ml de 12,3 M de formaldehid (la concentració final de l'agent desnaturalitzant és de 2,2 M). Afegir també 1 μ l de bromur d'etidi de concentració 10 mg/ml. D'aquesta manera obtindrem un gel al 1% d'agarosa. Aquesta concentració es escaient per separar molècules de RNA de 500 pb a 10 kb de mida.

3. Abans de córrer el gel, preequilibrar-lo amb tampó de la carrera MOPS 1x.

4. Afegir 1 volum de tampó de càrrega 5x per cada 4 volums de RNA. En cada carril del gel poden ser carregades des de 0,5 fins a 10 μ g de RNA. Mesclar i incubar 15 min a 55°C. Refredar a 4°C.

5. Dur a terme l'electroforesi a 5 V/cm en solució de tampó MOPS 1x en continua agitació.

Material i solucions

- Tampó MOPS 10x:

2 M MOPS (àcid 3-(N-morpholí)-propanosulfònic), pH 7,0,

5 M acetat de sodi,

0,1 M EDTA.

- Tampó de càrrega 5x:

2,7% (v/v) concentració final de formaldehid:

1 mM EDTA, pH 8,0,

0,25% (p/v) de blau bromofenol,

0,25% (p/v) xilencianol,

50 % (v/v) glicerol.

2.4.12.2. TRANSFERÈNCIA DE L'RNA DEL GEL A LA MEMBRANA

Una transferència de Northern de 10 μ g de RNA total ens ha estat suficient per detectar un mRNA que representava el 0,01-0,005% de la població total de RNA (com és el cas de l'expressió del gen *MNBH*). Per a trànscrits de baixa expressió hem emprat 3 μ g de la fracció de poli(A)⁺. Aquesta quantitat de RNA ha estat suficient per detectar aquells trànscrits que representaven menys del 0,0002% de la població d'mRNA. La sondes emprades tenien una activitat específica superior a 5 x10⁸ cpm/ μ g.

Procediment

1. Rentar diverses vegades el gel d'agarosa amb H_2O desionitzada lliure de RNasa per extreure el formaldehid, ja que aquest disminueix l'eficiència de la transferència.

2. Afegir uns 10 volums de 0,05 M NaOH/1,5 M NaCl i incubar-ho durant 30 min. Decantar i substituir-ho per 10 volums de 0,5 M TrisHCl, pH 7,4/1,5 M NaCl, durant 20 min per neutralitzar-lo.

3. Reemplaçar la solució amb 10 volums de 20x SSC i incubar-ho durant 45 min. Rentar la membrana de niló amb H₂O i després submergir-la en la mateixa solució de 20x SSC.

4. Dur a terme la transferència amb solució 20x SSC durant tota la nit.

5. Després de la transferència, rentar la membrana amb 2x SSC i assecar-la 2 hores a 80°C.

Val a dir que, de manera ocasional, alguns Northern blots foren comercials, bàsicament per la indisponibilitat de disposar de RNA d'un determinat teixit, o d'un determinat moment

Materials i solucions

- Membranes comercials de Northern blot:

Human Multiple Tissue Northern Blot (# PT1200-1 de Clontech).

Human Brain Multiple Tissue Northern Blot III (# 7750-1 de Clontech).

Human Brain Multiple Tissue Northern Blot II (# 7755-1 de Clontech).

Controls:

- Control de la transferència: tenyir el gel d'agarosa amb bromur d'etidi després de la transferència de l'RNA

3.3.13. CONSTRUCCIÓ DEL T-VECTOR

L'ús de la Taq polimerasa en la reacció de la PCR condueix a l'addició d'un nucleòtid (quasi en exclusiva d'una adenosina, Clark, 1988) a l'extrem 3' del producte de PCR.

Hem aprofitat aquesta particularitat per clonar aquest producte de l'amplificació en un vector modificat (T-vector), que conté un sol dTTP com a extrem protuberant (Marchuk -et al., 1991). Atès que l'eficiència i el tipus de nt que la Taq polimerasa afegeix al producte de la PCR està en funció de la seqüència de l'encebador (Costa et al., 1994), aquests han estat dissenyats de tal manera que tenen una T en la posició més 5' en llur seqüència.

També hem emprat el kit "TA cloning *Kit*" i el kit "pMOS*Blue* T-vector *kit*", però els millors resultats els hem obtingut amb el protocol que a continuació es detalla

Procediment

1. Digerir 3 μ g de DNA del vector plasmidi (pBluescript) amb *Eco*RV durant tota la nit a 37°C per obtenir una completa digestió. D'aquesta manera es generen dos extrems roms.

2. Fer una extracció amb fenol i una altra amb fenol/cloroform. Precipitar amb etanol al 96%.

3. Rentar el precipitat de DNA amb etanol al 70%.

4. Dissoldre el plasmidi linealitzat en la següent reacció d'incubació:

158 µl H2O,

20 µl tampó de PCR 10x que conté 15 mM MgCl₂,

20 µl 1,25 mM dTTP,

2 μl 5 U/μl Taq DNA polimerasa.

5. Incubar 2-3 hores a 75°C.

6. Afegir 200 µl de tampó TE.

7. Repetir els passos 2-3.

8. Dissoldre el precipitat de DNA en 100 μ l de tampó TE. Distribuir en alíquotes de 50 ng. Congelar a -70°C.

Material i solucions

- Kit "TA cloning Kit", (# K2000-01 d'Invitrogen).

- Kit "pMOSBlue T-vector kit", (# RPN 1719 d'Amersham Life Science).

3.3.14. LA DNA GLICOSILASA D'URACIL

A diferència del mètode anterior, el clonatge de productes de PCR mitjançant la DNA glicosilasa d'uracil (UDG) requereix la síntesi d'oligonucleòtids amb una cua 5'-CUACUACUACUA afegida a la seqüència específica de l'encebador. Quatre dels dotze dNTP són dUTP que serveixen com a dianes per a l'enzim UDG. D'aquesta manera, la digestió de l'insert i del vector amb l'enzim UDG forma uns extrems protuberants de 12 bases. L'avantatge d'aquest mètode és l'alta eficiència de clonatge (1 x10⁷ clons per μ g DNA), ja que la reacció d'aparellament entre les cadenes protuberants de l'insert i del vector té lloc espontàniament i ràpidament, sense el concurs de cap enzim.

Procediment

1. Afegir 2 µl (100 ng del producte de PCR),

2 µl 25 ng/µl de DNA del plasmidi pAMP10,

1 μ l tampó de la *Taq* DNA polimerasa 10x,

1 μ l 1 U/ μ l enzim UDG,

 H_2O fins a 10 µl de volum final. Mesclar i incubar 30 min a 37°C.

2. Transformar 100 µl de cèl·lules competents d'E. coli.

3.3.15. BIOINFORMÀTICA

La manipulació informàtica del DNA ha esdevingut una eina cabdal en la recerca tan important com ho pugui ésser la reacció de la PCR. A continuació detallaré per la seva importància en la realització d'aquesta tesi, aquells recursos informàtics que han estat més emprats.

3.3.15.1. EDITORS DE SEQÜÈNCIA

Comprenen aquells programes que permeten editar una seqüència, manipular-la, dur a terme mapes de restricció, trobar pautes de lectura oberta o ORF, traduir una seqüència de DNA, emmagatzemar-la i exportar-la en el format més escaient per a cada aplicació. Cal destacar els programes DNA Strider, MacVector i SeqEd.

3.3.15.2. PREDICCIONS DE L'ESTRUCTURA DEL DNA I DELS DOMINIS DE LES PROTEÏNES

El programa GCG STATPLOT ens mostra el contingut de GC i AT d'una sequència. El pograma GCG MFOLD realitza una predicció de l'estructura secundària de l'RNA. És útil per reconèixer regions de control de l'mRNA. PSORT és una altre programa per analitzar diversos aspectes i paràmetres de les proteïnes, o bé els possibles dominis funcionals d'aquestes.

3.3.15.3. DISSENY D'OLIGONUCLEÒTIDS

El programa PRIMER selecciona encebadors compatibles (directe i revers) per a la PCR. Personalment recomano el disseny manual d'encebadors, atenen els criteris mostrats en l'apartat 3.3.16.

3.3.15.4. IDENTIFICACIÓ DE REGIONS CODIFICANTS DE PROTEÏNES

El programa GRAIL no solament és un programa de predicció de regions codificants, sinó que també fa prediccions de seqüències promotores, identificació d'illes CpG, caracterització virtual dels límits entre els introns i els exons, identificació de seqüències repetitives i la recerca d'altres motius.

3.3.15.5. RECERCA DE SIMILITUDS I IDENTITATS

En destacaré els dos programes principals utilitzats en aquesta tesi: BLAST i FASTA.

3.3.15.5.1. BLAST: És una eina informàtica basada en complicats algoritmes matemàtics i mètodes estadístics, que ens donen una idea de la similitud entre la nostra seqüència i qualsevol altra que pugui haver en el banc de dades de seqüències. Aquesta recerca es pot dut a terme amb la seqüència de nucleòtids (BLASTN) o amb la d'aminoàcids (BLASTX).

3.3.15.5.2. FASTA: És una molt similar a l'anterior. La principal diferència és que FASTA és capaç de generar forats o *gaps* en l'alineament per trobar una millor optimització entre les seqüències, ja que és capaç de trobar similituds significatives entre dues seqüències que comparteixen un alt grau de similitud, encara que una d'elles difereixi de l'altra per una inserció o una deleció d'un fragment de DNA. En aquest cas, BLAST podria no trobar similitud significativa.

3.4.15.6. RECERCA D'INFORMACIÓ

3.3.15.6.1. Entrez

Representa una base de dades integrada molt útil per buscar informació sobre publicacions, referències literàries o cercar informació sobre un tema científic mitjançant paraules clau.

3.3.15.6.1. MEDLINE

Representa un programa informàtic de característiques i prestacions similars a l'anterior.

Tots aquests programes i una àmplia col·lecció d'altres programes informàtics amb interès biològic es poden trobar a la pàgina principal o *Home Page* que el Centre de Genètica Mèdica i Molecular de l'Institut de Recerca Oncològica posa a l'abast de tothom a l'adreça de web: http://www.iro.es/gm.html, dissenyada, caracteritzada i actualitzada per en Carles Pucharcós i en Marià Tarragó del nostre departament.

3.3.16. DISSENY MANUAL D'ENCEBADORS

Atès que no sempre els encebadors dissenyats per programes informàtics són els més escaients, sovint s'ha obtat per un disseny manual dels encebadors atenen els següents criteris:

- Les sequències riques en M(G ó C) en el 5' de l'oligonucleòtid confereixen una gran estabilitat d'aparellament amb la sequència diana.

- Els encebadors amb 3' inestables ($\Delta G > -9$ kCal/mol) sovint augmenten l'especificitat del producte amplificat per la PCR, ja que la possibilitat d'aparellar-se amb cadenes no específiques és més petita.

 Incorporar dues M (C ó G) en les darreres 3 bases de l'encebador, i millor si una de les M es troba en la posició més 3'.

- Incorporar una G o una C com a darrer nucleòtid en el 3' de l'encebador.

- Incorporar una T en el 5' de cada encebador per facilitar el clonatge posterior del producte de la PCR en el T-vector.

- Fer un anàlisi de BLASTN de la sequencia de DNA a partir de la qual es dissenya l'encebador, per evitar els encebadors amb motius repetitius del genoma.

- Calcular una temperatura mitjana d'aparellament (T_m) entre 55 i 58°C. Per estimar aproximadament la T_m hem emprat la fórmula següent:

 $T_{\rm m}$ (°C) = 2(N_a + N_t) + 4(N_g + N_c)

on NA és el nombre d'adenines,

on NT és el nombre de timines,

on NG és el nombre de guanines,

on Nc és el nombre de citosines.

- Què el nombre total de nucleòtids de l'encebador sigui entre 20 i 23.

3.4. MAPA DE COSMIDIS

3.4.1. CONSTRUCCIÓ DE GENOTEQUES DE COSMIDIS

Atès que: 1) el nostre objectiu principal fou la construcció d'un mapa transcripcional d'alta resolució, i 2) l'eficiencia de la recerca d'EST en la tècnica de la selecció de cDNA és més elevada quan s'utilitzen cosmidis en comptes de YAC com a material genòmic, vàrem optar per la construcció d'una genoteca de cosmidis a partir dels YAC seleccionats.

3.4.1.1. ANÀLISI DELS YAC

La nostra primera estratègia fou definir un grup de YAC amb un solapament mínim entre ells per tal de cobrir les regions cromosòmiques de l'HC21 d'interès. Aquestes foren tres: una regió al voltant del marcador de referència D21S55 dintre la DSCR (YAC 336G11 i 238B1), una regió proximal (YAC 72H9), i una altra distal (YAC 221B9 i 552A3). Els 5 YAC de l'HC21 foren escollits entre els reportats per Chumakov *et al.*, (1992a) i atenent als criteris de selecció que a continuació s'esmenten:

- 1) Longitud: entre 300 i 600 kb.
- 2) Estabilitat: sense reordenacions.
- 3) Quimerisme: no quimèrics.

3.4.1.1.1. Anàlisi de la mida i estabilitat

El DNA de cada llevat que conté un determinat YAC fou immobilitzat dins un cos esfèric i massís d'agarosa per preservar la seva integritat. Aquestes boletes massises d'agarosa que contenen el DNA íntegre s'han sotmès a una electroforesi de camps polsants o PFGE, juntament amb un marcador de pesos moleculars de cadascun dels cromosomes naturals del llevat (# 27-4520 Pharmacia). El fragment de DNA corresponent al YAC s'ha detectat mitjançant l'hibridació de la transferència de *Southern* amb una sonda de DNA específica dels extrems del vector del YAC. Aquells YAC que van donar bandes de mida diferent a la descrita foren descartats per la susceptibilitat a tenir reordenacions genètiques.

3.4.1.1.2. Anàlisi de quimerisme

L'estudi del possible quimerisme de cada YAC es dugué a terme per dos mètodes diferents: 1) anàlisi de FISH amb sondes inter-*Alu*, i 2) amplificació mitjançant la PCR de la bombolla.

3.4.1.1.2.1. Generació de productes Alu PCR a partir dels YAC per a estudis de FISH

S'ha obtingut un producte de PCR inter-Alu a partir del DNA de cada llevat que contenia el YAC d'interès, emprant els encebadors específics del 5' i del 3' de les seqüències repetitives Alu mencionats en l'apartat 1.4.4. D'aquesta manera, solament les seqüències humanes del YAC són amplificades per aquest mètode a partir del DNA del llevat. La mida dels productes resultants de la PCR són entre 500 i 4.000 pb de longitud. Aquesta col·lecció heterogènia de fragments amplificats per PCR representen parts de l'insert de DNA humà del YAC i l'hem emprat com a sonda específica en els anàlisis de FISH sobre metafases de cromosomes humans.

Procediment

1. Amplificar per la PCR un volum final de 100 μ l, a partir de 10 ng de DNA de llevat que conté el YAC d'interès i amb 30 pmols de cadascun dels encebadors específics de les sequències Alu (Alu-3' i Alu-5') (Fuentes et al., 1997). Atès que les sequències Alu són un consensus, hem afegit a la reacció de la PCR una quantitat

84 Material i mètodes

extra de MgCl₂ fins a obtenir una concentració final de 5 mM. Les condicions de la PCR són: 3 min a 95°C, 35 cicles de 1 min a 94°C, 1 min a 56°C i 2 min a 74°C, amb una extensió final de 5 min a 74°C.

2. Purificar el producte de la PCR amb fenol/cloroform/isoamílic. Precipitar el DNA amb etanol al 96%. Rentar amb etanol al 70%.

3. Dissoldre el DNA precipitat en 50 μl d'H₂O destil·lada.

3.4.1.1.2.2. PCR de la bombolla

Mitjançant la tècnica de la "PCR de la bombolla" o *bubble PCR* (Ausubel *et al.*, 1994) vàrem aïllar els fragments de DNA de cadascun dels extrems del YAC. Hem utilitzat aquests extrems del YAC com a sonda en una transferència de *Southern* que contenia DNA digerit amb *Eco*RI de la línia cel·lular híbrida WAV-17, la qual és específica de l'HC21.

Procediment

1. Digerir 3 μ g del DNA de llevat que conté el YAC amb l'enzim *Rsa*I. Aquesta digestió genera diversos fragments, entre els quals hi ha dos que, a més de contenir DNA humà de l'insert, contenen la seqüència dels braços dret i esquerra del YAC, respectivament.

2. Inactivar els enzims durant 15 min a 65°C.

3. Preparar la bombolla d'oligonucleòtids (bubble-top i bubble-bottom). Els dos oligonucleòtids són complementaris en la seva seqüència excepte en la part central (nucleòtids en cursiva). Quan aquests oligonucleòtids s'aparellen entre sí generen una "bombolla" de no complementarietat nucleotídica. Ajustar els dos tipus d'oligonucleòtids a una concentració de 4 nmol/ml amb H₂O. Mesclar-los i incubar-los 15 min a 68°C. Refredar-los lentament fins a temperatura ambient (60 min). D'aquesta manera s'obté un adaptador de doble cadena amb una "bombolla" al mig.

4. Preparar la següent reacció de lligament:

250 ng de DNA digerit,

2 pmols dels oligonucleòtids aparellats en forma de bombolla,

5 µl del tampó de la lligasa 10x,

2 U de la T4 DNA lligasa,

 H_2O fins a un volum final de 50 μ l.

5. Incubar 2 hores a 37°C.

6. Afegir 200 µl de H₂O per tenir una concentració final de DNA de 1 ng/µl.

7. Amplificar aquells fragments de DNA que contenen la sequència dels braços del vector pYAC4 emprant els encebadors específics de cadascun dels braços i

l'encebador específic de la sequència no complementària de la bombolla, en experiments paral·lels. Les condicions de la PCR són 35 cicles de 1 min a 92°C, 2 min a 65°C, i 2 min a 72°C. Per amplificar el fragment corresponent al braç dret del YAC emprar l'encebador HYAC-C. Per amplificar el braç esquerra emprar l'encebador HYAC-D. L'encebador específic de la bombolla és el 224.

Material i solucions

- Encebadors:

Bubble-top: 5'-gaaggaggaggacgctgtctgtcgaaggtaaggaacggagaggaggaggaggag-3'.
Bubble-bottom: 5'-ctctcccttctcgaatcgtaaccgttcgtacgagaatcgctgtcctctccttc-3'.
HYAC-C: 5'-gctacttggagccactatcgactacgcgat-3'.
HYAC-D: 5'-ggtgatgtcggcgatataggcgccagcaac-3'.
224: 5'-cgaatcgtaaccgttcgtacgagaatcgct-3'.

3.4.1.2. OBTENCIÓ DE DNA ENCAPSULAT EN BOLETES D'AGAROSA

Em emprat aquest mètode per a l'extracció intacta dels cromosomes a partir de cèl·lules de llevat que contenen el YAC d'interès. El mètode és una modificació del mètode original d'Overhauser i Radic, 1987. L'aïllament del DNA per mètodes tradicionals sotmeten el DNA a forces de cisalla, trencant-lo. D'aquesta manera, ens calgut emprar mètodes d'extracció de DNA que minimitzin la seva degradació. Aquest mètode d'encapsulament de les cèl·lules en boletes d'agarosa o *beads* ens ha permés purificar el DNA de les cèl·lules de llevat, preservant la seva integritat física i protegint-lo de la degradació.

Procediment

1. Fer un cultiu en estria en una placa de Petri amb medi AHC, a partir del glicerol del llevat amb el YAC d'interés.

2. Inocular una colònia singular de llevat en 150 ml de medi AHC. Agitar a 250 rpm durant 4-5 dies a 30°C fins que el cultiu arribi a una $DO_{600} = 1$.

3. Centrifugar les cèl·lules a 500 x g, 3 min a 4°C.

4. Descartar el sobrenedant. Dissoldre el sediment de cèl·lules en 10 ml de SE i incubar-les durant 10 min a temperatura ambient.

5. Centrifugar les cèl·lules a 500 x g, 2 min a 4°C.

6- Repetir els dos passos anteriors dos cops més. La segona, però, dissoldre les cèl·lules en tan sols 1 ml d'SE.

7. Transferir les cèl·lules en suspensió a un falcon de 50 ml i escalfar-ho a 42°C.

8. Afegir 4 ml d'agarosa de baix punt de fusió o low meelting point en SE a les cèl·lules i mesclar.

86 Material i mètodes

9. Afegir 15 ml d'oli mineral i 70 ml d'SE fred. Mesclar-ho tot vigorosament en un bany amb gel durant minuts fins que es formi una emulsió uniforme.

10. Afegir 100 μ g/ml de proteïnasa K i incubar les esferes 16 hores a 50°C.

11. Centrifugar a 500 x g durant 5 min. Després de centrifugar-les, veurem dues fases. Recuperar la fase aquosa (fase inferior) que conté les esferes d'agarosa.

12. Dissoldre les esferes d'agarosa amb SE fins a obtenir un volum final de 50 ml. Centrifugar a 500 x g durant 5 min. Rentar les esferes amb SE dos cops més.

13. Afegir 0,5 ml de 2-mercaptoetanol i 5 mg de zimolasa. Portar el volum final a 10 ml amb SE.

14. Dissoldre les esferes i incubar 2 hores a 37°C.

15. Rentar 2 cops amb SE durant 10 min. Fer el darrer rentat amb sarcosil/EDTA, 10 min més.

16. Centrifugar a 700 x g, 5 min. Descartar el sobrenedant i dissoldre'l en 20 ml de TE.

Material i solucions

- Solució SE:

75 mM NaCl,

25 mM EDTA.

Ajustar el pH final a 8,0 amb NaOH.

- Sarcosil/EDTA:

sarcosil 1%,

25 mM EDTA.

Ajustar el pH final a 8,0 amb NaOH.

- Medi de cultiu AHC (ura⁻, trp⁻):

Representa un medi selectiu on solament els llevats que contenen un YAC poden sobreviure, ja que la soca mare AB1380 és (trp1⁻, ura3⁻)

0,85 g de base nitrogenada de llevat que no contingui aminoàcids.

2,5 g de sulfat d'amoni

5 g d'hidrolitzat de caseïna sense sal ni vitamines,

5 ml d'hemisulfat d'adenina,

450 ml d'H₂0 bidestil·lada.

Ajustar el pH a 5,8. Esterilitzar a l'autoclau i afegir 50 ml de D-glucosa al 20% (p/v) estèril i filtrada. Per a les plaques sòlides de AHC, afegir 10 g d'agar abans d'autoclavar.

3.4.1.3. RECUPERACIÓ DEL DNA DE LES BOLETES D'AGAROSA

Material i mètodes 87

Atès que estàvem interessats en obtenir grans fragments intactes de DNA (>300 kb) s'ha optat per fer una digestió enzimàtica amb ß-agarasa. Aquest enzim digereix els polisacàrids de cadena llarga presents a l'agarosa a monosacàrids; aquest són fàcilment extrets amb una diàlisi o bé amb una purificació amb fenol.

Procediment

1. Rentar dues vegades les boletes d'agarosa que contenen el DNA (400 μ l) amb 2 vol del tampó de la β -agarasa 1x durant 30 min a 4°C. Afegir 4 μ l del tampó de la β -agarasa a les boletes (0,5x concentració final).

2. Desfer el gel completament, escalfant-lo 5 min a 70°C.

3. Equilibrar l'agarosa desfeta a 45° C (10 min). Afegir 1 U de l'enzim ß-agarasa per cada 200 µl d'agarosa al 1%, i continuar la incubació durant 2 hores més. Afegir 2 µl més de ß-agarasa. Incubar 2 hores.

4. Dialitzar la solució de DNA per extreure els carbohidrats i la ß-agarasa. Ajustar la concentració de sals de la solució de DNA a 0,3 M d'acetat de sodi i afegir un volum d'isopropanol. Incubar 15 min a 4°C.

5. Precipitar el DNA dialitzat amb etanol al 96%. Centrifugar a 15.000 x g, 15 min.

6. Rentar el sediment de DNA amb etanol al 70%. Dissoldre el sediment final amb tampó TE, pH 8,0.

Material i solucions

- L'enzim ß-agarasa 1 (# 3925 de New england Biolabs)

3.4.1.4. FRACCIONAMENT PARCIAL DEL DNA DE YAC

El mètode emprat fou una digestió parcial amb quantitats limitades d'un enzim de restricció de tall freqüent (*Mbol*) durant un temps limitat, per obtenir una col·lecció de fragments tallats a l'atzar. *Mbol* és un enzim de tall freqüent, de manera que és possible generar fragments que se solapin entre ells i que cobreixin tot el material genètic del YAC.

Procediment

La mida del DNA és crític en aquest experiment, de manera que els DNAs dels llevats foren prèviament encapsulats en boletes d'agarosa, a fi i efecte de preservar la integritat dels YAC.

1. Rentar les boletes d'agarosa dues vegades amb tampó TE 0,5x.

2. Afegir 50 ml de tampó TE 0,5x a les esferes d'agarosa.

3. Dialitzar >3 hores a 4° C.

4. Canviar el tampó una altra vegada, i dialitzar 3 hores a 4ºC.

5. Transferir les esferes, a 50 ml del tampó de l'enzim de restricció *Mbo*I sense MgCl₂. Dialitzar 2 hores a 4°C.

6. Extreure la solució tampó i pesar les esferes. Assumir que 1 g = 1 ml agarosa i una concentració de 300 μ l DNA/ml agarosa.

 Afegir tampó fresc d'Mbol sense MgCl₂ fins a un volum final de 1 ml. Incubar 15 min a 37°C.

8. Afegir 10 μ l de MgCl₂ 1 M per obtenir una concentració final de 10 mM. Incubar 15 min a 37°C.

9. Introduir 0,01 U d'enzim *Mbo*I. Extreure una alíquota de 30 μ I de la digestió cada 10 min (incloent el temps zero). Aturar la reacció de la digestió amb solució *stop*. Analitzar la mida en un gel d'agarosa al 0,5%. El temps de la digestió desitjat és aquell a on la mitjana de la mida dels fragments de DNA digerits és entre 30 i 50 kb.

10. Un cop el temps i la concentració de l'enzim han estat determinats, els volums de la reacció són incrementats proporcionalment per contenir 1 mg de DNA.
Preescalfar el volum a 37°C.

11. Després que la digestió parcial arribi al temps òptim, aturar la reacció ràpidament amb 100 μ l d'EDTA 0,5 M durant 30 min a 4°C. Efectuar una extracció del DNA amb fenol/cloroform. Precipitar el DNA amb etanol.

12. Dissoldre el precipitat de DNA molt suaument amb 50 μ l de TE.

Material i solucions

Solució stop:
10 mM Tris·HCl, pH 7,5,
20% de glicerol,
0,1% de SDS,
0,1% de blau de bromofenol.

3.4.1.5. DESFOSFORILACIÓ DE LA DIGESTIÓ PARCIAL DEL DNA DEL LLEVAT

El DNA del llevat amb el YAC es desfosforila per evitar la formació de concatenàmers de petits fragments de DNA en la reacció de lligament amb el vector i evitar la formació de clons quimèrics.

Procediment

Preparar la següent reacció:
 μl de tampó CIAP 10x,
 μl de CIAP (enzim fosfatasa),
 μl d'H₂O,

50 µl del DNA digerit de la preparació anterior (apartat 3.4.1.4.).

2. Incubar 1 hora a 37°C.

3. Afegir 3 μ l de 0,5 M EDTA. Incubar 10 min a 68°C.

4. Purificar la digestió parcial desfosforilada amb fenol/cloroform i posterior precipitació i rentat en etanol.

Material i solucions

- Solució CIAP 10x:

500 mM Tris HCl, pH 9,5.

100 mM EDTA.

3.4.1.6. LLIGAMENT DEL DNA FRACCIONAT DEL YAC ALS COSMIDIS

Els fragments de DNA desfosforilats de la digestió de cadascun dels llevats que contenen els YAC d'interès foren lligats a un vector cosmidi anomenat Supercos1 (Fig. 9). Aquest cosmidi havia estat prèviament linealitzat amb un enzim de restricció per generar extrems cohesius amb els extrems dels fragments de la digestió parcial del DNA del llevat generats amb *Mbo*I.

3.4.1.6.1. Digestió del cosmidi amb XbaI

La digestió dels cosmidis amb Xbal linealitza el vector Supercos1.

Procediment

1. Digerir uns 50 μ g del vector cosmidi Supercos1 durant 1 hora a 37°C amb 9 U/ μ g de l'enzim XbaI, en un volum final de 200 μ l en condicions estàndards.

2. Comprovar la seva completa digestió en un gel d'agarosa.

3. Purificar el DNA digerit amb una extracció de fenol/cloroform isoamílic. Rentar amb cloroform. Precipitar amb etanol al 96%. Rentar les sals amb etanol al 70%.

4. Dissoldre el precipitat de DNA digerit en H₂O desionitzada fins a obtenir una concentració de 1 μ g/ μ l.

3.4.1.6.2. Desfosforilació del cosmidi digerit amb XbaI

5. Desfosforilar el DNA del vector digerit amb XbaI amb la fosfatasa alcalina (CIAP), emprant 2,5 U/50 pmols de DNA en tampó de CIAP 1x, 1 hora a 37°C.

6. Inactivar la fosfatasa alcalina, ajustant la reacció a 15 mM EDTA i incubant-la durant 10 min a 68°C.

7. Purificar el DNA amb fenol, i posteriorment amb cloroform.

8. Precipitar el DNA amb etanol al 96%. Rentar amb etanol al 70%.

9. Dissoldre el precipitat de DNA desfosforilat amb tampó TE fins a tenir una concentració de 1 μ g/ μ l.

Material i solucions

- Tampó CIAP 10x: 500 mM Tris·HCl, pH 9,5. 100 mM EDTA.

3.4.1.6.3. Segona digestió amb BamHI

10. Digerir el DNA del cosmidi Supercos 1 desfosforilat (20 μ g) del pas anterior amb l'enzim *Bam*HI en un volum final de 200 μ l, 1 hora a 37°C. Comprovar la total digestió en un gel d'agarosa.

11. Purificar el producte de la segona digestió amb fenol-cloroform, seguit d'un cloroform. Precipitar i rentar amb etanol.

12. Dissoldre el DNA en H₂O fins a tenir una concentració final de 1 μ g/ μ l.



Figura 9. Esquema del cosmidi Supercos1.

3.4.1.6.4. Lligament del DNA fraccionat al cosmidi Supercos1 doblement digerit.

Procediment

1. Preparar una reacció de lligament tal com segueix:

2,5 µg del DNA del llevat digerit i desfosforilat de l'apartat 3.4.1.5.,

1 µg de DNA de SuperCos1 de l'apartat 3.4.1.6.3.,

2 μl del tampó de la lligasa 10x,

2 μl d'ATP 10 mM,

1 µl de l'enzim T4 lligasa (1-2 U),

 H_2O fins a un volum final de 20 μ l.

Incubar la reacció tota la nit a 4ºC.

Després d'aquesta reacció, els dos extrems de cada fragment de DNA genòmic del llevat amb el YAC estan lligats a una seqüència de DNA del vector anomenada COS. Aquesta seqüència COS que es troba a banda i banda del fragment del YAC serà posteriorment reconeguda per l'enzim que empaqueta el DNA dintre la càpsida del fag.

3.4.1.7. EMPAQUETAMENT DEL CDNA FRACCIONAT DEL YAC

Per tal d'introduir els cosmidis que contenen el DNA de la digestió parcial del YAC a cèl·lules d'*E. coli* s'ha optat per l'empaquetament de DNA en partícules del fag lambda, bàsicament per dues raons: 1) per la seva gran eficiència de transformació i 2) perquè les mides del DNA encapsulat són entre 37 i 45 kb.

3.4.1.7.1. Preparació de les cèl·lules hostes

Procediment

1. Fer un cultiu en estria d'*E. coli* XL1-Blue MR sobre placa sòlida d'LB per obtenir una colònia senzilla de cèl·lules hostes.

2. A partir d'aquesta colònia senzilla, fer crèixer un cultiu de cèl·lules hostes en medi LB suplementat amb MgSO₄ i maltosa. Agitar durant 4-6 hores a 37° C, sense arribar a una DO₂₆₀ = 1,0.

3. Centrifugar 3.000 x g, 10 min.

4. Dissoldre el sediment de cèl·lules hostes en 10 mM MgSO₄ estèril. El volum és la meitat del volum inicial.

5. Diluir les cèl·lules hostes amb 10 mM MgSO₄ estèril fins a tenir una $DO_{260} = 0,5$.

3.4.1.7.2. Empaquetament dels cosmidis

S'ha utilitzat el kit d'empaquetament "Gigapack®II Packaging Extract" (# 200217 de Stratagene). S'ha procedit segons les indicacions de la casa comercial, emprant 5 µl de la reacció de lligament de l'apartat 3.4.1.6.4. D'aquesta manera s'han obtingut partícules del fag lambda amb la capacitat d'infectar cèl·lules hostes. Aquests fags contenen dos braços amb les seqüències COS del vector cosmidi, flanquejant el fragment de DNA del YAC digerit. Aquest DNA empaquetat en el fag lambda esdevindrà un cosmidi quan penetri dins la cèl·lula hoste d'*E.coli*.

3.4.1.7.3. Infecció de les cèl·lules hostes

Primerament és necessari titular la genoteca de fags. Un cop s'ha calculat el títol, s'ha diluit la solució de fags fins a tenir unes 433 ufc/placa de Petri gran (15 cm). S'han obtingut 5 plaques per cada YAC, resultant un total de 2.654 colònies de cosmidis. Amb

aquest nombre de colònies, la possibilitat de tenir tot el material genòmic del llevat representat en cosmidis és bastant alta, segons els càlculs següents:

13 x10³ kb del llevat x 35⁻¹ kb de mida mitjana del cosmidi x 5 cops representat = 1.857 Tot això s'ha fet separadament per cada YAC.

Procediment

1. Afegir 500 μ l d'SM als 20 μ l de la reacció d'empaquetament de l'apartat 3.4.1.7.2.

2. Afegir 1 µl d'aquesta dil·lució a 25 µl d'SM.

3. Afegir-hi 25 µl de cèl·lules hoste *E. coli* XL1-Blue MR preparades segons l'apartat 3.4.1.7.1.

4. Incubar les cèl·lules hostes infectades amb els fags durant 30 min a 37°C. Quan el DNA del fag penetra en la cèl·lula hoste esdevé un DNA circular o cosmidi. Només aquelles cèl·lules amb el cosmidi podran sobreviure en un medi d'LB amb kanamicina.

5. Afegir-hi 200 µl d'LB. Créixer les cèl·lules durant 45-60 min a 37°C.

6. Dispensar el creixement en plaques d'LB sòlid amb kanamicina.

7. Incubar les cèl·lules durant tota la nit a 37°C.

D'aquesta manera s'ha obtingut un títol de 1,5 x106 ufc/ml

Material i solucions

- Kanamicina:

Concentració final 50 µg/ml medi.

3.4.1.7.4. Primera selecció d'aquells clons amb DNA del YAC

S'han fet créixer sobre una membrana de niló cadascuna de les 2.654 colònies de cosmidi esmentades anteriorment. Aquestes membranes s'han hibridat amb una sonda radioactiva de DNA humà per detectar aquells cosmidis que tenien material del YAC, descartant així aquells que contenien DNA del llevat. La proporció de clons positius per la sonda fou de 1 cada 40-60 colònies. D'aquesta manera, a partir dels 5 YAC s'han obtingut 502 cosmidis amb material humà de l'HC21.

3.4.1.7.5. Segona selecció d'aquells clons amb DNA del YAC

Em realitzat una digestió amb un enzim de restricció de tall freqüent (*Eco*RI i *Hind*III) cadascun dels 502 cosmidis obtinguts en l'apartat 3.4.1.7.4. Aquesta digestió enzimàtica fou separada electroforèticament i transferida a una membrana de niló (apartat 3.4.2.2.). Aquesta membrana s'ha hibridat amb la mateixa sonda humana emprada en el pas anterior i amb una sonda específica del DNA de llevat, en experiments paral·lels.
Aquest segon clivellatge ens ha havia de permetre detectar els falsos positius de la primera hibridació, o bé detectar els clons quimèrics de llevat-humà.

3.4.2. CONSTRUCCIÓ D'UN MAPA CONTIGU DE COSMIDIS

Es tracta d'obtenir una col·lecció de cosmidis a partir d'un YAC, endreçats de manera seqüencial, i que entre tots, representin el total del material genètic del YAC. Abans de començar a analitzar la contigüitat dels cosmidis necessitem identificar els dos cosmidis que contenen els fragments de DNA humà de cadascun dels extrems del YAC. Això s'ha aconseguit hibridant la genoteca de cosmidis de cada YAC amb una sonda específica de cadascun dels braços del YAC, en experiments paral·lels. Aquests dos cosmidis representen els extrems dreta i esquerra de cada YAC, i s'han emprat com a punts de referència per a la construcció del mapa contigu de cosmidis. L'ordre de la resta de cosmidis s'ha establert atenen els següents criteris:

1) Establiment d'un ordre preliminar dels cosmidis mitjançant la comparació de llurs patrons de digestió amb enzims de restricció.

- 2) Anàlisi del contingut d'STS.
- 3) Anàlisi de les ribosondes de cada cosmidi (sondes de RNA).
- 4) Anàlisi per PCR linial de cada cosmidi.

D'aquests experiments es desprenen dos tipus d'informació: l'ordre d'un cosmidi respecte als altres i la seva orientació respecte al cromosoma.

3.4.2.1. LOCALITZACIÓ PEL CONTINGUT D'STS

Aquells cosmidis que comparteixin un mateix STS de còpia única hauran de solapar part del seu contingut de seqüència. La presència o no d'un determinat marcador de DNA s'ha dut a terme mitjançant la tècnica de la PCR amb encebadors específics de cada STS.

3.4.2.2. MAPA DE RESTRICCIÓ

La comparació del patrons de digestió amb un enzim de restricció de cadascuna de les minipreparacions de DNA de cada cosmidi ha servit també per definir un ordre preliminar dels cosmidis, identificant aquells que solapaven un fragment de DNA. Les endonucleases de restricció emprades han estat *Eco*RI i *Hind*III.

Procediment

1. Digerir el DNA de la maxipreparació de cada cosmidi amb els enzims de restricció *Eco*RI i/o *Hind*III, tal com segueix:

10 µg de DNA del cosmidi,

5 µl del tampó de l'enzim de restricció (1x final),

94 Material i mètodes ·

 $1 \ \mu l \ 50 \ mM$ espermidina (1 mM final).

 H_2O fins a un volum final de 50 μ l.

2. Afegir 30 U de l'endonucleasa i incubar la reacció durant 4 hores a 37ºC.

3. Afegir tampó de càrrega 10x fins a obtenir una concentració final de 1x per aturar la reacció.

4. Carregar la digestió en un gel d'agarosa, el suficientment gran perquè les bandes de la digestió se separin bé (30 cm).

3.4.2.3. RIBOSONDES

Per tal de construir un mapa contigu de cosmidis, vàrem utilitzar el DNA genòmic dels cosmidis com a motlle per a la seva transcripció *in vitro*, mitjançant l'acció de les RNA polimerases T7 i T3. Aquestes polimerases actuen a partir de llurs promotors presents en el vector Supercos1, flanquejant cadascun dels extrems de l'insert. Em utilitzat aquestes ribosondes de RNA, específiques de cada extrem de l'insert, con a sondes en hibridacions de membranes que contenen la col·lecció de cosmidis específics de cada YAC.

Procediment

1. Linealitzar 200-400 ng de DNA del cosmidi amb l'endonucleasa *Rsa*I o *Hae*II. S'utilitzen aquests enzims de tall freqüent per obtenir sondes curtes i el suficientment específiques perquè el soroll de fons sigui tan petit com sigui possible.

- 2. Purificar el DNA de la digestió amb fenol/cloroform.
- 3. Precipitar en etanol. Rentar les sals.
- 4. Dissoldre el DNA precipitat en 10 μ l d'H₂O tractada amb DEPC.
- 5. Sintetitzar ribosondes segons la següent reacció:
- 4 µl de tampó de transcripció 5x,
- 1 µl DTT de 20 mM,
- 20 U d'inhibidor de l'RNasa,
- 4 µl d'una mescla d'ATP, GTP, CTP a 2,5 mM cadascun,
- 1,2 µl d'UTP 200 µM,
- 5 μ l de DNA linealitzat (200 ng -1 μ g),
- 1 µl d'alfa-³²P UTP (50 µCi/ml),
- 15 U de la polimerasa T3 o de la T7,
- H_2O fins a un volum final de 20 µl.

Afegir els components de la reacció a temperatura ambient per evitar que aquests precipitin amb l'espermidina:

6. Incubar durant 60-120 min a 37-40°C.

Les sondes obtingudes d'aquesta manera han tingut una gran activitat específica (5 $\times 10^8$ cpm/µg), i fàcilment s'ha obtingut uns 5-10 µg de RNA/µg de plasmidi).

7. Digerir el DNA de la reacció anterior amb l'enzim DNasa I.

El producte de RNA obtingut potser utilitzat com a sonda, sense la necessitat de practicar cap altre pas de purificació.

Material i solucions

- Tampó de transcripció 5x: 200 mM Tris·HCl, pH 7,5, 30 mM MgCl₂, 50 mM NaCl.

3.4.2.4. SONDES PER PCR LINEAL

Una altra tècnica complementària a la síntesi de ribosondes fou la síntesi de sondes de cDNA de cadena senzilla mitjançant la PCR lineal.

Procediment

1. Preparar una reacció de PCR lineal tal com segueix:

2 µl dels encebadors T7 o T3 a 10 µM,

2 μl [alfa-32P] dCTP, 800 Ci/mmol,

10 ng del cosmidi digerit del pas núm. 1 de l'apartat 3.4.2.3.

 $2 \mu l dCTP de 6 \mu M$,

1 µl d'una mescla de dTTP, dATP i dGTP de 2 mM,

2,5 µl del tampó de la Taq polimerasa,

0,4 µl de la Taq polimerasa,

 H_2O fins a un volum final de 25 µl.

2. Les condicions de la PCR són 3 min a 94°C, 35 cicles de 30 s a 94°C, 40 s a 46°C i 1 min a 74°C.

3. Utilitzar aquests productes de PCR com a sonda en l'hibridació de les membranes que contenen tots els cosmidis d'un determinat YAC.

3.5. TÈCNICA DE LA SELECCIÓ DE cDNA

3.5.1. PREPARACIÓ DE RNA POLI(A)+

Aproximadament l'1% de l'RNA total de la cèl·lula correspon a RNA poli(A)⁺. S'ha recuperat l'RNA poli(A)⁺ a partir de l'RNA total de cervell fetal humà mitjançant una cromatografia d'afinitat amb cel·lulosa-oligo(dT).

Procediment

1. Equilibrar la columna de la reïna poli-(dT) amb 15 ml de tampó de càrrega. Cal que el pH del líquid recuperat sigui a prop de 7,5 al final del rentat.

96 Material i mètodes

2. Escalfar uns 2 mg de RNA total en H₂O durant 10 min a 70°C (per desfer les estructures secundàries de l'RNA). Afegir LiCl 10 M fins a tenir una concentració final de 0,5 M.

3. Passar la solució de RNA total a través de la columna d'oligo(dT).

4. Rentar la columna amb 1 ml de tampó de càrrega poli(A)+.

5. Passar el tampó de càrrega recuperat a través de la columna dos cops més.

6. Rentar la columna amb 2 ml de tampó de rentat suau.

7. Recuperar l'RNA a un tub fresc amb 2 ml de 2 mM EDTA/0,1% SDS.

8. Reequilibrar una columna nova de poli(dT) com en el pas 1. Repetir la selecció de RNA poli(A)⁺ amb l'RNA recuperat de la columna anterior, segons els passos 3-6. Aquesta segona selecció elimina petits nivells de RNA poli(A)⁻ (RNA estructural, no missatger).

9. Ajustar la concentració de sals de l'RNA recuperat a 0,3 M d'acetat de sodi. Precipitar l'RNA amb 2,5 volums d'etanol al 96%.

10. Incubar l'RNA durant tota la nit a -20°C. Recollir el precipitat centrifugant a 10.000 x g, 30 min.

11. Assecar el precipitat de RNA i dissoldre'l en 150 µl de tampó TE lliure de RNasa.

L'RNA cal que aparegui com una cua llarga o *smear* amb una mida inferior a unes 20 kb en un gel d'agarosa i sense evidències de bandes de rRNA.

Material i solucions

Tampó de càrrega poli(A)+:
0,5 M LiCl,
10 mM Tris·HCl, pH 7,5,
1 mM EDTA,
0,1% de SDS.
Tampó de rentat suau:
0,15 M LiCl,
10 mM Tris·HCl, pH 7,5,
1 mM EDTA,
0,1% de SDS.
Columna oligo(dT):

Oligo(dT)-cel·lulosa (# 300138 d'Stratagene).

3.5.2. PREPARACIÓ DE CDNA A PARTIR DE RNA POLI(A)⁺

La conversió d'mRNA a cDNA és un element imprescindible per a la tècnica de la selecció de cDNA. La qualitat del cDNA, o sia, l'existència de petites quantitats de DNA genòmic i d'mRNAs no representats en el cDNA determina l'eficiència de la tècnica mencionada. La qualitat de la preparació de la genoteca de cDNA (Fig. 10a), o bé la puresa del DNA -genòmic (Fig. 10b) determinaran l'èxit o el fracàs de la tècnica de selecció de cDNA, tal i com es discuteix en l'apartat 5.2.

3.5.2.1. CONVERSIÓ D'MRNA A dscDNA D'EXTREMS ROMS

Per tal de poder lligar uns adaptadors d'amplificació al cDNA, cal que aquest darrer sigui de doble cadena.

3.5.2.1.1. Síntesi de cDNA

Procediment

1. Preparar 10 μ g de RNA poli(A)⁺ a la concentració de 0,5 μ g/ μ l.

2. Escalfar l'RNA (10 μ g en 20 μ l) durant 5 min a 65°C. Incubar-ho immediatament després a 4°C.

3. Preparar la mescla per a la reacció de la transcripció inversa tal com segueix:

20 µl 5 mM dNTPs (mescla dels 4),

40 µl tampó de la transcripció inversa 5x,

10 µl 200 mM DTT (agent reductor),

20 µl 0,5 mg/ml oligo-(dT)15,

2 µl [alfa-32P] dCTP, 3.000 Ci/mmol,

58 μl H₂O,

10 µl (10U) inhibidor de RNases.

4. Mesclar i afegir els 160 μ l de la mescla del pas anterior al tub que conté l'RNA desnaturalitzat del darrer pas núm. 2.

5. Afegir 20 μl (200 U) la transcriptasa inversa Superscript (el volum final és de 200 μl).

6. Mesclar i incubar durant 5 min a temperatura ambient.

7. Incubar la reacció durant 1,5 hores a 42ºC.

8. Preparar una altra reacció exactament igual i de manera paral·lela a l'anterior, però se substitueixen els oligo(dT)₁₅ per oligòmers a l'atzar (dN)₆. Incubar la reacció durant 1,5 hores a 37°C.

9. Aturar la reacció a 4ºC.

Material i solucions

- Tampó 5x de l'enzim SuperScript (transcriptasa inversa):

250 mM Tris HCl, pH 8,3,

375 mM KCl,

15 mM MgCl₂.

Cal que totes les solucions siguin lliures de RNases. L'H2O cal que sigui bidestil·lada i tractada amb DEPC.



Figura 10. Esquema general de l'obtenció de dscDNA amb adaptadors d'amplificació, i de DNA genòmic d'alta puresa.

3.5.2.1.2. Conversió del sscDNA a dscDNA

Procediment

9. Afegir els següents reactius amb l'ordre indicat al sscDNA (cDNA de cadena senzilla) del pas anterior (apartat 3.5.2.1.1.) en un volum fianl de 1.000 μ l d'H₂O: 400 μ l 5x del tampó de la segona cadena (1x final),

2 µl [alfa-32P] dCTP, 3.000 Ci/mmol,

10 µl (10 U) de RNasa H,

50 µl (500 U) de la DNA lligasa d'E. coli (no T4 DNA lligasa),

10 µl (100 U) polimerasa I d'E. coli,

 $338 \ \mu l \ H_2O.$

Mesclar i incubar la reacció durant 120 min a 12°C. Després 60 min a 22°C, i 10 min a 65°C.

Material i solucions

- Tampó de la segona cadena 5x:

100 µl 1 M Tris HCl, pH 7,5,

500 µl 1 M KCl,

25 μl 1 M MgCl₂,

50 μl 1 M (NH4)2SO4,

50 μl 1 M DTT,

50 μ l 5 mg/ml BSA,

225 μl H2O.

10. Afegir a la reacció anterior 40 µl de la T4 DNA polimerasa (20 U).

11. Mesclar i incubar la reacció durant 45 min a 37°C.

13. Afegir 120 μ l EDTA 0,2 M, pH 7,2. Purificar el dscDNA amb fenol-cloroform. Precipitar amb etanol al 96%. Dessalar amb etanol al 70%.

Material i solucions

- RNasa H (lliure de DNasa):

Dissoldre 1 mg/ml de RNsa H en tampó TE. Bullir 10 min per extreure les DNases contaminants. Deixar refredar lentament fins a temperatura ambient. Diluir a concentració de 2 μ g/ml en tampó TE.

- Control:

La incorporació de dCTP al cDNA cal que sigui > $2x10^5$ cpm, tant en la síntesi de la primera cadena, com en la síntesi de la segona. Tots dos productes de síntesi es verifiquen en un gel d'agarosa desnaturalitzant; la mida de tots dos cal que siguin equiparables. D'aquesta manera s'obté 3 µg de cDNA.

3.5.3. OBTENCIÓ DE dscDNA AMB ADAPTADORS

3.5.3.1. CREACIÓ DE L'ADAPTADOR DE DOBLE CADENA

Procediment

1. Aparellar quantitats equimolars (p.e. $2 \mu g$) de cadascun dels encebadors OLIGO3 i OLIGO4 (l'encebador OLIGO3 ha estat prèviament fosforilat amb l'enzim T4 polinucleòtid quinasa, segons l'apartat 3.3.10.).

2. Incubar 1 volum d'aquesta mescla amb un volum de solució d'hibridació 2x durant 5 min a 98°C i deixar-los refredar lentament en el bany fins a temperatura ambient.

3. Purificar l'adaptador de doble cadena amb fenol/cloroform i posterior precipitació amb etanol.

100 Material i mètodes-

La nostra experiència ens indica que el cDNA purificat amb fenol/cloroform i posterior precipitació i rentat amb etanol dessala millor que les columnes de Sephadex G-50.

Material i solucions

- Solució d'hibridació 2x:

1,5 M NaCl,

40 mM fosfat de sodi, pH 7,2,

10 mM EDTA, pH 8,0,

10x solució Denhardt,

1,2% (p/v) de SDS.

- Solució Denhardt 1x:

0,02% Ficoll,

0,02% polivinilpirrolidona,

0,02% BSA.

- Oligonucleòtids

OLIGO3: 5'-ctcgagaattctgatcctc-3'.

OLIGO4: 5'ttgaggatcagaattctcgag-3'

Els encebadors han estat purificats mitjançant HPLC. L'OLIGO4 és el complementari de l'oligonucleòtid OLIGO3 amb un extrem protuberant de 2 timines, per evitar la formació de concatenàmers en la reacció de lligament amb el dscDNA.

3.5.3.2. UNIÓ DE L'ADAPTADOR DE DOBLE CADENA O *LINKER* AL dscDNA

Hem afegit uns adaptadors d'amplificació o *linkers* al dscDNAs primaris, o sia, no subclonat, per poder-los amplificar per la PCR amb l'encebador específic de l'adaptador.

Procediment

1. Fer una reacció de lligament entre l'adaptador de doble cadena i els dscDNA en un volum total de 30 μ l, afegint-hi:

19 μ l (3 μ g) de cDNA purificat,

3 μl tampó de la T4 DNA lligasa 10x,

2 μ l d'adaptador de doble cadena (2 μ g),

3 μl (3 U) de l'enzim T4 DNA lligasa,

3 µl d'ATP 5 mM.

2. Mesclar i incubar tota la nit a 4°C.

L'adaptador de doble cadena conté 2 bases protuberants a l'extrem 3' que no permet la formació d'adaptadors concatenats.

3. Inactivar l'enzim durant 10 min a 65°C.

4. Purificar el dscDNA amb l'adaptador. L'extracció dels adaptadors no lligats, encebadors de cadena senzilla i les sals es realitza primerament mitjançant una columna de Sephadex G-50, i posteriorment amb una extracció amb fenol i cloroform.

5. Precipitar amb etanol al 96%. Rentar amb etanol al 70%. Dissoldre el precipitat de DNA en H_2O per tal d'obtenir una concentració de 400 ng/µl.

Material i solucions

```
- Tampó de la T4 DNA lligasa 10x:
200 mM Tris·HCl, pH 7,6,
50 mM MgCl<sub>2</sub>,
50 mM DTT,
500 μg/ml BSA.
```

3.5.4. MARCATGE DELS COSMIDIS AMB BIOTINA

El DNA dels cosmidis purificats per gradient de CsCl són immobilitzats en la tècnica de la selecció de cDNA mitjançant el seu marcatge amb biotina. Aproximadament 50 molècules de biotina són incorporades per kb de DNA.

Procediment

1. Desnaturalitzar i linealitzar 2 μ g de DNA dels cosmidis purificats per gradient de CsCl en un volum de 34 μ l d'H₂O, durant 5 min a 98°C. Refredar en gel.

2. Afegir al DNA desnaturalitzat:

10 µl de tampó OLB 5x (amb els oligonucleòtids generats a l'atzar (dN)6),

5 μl de la mescla dNTP/biotina,

1µl (5U) fragment de Klenow,

5 μl [alfa-32P] dCTP, 3.000 Ci/mmol,

 H_2O suficient per a 50 µl de volum final.

2. Incubar la reacció 30 mim a 37°C.

3. Aturar la reacció amb 3 µl de 0,5 M EDTA, pH 8,0, o escalfar durant 10 min a 65° C.

4. Precipitar el DNA, afegint 5 μ l de LiCl 4 M i 150 μ l d'etanol fred al 96%. Incubar 30 min en gel. Rentar amb etanol al 70%.

5. Dissoldre el DNA en 50 µl de tampó TE, pH 7,5.

6. Repetir els passos 4-5, dues vegades més.

7. Determinar l'activitat específica. Aquesta cal que sigui superior a 10^9 cpm/µg. Material i solucions

- dNTP/biotina:
0,08 mM de cada dCTP, dTTP, dGTP,
0,1 mM biotina-14-dATP,
500 mM TrisHCl, pH 7,8,
50 mM MgCl₂.

3.5.5. TÈCNICA DE LA SELECCIÓ DE cDNA

Representa una tècnica per a obtenir una població heterogènia de fragments de cDNA, enriquida en seqüències que es transcriuen a mRNA en la cèl·lula i que es localitzen a la regió d'interés. Per a minimitzar les contaminacions de DNA genòmic i de RNA no processat (hnRNA) s'ha aïllat l'RNA del citoplasma. Per a maximitzar la complexitat del cDNA vàrem partir d'una genoteca primària o *primary library* de cDNA, a partir de l'RNA citoplasmàtic poliadenilat al que li hem afegit un adaptador per a la PCR. Per visualitzar l'enriquiment de la tècnica en cada cicle de selecció és imprescindible l'ús d'un gen com a control positiu. En aquest cas, el gen emprat fou *DSCR1*, caracteritzat en el nostre Departament (Fuentes *et al.*, 1995) i localitzat a l'HC21, a una de les regions a la qual s'ha aplicat la tècnica de la selecció de cDNA. Ja que la tècnica de la selecció de cDNA està basada en la reacció de la PCR, com a norma general, l'enzim *Taq* polimerasa no s'ha afegit fins al primer cicle o *hot start*, per guanyar especificitat.

3.5.5.1. AMPLIFICACIÓ DE LA GENOTECA DE CDNA

S'amplifica la genoteca primària de dscDNA amb els adaptadors mitjançant la PCR (*hot start*) amb l'encebador OLIGO3 (apartat 3.5.3.1.) el qual és específic de l'adaptador dels cDNA. Les condicions d'amplificació són les estàndards, amb una temperatura d'aparellament de 60°C. El producte d'amplificació de la PCR és purificat amb columnes de sephadex G-50, precipitat i dessalat amb etanol.

3.5.5.2. SUPRESSIÓ DE LES SEQÜÈNCIES REPETITIVES DEL CDNA PRIMARI Procediment

1. Bloquejar 2 μ g del cDNA amplificat de l'apartat 3.5.5.1., hibridant-lo amb 2 μ g COT-1 DNA humà, 100 ng de DNA ribosomal, 10 ng de DNA linealitzat del vector del cosmidi Supercos1, en un volum total de 10 μ l.

2. Bullir la mescla 5 min a 95°C. Incubar 10 min a 4°C. Afegir 10 μ l de la solució d'hibridació 2x.

3. Incubar 4 hores a 65°C amb agitació suau. La solució de prehibridació té un total de 20 μ l i una concentració de DNA de 200 μ g/ml. Això correspon a un $C_0 t_1/_2$

de 10 mols de nucleòtid x segon x litre⁻¹. Per calcular el $C_0 t_1/2$, com a regla aproximada, podem dir que 1 hr d'hibridació a una concentració de 80 µg/µl té un $C_0 t_1/2$ de 1.

Material i solucions

Solució d'hibridació 2x:
40 mM fosfat de sodi, pH 7,2,
10 mM EDTA,
1,5 M NaCl,
10x solució de Denhardt,
0,2% SDS.

3.5.5.3. PRIMER CICLE DE SELECCIÓ DE cDNA

Procediment

1. Desnaturalitzar 100 ng (5 μ l de 20 μ g/ml) de DNA dels cosmidis biotinilats i marcats radioactivament durant 5 min a 98°C (200 pg cada cosmidi). Incubar durant 4 min a 95°C. Afegir 5 μ l de solució d'hibridació 2x.

2. Mesclar les dues solucions (els cDNA prehibridats de l'apartat 3.5.5.2. i cosmidis biotinilats i desnaturalitzats). El volum final serà de 30 μ l. La hibridació és també en excés de DNA genòmic (50 pg) respecte als cDNA (1 pg).

3. Hibridar tota la mescla durant 54 hores a 65°C amb agitació suau. Això correspon a un $C_0t_{1/2}$ total de 100 mols de nucleòtid x segon x litre⁻¹. La concentració de DNA i els temps d'hibridació poden modificar-se, a fi i efecte d'obtenir el mateix valor final de $C_0t_{1/2}$.

4. Rentar amb solució d'hibridació uns 100 μ l de la solució de Dynabeads M280 (1 mg total), les quals contenen les esferes d'estretavidina. Rentar tres vegades més en solució d'hibridació. Les esferes d'estreptavidina són extretes de la solució de Dynabeads emprant el concentrador magnètic de partícules per un temps de 1 min en cada rentat. Dissoldre les esferes d'estreptavidina fins a 10 mg/ml en tampó d'hibridació.

5. Recuperar aquells cDNA que s'han hibridat amb el DNA genòmic de l'hibridació del pas núm. 3, incubant la hibridació amb les esferes d'estreptavidina durant 15 min a temperatura ambient. Extreure el complex d'esferes d'estreptavidina-biotina-DNA genòmic-cDNA específic de la solució d'hibridació mitjançant el concentrador magnètic de partícules durant 1 min.

104 Material i mètodes-

6. Rentar el complex recuperat amb dos rentats de 15 min en 1 ml de 1x SSC/0,1% SDS a temperatura ambient, seguit de 4 rentats de 15 min en 0,1x SSC/0,1% SDS, a 65°C, emprant el concentrador magnètic.

7. Incubar el complex recuperat de l'últim rentat en 50 μ l de 100 mM NaOH, 10 min a temperatura ambient per deshibridar el cDNA específic del DNA genòmic. Recuperar el cDNA específic deshibridat del DNA genòmic (Fig. 11).

8. Neutralitzar el cDNA recuperat amb 50 μ l de 1 M Tris HCl, pH 7,5. Purificar aquest material mitjançant una columna de cromatografia de Sephadex G-50.

Una novetat important en el protocol presentat anteriorment respecte a altres descrits amb anterioritat és l'obtenció del cDNA específic del complex Fe⁺⁺-estreptavidina-biotinacosmidi mitjançant una desnaturalització alcalina de NaOH i no per calor, ja que aquest darrer també trenca les unions entre l'estreptavidina i la biotina (Fig. 11). D'aquesta manera el material recuperat està lliure de material genòmic, facilitant els processos posteriors.



Figura 11. Mètode alcalí per a la hidròlisi de l'híbrid DNA-cDNA.

Material i solucions

- Concentrador magnètic de partícules (# 120.04 de Dynal®).

3.5.5.4. AMPLIFICACIÓ DE LA PRIMERA SELECCIÓ DE CDNA

Procediment

1. Preparar una PCR de 50 µl volum total, amb 1 µl del cDNA eluit del pas anterior, emprant l'encebador OLIGO3 que és específic de l'adaptador d'amplificació lligat al cDNA. Les condicions de la PCR són estàndards, a una temperatura d'aparellament de 60°C.

2. Separar electroforèticament els diferents productes de la PCR en un gel d'agarosa al 1%. Cal obtenir una cua o *smear* de productes de mida igual o lleugerament inferior a la mida del DNA obtinguda en la PCR de la genoteca original de cDNA (apartat 3.5.5.1.).

Purificar la resta del producte de la PCR mitjançant una columna de Sephadex G Mesurar la concentració de DNA de la mostra purificada.

3.5.5.5. SEGON CICLE DE SELECCIÓ

Procediment

1. Aplicar un segon cicle de selecció, o sia, bloquejar el cDNA específic capturat en el cicle de selecció anterior segons les condicions del pas núm. 1 de l'apartat 3.5.5.2. i tornar-lo a hibridar en les mateixes condicions dels pas núm. 3 de l'apartat 3.5.5.3. amb 100 ng de DNA genòmic biotinilat.

2. Capturar el cDNA específic, rentar i purificar de manera igual al descrit en la primera selecció de cDNA.

3.5.5.6. AVALUACIÓ DE L'ENRIQUIMENT

Procediment

1. Amplificar per la PCR amb l'encebador OLIGO3 les següents fonts de DNA:

- De la genoteca inicial de cDNA no sotmès a cap cicle d'enriquiment.
- Del cDNA seleccionat després d'un primer cicle de selecció.
- Del cDNA seleccionat després d'un segon cicle de selecció.
- Del cDNA seleccionat després d'un tercer cicle de selecció.
- Del cDNA seleccionat després d'un quart cicle de selecció.

- Del material no específic que resulta del 6è rentat després del segon cicle de selecció.

- PCR amb l'encebador OLIGO3, però sense cDNA específic obtingut en el segon cicle de selecció (control negatiu).

- PCR amb 1 μl del cDNA específic després del segon cicle de selecció, però sense encebadors (doble control negatiu).

Amplificar per la PCR amb encebadors específics dels gens següents:

- Un fragment del gen DSCR1 que actua com a control positiu.

- Un fragment del gen del Norrie (cromosoma X) que actua com a control negatiu.

2. Fer una electroforesi de 100 ng de cadascun dels productes amplificats amb l'encebador OLIGO3, i 25 pg dels dos darrers. 3. Transferir el DNA a una membrana de niló o Southern blot.

4. Hibridar la transferència de Southern utilitzant com a sonda un fragment del gen control positiu DSCR1.

3.5.5.7. CLONATGE I CARACTERITZACIÓ DELS CDNA SELECCIONATS

Procediment

1. Amplificar per la PCR els cDNA obtinguts en la segona selecció d'hibridació amb l'encebador OLIGO3CUA (l'encebador OLIGO3 duu en el seu extrem 5' una seqüència (CUA)₄).

2. Dur a terme una reacció de lligament entre el producte de la PCR dels pas anterior i el vector de seqüenciació pAMP10, segons les condicions de l'apartat 3.3.14.

3. Transformar les cèl·lules *E. coli* competents amb el producte de la reacció de lligament del pas anterior.

4. Recuperar les colònies blanques i transferir-les a plaques del tipus elisa de 96 pous (s'ha obtingut 6 plaques de 96 pous, amb un total de 576 clons recombinants).

Material i solucions

- Cèl·lules competents:

"MAX efficiency STBL2™ Competent Cells" (# 90268 de Life Technologies).

- Vector de seqüenciació:

pAMP10 (# 18552-018 de Life Technologies).

3.6. MAPA TRANSCRIPCIONAL

Amb l'objectiu de construir un mapa transcripcional que integri tota la informació física i transcripcional obtinguda, vàrem localitzar en el mapa de cosmidis cadascun dels fragments de les unitats transcripcionals trobades pel mètode de la selecció de cDNA. D'aquesta manera, s'ha obtingut una col·lecció de cDNA localitzats de manera seqüencial al llarg de tot el mapa contigu de cosmidis generats a partir de cada YAC de les tres regions cromosòmiques de l'HC21.

3.6.1. DETERMINACIÓ DE L'ORIGEN DE CADA EST ALS YAC DE L'HC21

S'han fet créixer tots els clons recombinants del pas núm. 4 de l'apartat 3.5.5.7. sobre membranes de niló per a la seva posterior hibridació amb sondes de DNA repetitiu i de DNA ribosomal humà, en experiments paral·lels, per detectar aquells clons que contenien DNA repetitiu i ribosomal, respectivament.

Procediment

1. Després de descartar tots aquells clons que contenien seqüències de DNA ribosomal o de DNA repetitiu, amplificar per la PCR cadascun dels clons de cDNA putatius obtinguts en la tècnica de selecció de cDNA,. L'amplificació de cada clon de cDNA s'ha fet a partir del l'encebador específic de l'adaptador lligat als cDNA (OLIGO3) i no amb els encebadors específics del vector, per a minimitzar el contingut de seqüència de plasmidi en el producte amplificat.

2. Fer una electroforesi en un gel d'agarosa al 1,5% dels productes de la PCR de cadascun dels clons amplificats.

3. Tallar cadascuna de les bandes de DNA dels productes amplificats individualment a partir del gel d'agarosa amb una ganiveta estèril.

4. Utilitzar individualment cadascun d'aquests productes purificats de la PCR per dur a terme una reacció de marcatge radioactiu, emprant hexanucleòtids a l'atzar (dN)₆, segons s'ha esmentat en l'apartat 3.3.7.1.

5. Purificar cada sonda marcada amb una columna de Sephadex G-50.

6. Preparar dot blots amb els DNAs purificats de cadascuna de les següents maxipreparacions de DNA: 1) DNA dels YAC emprats en aquest treball (72H9, 336G11, 238B1, 221B9, 552A3), 2) DNA de la soca de llevat AB1380, a partir de la qual s'han originat els descrits anteriorment, 3) DNA genòmic total d'humà, i 4) H_2O .

7. Hibridar individualment cadascun d'aquests dot blots, emprant com a sonda específica, cadascun dels cDNA purificats i marcats del darrer pas núm. 5. Les condicions d'hibridació són tota la nit a 65°C en tampó fosfat d'hibridació. Les condicions de rentat són les generals (pasos 6-8 de l'apartat 3.3.6.1.).

8. Exposar les membranes a una pel·lícula de raigs X, 2 hores a temperatura ambient.

Materials i solucions

- Tampó fosfat d'hibridació 1M, pH 7,1:

Afegir una solució de 1 M de NaH₂PO₄·H₂O (138 gr/l) a una solució de 1 M Na₂HPO₄·2H₂O (178 gr/l) fins a obtenir un pH de 7,1.

3.6.2. DETERMINACIÓ DE L'ORIGEN DE CADA EST ALS COSMIDIS

Procediment

9. Utilitzar cadascuna de les sondes específiques dels cDNA del pas anterior núm. 5 per a hibridar una membrana de niló, en la qual estan representats tots els cosmidis d'aquell YAC en el que un determinat cDNA ja ha estat localitzat prèviament, segons s'indica en l'apartat 3.6.1.

108 Material i mètodes-

10. Les condicions d'hibridació, rentat i exposició són les mateixes que en l'apartat anterior 3.6.1.

3.6.3. Determinació dels clons redundants.

Els clons que contenen sequències redundants, parcialment o totalment, s'han detectat mitjançant l'hibridació de la membrana amb els 576 clons de la selecció de cDNA, emprant com a sonda cadascun dels EST aïllats.

3.7. CARACTERITZACIÓ D'UN NOU GEN HUMÀ: *MNBH* 3.7.1. AÏLLAMENT DE LA SEQÜÈNCIA CODIFICANT

3.7.1.1. IDENTIFICACIÓ DE L'EST DE MNBH

La seqüència del DNA del clon D1-34 (el clon del pou que corresponia a la filera <u>D</u>, de la columna núm. <u>1</u>, de de la selecció de cDNA amb l'encebador OLIGO<u>3</u>, de la placa núm. <u>4</u>) presentava una alta similitud amb el gen *mnb* de *Drosophila* (Tejedor *et al.*, 1995) mitjançant anàlisis de BLASTN i BLASTX. Fou en aquest moment quan pensàrem en la possibilitat que el clon D1-34 fos realment un fragment del gen humà homòleg de *mnb* de *Drosophila*, i fou a llavors, atès el fenotip tan interessant de la mosca *minibrain*, que decidírem caracteritzar la totalitat d'aquest gen humà.

3.7.1.2. CLIVELLATGE DE GENOTEQUES RECOMBINANTS DE CDNA

Es va amplificar per la PCR el fragment de DNA del clon D1-34 amb l'encebador OLIGO3. Aquest producte de PCR purificat s'ha emprat com a sonda radioactiva en el clivellatge d'una genoteca de cDNA de cervell fetal humà.

3.7.1.2.1. Primer clivellatge

Procediment

1. Determinar el títol la genoteca de cDNA (veure apartat 3.7.1.3. per saber quines genoteques de cDNA s'han emprat) per a determinar la concentració fags mitjançant dilucions seriades en solució d'SM del fags (normalment les dilucions són des de 10^{-3} fins a 10^{-9}). Calcular el títol (ufp/ml) de la següent manera:

ufp/ml = nombre de calbes observades x factor de dilució x $10^3 \,\mu$ l/ml x μ l⁻¹ sembrats

Cal assegurar-se de tenir el suficient nombre de fags com per tenir representat el clon que cerquem, tenint en compte que la quantitat de molècules de RNA que desitgem pot trobar-se a la cèl·lula a una concentració des de 1/10³ fins a 1/10⁶. En general, si es fa

un clivellatge de 500.000 a 1.000.000 clons de cDNA independents, la genoteca tindrà, amb una alta probabilitat, una còpia d'un determinat clon.

Procediment

1. Fer créixer un cultiu d'*E. coli* fins a $DO_{600} = 0,5$ en 20 ml d'LB suplementat amb 0,2% maltosa i 10 mM MgSO₄ a 37°C amb agitació (250 rpm). La maltosa indueix la producció del receptor h, i els ions Mg⁺⁺ també afavoreixen l'absorció.

2. Centrifugar les cèl·lules a 3.000 x g. Dissoldre-les en 4 ml de 10 mM MgSO4.

3. Mesclar 300 μ l del cultiu bacterià amb 50.000 ufp/placa de Petri gran (15 cm), i incubar 20 min a 37°C.

4. Afegir 8 ml de *top* agarosa al 0,7% equilibrada a 45°C al tub de cultiu que conté la mescla de fags i cèl·lules. Transferir immediatament aquesta mescla a plaques sòlides d'LB. Dispensar la mescla uniformement sobre la superfície de la placa.

5. Un cop el top agararosa amb les cèl·lules infectades s'ha solidificat, incubar les plaques a 37°C fins que les calbes cobreixin la superfície, sense arribar a la confluència (8-12 hores).

6. Incubar les plaques 1 hora a 4ºC abans d'aplicar la membrana de niló.

7. Aplicar la membrana de niló de manera orientada a la superfície de la placa durant 45 seg. Deixar-la assecar a temperatura ambient. Aplicar una segona membrana com a rèplica durant 1 min.

8. Desnaturalitzar les membranes en la solució desnaturalitzant durant 6 min (en dos vegades de 3 + 3 min).

9. Neutralitzar les membranes amb solució neutralitzant durant 3 min. Rentar-les amb 2x SSC, 2 min.

10. Assecar les membranes durant 120 min a 80°C per fixar el DNA covalentment.

11. Hibridar les membranes amb una sonda del cDNA de *MNBH* corresponent al producte de la PCR amb els encebadors mnb1D i mnb1R.

Material i solucions

Top agarosa/1 litre
10 g triptona,
8 g NaCl,
6 g agarosa,

Esterilitzar a l'autoclau. Mantenir a 45°C.

- Medi de suspensió (SM):

5,8 g NaCl,

 $2 g MgSO_4 \cdot 7 H_2O, pH 7,5,$

0,01% gelatina.

Bacteri hoste:
C600 *hfl*: per al fag lambda gt10.
Y1090: per al fag lambda gt11.
Solució desnaturalitzant:
500 mM NaOH,
1,5 M NaCl.
Solució neutralitzant:
500 mM Tris·HCl, pH 7,0,
1,5 M NaCl.
Encebadors:
mnbD1: 5'-ctcgggtgtattttggttg-3'.
mnbR1: 5'-cgttttccatctttggtcttc-3'.

3.7.1.2.2. Segon clivellatge

Cal fer un segon clivellatge de cadascuna de les calbes seleccionades com a positives en el primer clivellatge. Es pica la calba positiva i es posa en 1 ml d'SM. A partir d'aquesta dilució, es completa el segon clivellatge, repetint els passos del núm. 1 al 10 de l'apartat 3.7.1.2.1. Caldrà fins i tot dur a terme un tercer clivellatge per a aïllar calbes senzilles. Cadascuna de les seqüències de DNA d'aquests fags són seqüenciades. Amb les noves seqüències de DNA dels fags específics aïllats en el tercer clivellatge es torna a sintetitzar una nova sonda per hibridar un altre cop la genoteca de cDNA de cervell humà. Ha estat necessari dur a terme diversos clivellatges de diferents genoteques de cDNA fins a obtenir tota la seqüència que cercavem del gen *MNBH*.

3.7.1.3. GENOTEQUES DE CDNA

Per a obtenir la sequència de tot el gen MNBH em clivellat les següents genoteques comercials:

1) Genoteca de cDNA de cervell fetal humà de 21 setmanes, construïda a partir d'encebadors oligo-(dT)₁₅ i oligonucleòtids a l'atzar (dN)₆ en el vector lambda gt11 (# HL5015b de Clontech).

2) Genoteca de cDNA de cervell fetal humà de 23 setmanes, construïda a partir d'encebadors oligo-(dT)₁₅ i oligonucleòtids a l'atzar (dN)₆ en el vector lambda gt10 (# HL3003a de Clontech).

3) Genoteca de cDNA de fetge fetal humà de 22-26 setmanes, construida a partir d'oligonucleòtids (dT)₁₅ i oligonucleòtids a l'atzar (dN)₆, en el vector lambda gt11 (# HL503b de Clontech).

4) Genoteca de cDNA de cerebel adult humà, construïda a partir d'encebadors oligo-(dT)₁₅ i oligonucleòtids a l'atzar (dN)₆, en el vector lambda gt10 (# 1128a de ⁻Clontech).

5) Genoteca de cDNA d'ull adult humà, construïda a partir d'encebadors oligo- $(dT)_{15}$, en el vector lambda gt10 (# HL1047a de Clontech).

3.7.1.4. SEQÜENCIACIÓ DEL DNA DELS NOUS CLONS DE MNBH

Cada fragment de DNA s'ha seqüenciat automàticament pel mètode de seqüenciació dels dideoxinucleòtids (Sanger, 1977). Aquest mètode, respecte al de la seqüenciació manual, té les avantatges de no ésser radioactiu, d'ésser molt sensible i de llegir automàticament les seqüències, les quals queden en forma de llistats informàtics. S'ha emprat els seqüenciadors automàtics 373 i 370A (Applied Biosystems). Com a sistema de detecció s'han utilitzat els dideoxinucleòtids, cadascun d'ells marcat amb una fluorescència de color diferent. Aquests dideoxinucleòtids són incorporats en la reacció i impedeixen que la polimerasa continuï allargant la cadena.

Procediment

1. S'utilitzà el protocol recomanat per la casa comercial. S'ha seqüenciat DNA plasmídic i DNA amplificat per PCR. A diferència del DNA d'una minipreparació, el DNA de la reacció de la PCR cal purificar-lo dels altres components de la reacció (encebadors, nucleòtids, tampó) abans de dur a terme la reacció de seqüenciació. Les columnes amb les quals hem obtingut una eficiència en la recuperació del DNA i una qualitat millor d'aquest han estat les columnes del kit de Qiagen Qiaquick-spin. Es recomana purificar una PCR de 50 µl de volum.

2. La quantitat de DNA que actua com a motlle en la reacció de la PCR de seqüenciació varia des de 200 ng (si és un producte de la PCR) i 1 mg de DNA (si és una minipreparació del DNA plasmídic), o sia, >de 10^{10} còpies de DNA motlle. En la PCR de la reacció de seqüenciació s'ha utilitzat el kit "*Taq* DyeDeoxyTM Terminator Cycle Sequencing" amb un volum final de 10 µl en comptes dels 20 µl recomanats per la casa comercial.

3. La temperatura d'aparellament de la PCR és la T_m calculada per a l'encebador emprat, segons la fórmula expressada en l'apartat 3.3.16. El temps d'extensió a 60°C no és necessari que sigui de 4 min, i es pot reduir a 1,5 min amb idèntics resultats.

4. Purificar els productes de la reacció de seqüenciació amb la columna de Sephadex G-50 (TE 10/1).

5. Liofilitzar la mostra.

112 Material i mètodes-

6. Dissoldre el DNA en 5 μl de tampó de càrrega. Desnaturalitzar 2 min a 95°C. Incubar la mostra 5 min en gel.

7. Carregar les mostres en un gel d'acrilamida al 6%. Preequilibrar-lo durant 10 min a 40 W. Carregar les mostres. Córrer les mostres durant 10 hores a 1.700 volts

8- Les dades són recollides i analitzades pels programes d'ordinador Collection i Analysis d'Applied Biosystems.

Materials i solucions

- Kit de "Qiagen Qiaquick-spin" (# 28106 de Qiagen).

- Kit "Taq DyeDeoxy[™] Terminator Cycle Sequencing" (# 401113 d'Applied Biosystems).

- Sephadex G-50 (TE 10/1): pesar 30 g de G-50 i afegir a 300 ml de TE 10/1. Agitar durant 30 min a 65°C. Rentar diverses vegades el sobrenedant amb TE 10/1. Esterilitzar a l'autoclau.

- Tampó de càrrega:

1 volum d'EDTA 50 mM pH 8,0,

1 volum de formamida desionitzada.

3.7.2. AÏLLAMENT DE LA SEQÜÈNCIA SENCERA DEL cDNA 3.7.2.1. AÏLLAMENT DELS EXTREMS 5' I 3'-UTR DEL cDNA

Per determinar les sequències dels 5' i 3' de la regió no traduïda del cDNA (5'-UTR i 3'-UTR) del gen *MNBH* s'ha optat per tres mètodes diferents: 1)RACE, 2) RACE *Marathon*, i 3) clivellatge de genoteques de cDNA. En tot ells s'ha emprat sempre una polimerasa especial (*Advantage*) que permet amplificar llargs fragments de DNA amb una fidelitat millorada (taxa d'error més petita) respecte a la *Taq* polimerasa, ja que té un sistema de correcció de lectura o *proofreading* i un sistema de *Hot Start* incorporat.

Material i solucions

-Advantage cDNA polymerase mix (# 8417-1 de Clontech).

3.7.2.1.1. RACE de cDNA

S'ha emprat la tècnica de RACE (amplificació ràpida dels extrems dels cDNA) per trobar el 5'-UTR del cDNA de *MNBH*. A partir de l'RNA total i de l'RNA poli(A)⁺ de cervell adult humà s'ha sintetitzat la primera cadena de cDNA amb un encebador específic del gen *MNBH*.

Procediment

1. Dur a terme una transcripció inversa emprant el kit "GeneAmp Thermostable r*Tth* Reverse Transcritase RNA PCR Kit" segons les recomanacions de la casa comercial. Atès que aquest kit utilitza una transcriptasa inversa que és termoestable, la reacció de la transcripció inversa s'ha dut a terme a temperatures elevades (60° C), emprant un encebador específic del gen *MNBH* (mnb4R). S'ha utilitzat 500 ng de RNA total i 100 ng de RNA poli(A)⁺.

2. Dur a terme una extracció de la cadena de cDNA específica del gen *MNBH* amb fenol/cloroform i posteriorment amb cloroform. Precipitar el DNA amb etanol al 96%. Rentar les sals amb etanol al 70%.

3. Efectuar una segona purificació emprant les columnes Qiagen de purificació de PCR (esmentades en el punt 3.7.1.4) per a extreure els encebadors.

4. Preparar la reacció de fosforilació de l'encebador Anchor1 segons l'apartat 3.3.10.

5. Purificar l'encebador fosforilat mitjançant una columna de Sephadex G-25.

6. Dur a terme una reacció de lligament entre els sscDNA específics de MNBH i l'encebador Anchor1 fosforilat segons la següent reacció:

7 pmols de l'encebador Anchor1 desnaturalitzat prèviament durant 5 min a 65°C i refredat a 4°C,

1 µl tampó de la lligasa 10x,

1 mM ATP,

1 µl T4 RNA lligasa,

 $1 \mu l DMSO,$

 H_2O fins a 10 μ l volum final.

Incubar la reacció durant més de 16 hores a 4ºC.

8. Amplificar per la PCR, 2 μ l de la reacció de lligament anterior, emprant l'encebador interior o *nested* mnb46R que és específic del 5'-UTR del cDNA de *MNBH* i l'encebador Anchor2 que és complementari a l'encebador Anchor1. Les condicions són les estàndards, llevat que s'ha afegit DMSO a una concentració final del 4%.

9. Sotmetre una alíquota d'aquesta PCR a una segona amplificació seminested amb l'encebador mnb16R que és específic de MNBH i l'encebador Anchor 2

10. Subclonar els productes de la PCR en un vector de seqüenciació. Seqüenciar i analitzar les seqüències.

Material i solucions

- DMSO:

El DMSO, tal com altres cosolvents com el glicerol, milloren la desnaturalització de regions de DNA riques en GC.

- Kit "GeneAmp Thermostable r*Tth* Reverse Transcritase RNA PCR *Kit*" (# N808-0069 de Perkin Elmer)

- Oligonucleòtid:

Anchor1: 5'-cgcgcggccgcctatcgattctggaaccttc-3'.

Anchor2: 5'-gaaggttccagaatcgatagg-3'.

mnbD4: 5'-caggccaagggtctacaattc-3'.

3.7.2.1.2. RACE Marathon

S'ha emprat el kit "Race Marathon" per trobar els 5'-UTR i el 3'-UTR del cDNA del gen MNBH. Aquesta tècnica està basada en el mètode de la RACE. A partir de diferents teixits de RNA i de diferents edats fetals i adults se sintetitza una primera cadena de cDNA, per a després sintetitzar la segona cadena i generar un cDNA d'extrems roms. Els extrems d'aquests darrers dscDNA són lligats a un adaptador asimètric, de manera que un dels dos oligonucleòtid emprats per a la construcció d'aquest adaptador de doble cadena és més curt que l'altre, descobrint una seqüència de DNA de cadena senzilla protuberant, a partir de la qual se sintetitza un encebador d'idèntica sequència. Aquest oligonucleòtid no podrà servir com a encebador fins que no s'hagi sintetitzat la seva cadena complementària, a partir d'un encebador específic de la sequència del gen que coneixem, i que volem estendre. Sovint es realitza una segona PCR interna o nested per guanyar en especificitat. Els productes de la PCR són subclonats en vectors plasmidics per transformar cel·lules d'E. coli. Am l'ajuda del robot es disposa en una membrane de niló cada clon recombinant. Emprant un oligonucleòtid específic de MNBH més interior com a sonda (mnb16R) en la hibridació d'aquesta membrana, s'obtenen aquelles colònies amb DNA específic del gen MNBH. Els clons recombinants positius són següenciats i analitzats. Vàrem aplicar aquesta tècnica per trobar el 5'-UTR de MNBH, a partir de quatre tipus de teixit humà:

- 1) cDNA de fetus humans sencers d'avortaments espontanis de 7 i 8 setmanes.
- 2) cDNA de cerebels adults humans d'edats compreses entre els 22 i 70 anys.
- 3) cDNA de cervell total humà de 37 anys d'edat.
- 4) cDNA de fetge adult humà d'edats entre els 22 i 70 anys.

En el cas del 5'-UTR de MNBH, s'han dut a terme diverses reaccions de RACE Marathon amb l'encebador mnb46R per a la primera PCR, i amb els encebadors mnb47R, mnb80R i mnb86R per a les PCR nested en experiments paral·lels. Cadascun d'aquests encebadors interns es localitzen en exons diferents. Cadascun d'aquests experiments RACE s'ha realitzat amb diferents combinacions d'encebadors i amb cadascuna de les fonts de cDNA mencionats anteriorment. Pel que fa a l'extrem 3'-UTR del cDNA de MNBH s'ha utilitzat la mateixa tècnica, però amb parelles d'encebadors específics del 3' de MNBH: mnbD4-mnb89D.

Materials i solucions

-RACE Marathon Kit:
Cervell Fetal humà de 19-23 setmanes (# 7402-1 de Clontech)
Cervell aldult de ratolí de 9-11 setmanes (# 7450-1 de Clontech)
Fetus humà de 7-8 setmanes (# 7438-1 de Clontech)
Encebadors:
mnb4D: 5'-caggccaagggtctacaattc-3'.
mnb89D: 5'-gtccagttttgtgtgctttactacacag-3'.
mnb46R: 5'-gtttggctggcgacggtcactgtactg-3'.
mnb47R: 5'-ggggaagggagggaagggtcaggtc-3'.
mnb80R: 5'-acaccttctcggtttccagagtc-3'.
mnb86R: 5'-ggggcgctcgcgccgcttc-3'.
mnb86R: 5'-ggggcgctcgcgccgcttc-3'.

3.7.2.1.3. Genoteques de cDNA

Sondes de DNA específiques de les parts 5' i 3' del cDNA de *MNBH* obtingudes per la PCR amb parelles d'encebadors específics de cadascun dels extrems de la part codificant del cDNA de *MNBH*, s'han utilitzat en experiments paral·lels per hibridar genoteques de cDNA de cervell humà fetal, i obtenir així una col·lecció de fags amb seqüència de DNA nova de cadascun dels extrems del fragment del cDNA de *MNBH* que ja coneixiem. Per a corroborar els inicis de transcripció del cDNA de *MNBH* s'han dissenyat uns oligonucleòtids de 60 residus (mnb77R i mnb78R) específics de cada trànscrit (MNBHa i MNBHb, respectivament). Aquests sondes han estat emprades per clivellar les genoteques de cDNA mencionades anteriorment.

Materials i solucions

- Sondes:

mnb78R: 5'-caggatattaataaagagttttgatggttaaggacaagcaggcgaattaa-3'. mnb77R: 5'-cgggagcccgaggctgagactcaccggaggaagcggcgcggggcgcccg-3'.

3.7.3. HIBRIDACIÓ IN SITU

Per a estudiar el patró d'expressió de *Mnbh* s'han utilitzat seccions sagitals i coronals de cervells adults de ratolí i de rata d'edats de 2, 7 i 16 dies (3 animals de cada dia).

Procediment

116 Material i mètodes-

1. Congelar ràpidament el teixit en neu carbònica i fer seccions de 20 μ m amb l'ajuda d'un criòstat a -20°C. Muntar-les sobre un porta de vidre.

2. Assecar i fixar amb paraformaldehid al 4% (tamponat en solució salina de fosfat 0,1 M) i deshidratar les seccions amb etanol al 70% durant 5 min. Després en etanol al 96%.

3. Preparar les sondes d'oligonucleòtids de 40 residus de la sequència de *MNBH* de la cadena sense sentit o *antisense strand* i de la seva complementaria o cadena amb sentit o *sense strand* com a control negatiu. Marcar les sondes radioactivament amb [alfa-³²P]dATP >3000 Ci/mmol) i purificar-les mitjançant columna de Sephadex G-25.

4. Hibridar les membranes tota la nit a 42°C en solució d'hibridació. Rentar posteriorment amb la solució de rentat 2x-0,1x.

5. Deshidratar les seccions en etanol al 70% durant 5 min. Després en etanol al 96%, tots dos amb acetat amònic 0,3 M pH 7,0. Deixar assecar.

6. Exposar les seccions marcades a una pel·lícula de raigs X a -80°C, 4 dies.

Material i solucions

Solució d'hibridació:

0,6 M NaCl,

1x solució Denhardt,

40% de formamida desionitzada,

10 mM Tris HCl pH 7,5,

1 mM àcid etilèdiamintetraacètic,

0,5 mg/ml de tRNA de llevat,

10% dextran sulfat.

- Solució de rentat SSC 1x:

150 mM NaCl,

15 mM citrat de sodi, pH 7,0.

3.8. ANÀLISI DE MICROSATÈL·LITS

3.8.1. ANÀLISI DELS MICROSATÈL·LITS DEL TIPUS (CA)n

3.8.1.1. DETERMINACIÓ DE L'HETEROZIGOSITAT DELS MICROSATÈL·LITS

A partir dels microsatèl·lits del tipus (CA)_n específics de l'HC21 caracteritzats en el nostre Departament per la Dra. A. Bosch *et al.*, (1993a, 1993b, 1994b, 1995a, 1995b i 1995c) s'han dissenyat oligonucleòtids específics per a cada microsatèl·lit. Després de trobar les condicions més escaients per l'amplificació per la PCR, s'han calculat llurs valors d'heterozigositat i PIC mitjançant la seva amplificació i anàlisi per la PCR en 40 mostres de DNA d'individus no relacionats (els pares del pannell de les famílies CEPH).

3.8.1.2. GENOTIPATGE O DETERMINACIÓ AL·LÈLICA

Cadascun dels microsatèl·lits polimòrfics nous de l'HC21 han estat genotipats mitjançant una PCR asimètrica dels DNA dels 532 individus del pannell de referència CEPH. Aquesta PCR es dur a terme amb un encebador marcat radioactivament, el qual es troba a una concentració limitant de 1,5 pmols respecte als 10 pmols de l'encebador no marcat. Les condicions de PCR són 5 min a 92°C, 28 cicles de 20 seg a 95°C, 20 seg a 56-60°C i 20 seg a 74°C. Una extensió final de 5 min a 74°C. Aquestes condicions han estat optimitzades per a un aparell de PCR Perkin Elmer Cetus 9600.

Després de l'amplificació del microsatèl·lit hom barreja 2 μ l del producte de la PCR amb 2 μ l del tampó de càrrega, es desnaturalitza la barreja 10 seg a 96°C i es carreguen en un gel desnaturalitzant d'urea. D'aquesta manera les variants al·lèliques es discriminen electroforèticament a 60 W, durant 3-4 hores segons la mida del fragment amplificat. El gel s'asseca a 80°C i s'exposa a una pel·lícula de raigs X durant 8-24 hores a -80°C.

Material i solucions

- Famílies de referència CEPH

El pannell de referència CEPH (Centre d'Etude du Polymorphism Humain) consisteix en una col·lecció de 40 famílies amb una llarga descendència i seleccionades de diferents regions geogràfiques: 27 provenen dels mormons d'Utah (USA), 10 són d'origen francès, 2 són de procedència veneçolana i 1 família correspon als Amis de Pensilvània (USA). Representen un total de 656 meiosis. El nombre mitjà de fills és de 8,3 encara que cap en té menys de 6. Els quatre avis són presents en 29 pedigrís Els resultats de tots els grups d'investigació implicats en el genotipatge de les famílies CEPH són accessibles per a tothom en el Banc de dades del genoma "GDB" (Cuticchia *et al.*, 1993).

- Pannell d'híbrids somàtics de tot el genoma humà

El pannell núm. 2 d'híbrids somàtics de tot el genoma humà està format per 24 línies d'híbrids rosegador-humà, caracteritzades per diferents laboratoris (Drwinga *et al.*, 1993) i distribuïts pel NINGS (Human Genetic Mutant Cell Repository). Aquesta col·lecció d'híbrids permet localitzar un determinat STS a un determinat cromosoma humà.

- Línia cel·lular específica del cromosoma 21 humà (WAV-17)

Es tracte d'una línia cel·lular híbrida entre ratolí i humà que conté el cromosoma 21 com l'únic component humà (Raziuddin et al., 1984). Aquesta línia s'ha

118 Material i mètodes-

emprat per a confirmar l'origen dels nous STS i EST caracteritzats a l'HC21, ja sigui mitjançant la PCR o la de la transferèfencia de *Southern*.

- Pannell d'híbrids somàtics del cromosoma 21

Representa un pannell de 18 línies cel·lulars híbrides entre hàmster xinès i humà. Cada una de les línies conté un fragment diferent de l'HC21 (Gardiner et al 1988; 1990a). S'han emprat per a la sublocalització cromosòmica dels microsatèl·lits

- YAC del CEPH/GÉNÉTHON

S'han utilitzat els YAC específics de l'HC21 descrits per Chumakov et al., 1992a, 1992b, tan pel que fa a la localització dels microsatèl·lits com pel que fa al material genòmic de partida per a la construcció de la genoteca de cosmidis per a la seva posterior utilització en la tècnica de selecció de cDNA, o bé en la posterior caracterització del gen *MNBH*.

- YAC del JYSE

Com s'ha comentat en l'apartat 1.2.3.5., el mapa de YAC CEPH conté uns forats o *gaps*. S'han utilitzat els YAC caracteritzats pel projecte "JYSE" (Chromosome 21 Joint YAC Screening Effort) (Cox i Shimizu, 1990) per tal d'omplir aquest forats.

- Tampó de la Taq polimerasa10x:

100 mM Tris HCl, pH 8,3,

15 mM MgCl,

500 mM KCl, pH 8,3.

Per alguns microsatèl·lits ha estat necessari variar la concentració de MgCl₂ (des de 1,2 fins a 2,4 mM) i de KCl (des de 40 fins a 80 mM).

- Reacció de la PCR

80 ng de DNA,

200 µM de cada dNTP,

1 U de Taq polimerasa,

1,5 pmols de l'encebador marcat o calent, segons l'apartat 3.3.7.5.

10 pmols de l'encebador no marcat o fred,

 H_2O fins a un volum de 25 μ l.

- Gel desnaturalitzant d'urea al 6%:

30 ml de solució poliacrilamida (acrilamida/bisacrilamida, 19:1) al 40%,

50 ml H₂0 destil·lada,

97 g d'urea,

40 ml de TBE 5x

Ajustar el volum a 200 ml amb H₂0 destil·lada.

Agitar per dissoldre les sals.

Tampó TBE 5x:
450 mM Tris,
440 mM àcid bòric,
9,5 mM EDTA.
Ajustar el pH a 8,3 amb àcid acètic.
Afegir H₂O fins a un volum final de 10 l.
Tampó de càrrega
0,25% de blau de bromofenol (,# B-7021 de Sigma),
40% de sucrosa.

3.8.2. CONSTRUCCIÓ DE MAPES GENÈTICS

S'han seleccionat els marcadors més polimòrfics desenvolupats pel nostre grup i altres marcadors de la base de dades del CEPH com a referència (Weissenbach *et al.*, 1992; Bowcock *et al.*, 1994). Els experiments de lligament genètic s'han dut a terme per la Dra. Assumpció Bosch del grup del Cromosoma 21 del nostre Departament, emprant el programa CRI-MAP (Bosch *et al.*, 1996). La meva tasca s'ha limitat a trobar les condicions d'amplifiació per a la PCR, genotipar cadascun dels microsatèl·lits als 533 individus de les famílies del pannell CEPH, caracteritzar les mutacions *de novo* dels microsatèl·lits genotipats, i corroborar el genotipatge de cada un dels individus que presentaven doble recombinació en el mapa de lligament genètic i d'aquells que eren susceptibles de contenir errors de genotipatge o bé la localització exacta d'alguns microsatèl·lits en el mapa d'híbrids somàtics i en el mapa de YAC.

.

•

.

-

4.1. RESULTATS DIRECTES DE LA TESI

ANTECEDENTS

El genotipatge dels marcadors de DNA polimòrfics del tipus $(CA)_n$ específics de l'HC21 amb les famílies CEPH ens ha permès construir un mapa de lligament genètic i millorar els mapes físics existents. La culminació d'aquest treball arribà amb la localització de tots els marcadors en el mapa físic de YAC i la seva integració en el mapa citogenètic i genètic (veure apartat 4.2.8.). Els mapes genètics han estat claus per al descobriment de moltes malalties hereditàries importants en els darrers anys. Aquest és el cas de la fibrosi quística, la distròfia muscular de Duchenne, la corea d'Huntington, la neurofibromatosi, la síndrome del cromosoma X fràgil, la cistinúria i moltes altres. Pel que fa al mapa genètic de l'HC21, fou el primer i actualment és un dels més densos en el seu contingut d'STS polimòrfics.

Aquests marcadors de DNA hipervariables aïllats i caracteritzats han constituït unes eines imprescindibles per a la construcció sistemàtica i millorament de la resolució dels mapes físics i genètics de l'HC21, i que posteriorment han conformat el marc idoni per a:

1) Estudis de lligament genètic entre un locus i una malaltia localitzada a l'HC21.

2) Estudis de la determinació de l'origen parental de l'HC21 extra en les trisomies del 21.

3) Estudis de les relacions entre el genotip-fenotip de les trisomies parcials de l'HC21 i la DS

4) Estudi genètic de la DS, amb la construcció de mapes físics de clons solapats per a ésser utilitzat posteriorment en la recerca de gens dirigida a determinades regions de l'HC21.

PROJECTE

En el moment de començar el projecte d'aquesta tesi, el nombre de polimorfismes $(CA)_n$ de l'HC21 començava a estar saturat. Malgrat que el mapa genètic de l'HC21 s'havia millorat considerablement en els darrers anys, la construcció de mapes físics d'alta resolució, de mapes transcripcionals i la recerca de gens no havia avançat amb tanta celeritat. La construcció d'un mapa físic d'alta resolució és un prerequisit per a l'avenç dels estudis moleculars de l'HC21. Més concretament, possibilita la construcció immediata d'un mapa de transcripció detallat, el qual és essencial vers la identificació de gens importants en la etiologia de les malalties associades al cromosoma 21 i en la DS. Donat que un dels objectius fonamentals en la investigació de la DS és comprendre les bases moleculars de la seva etiologia i trobar els gens involucrats en aquesta síndrome, el treball que aquí es presenta s'emmarca dins d'aquesta prioritat. Aquest esforç vers l'HC21 està plenament justificada per la rellevància mèdica de les malalties monogèniques d'aquest cromosoma, per l'interès científic pels desordres multifactorials i per la importància social de la DS.

El primer treball presentat, "Cosmid contig and transcriptional map of three regions of human chromosome 21q22: identification of 37 novel transcripts by direct selection", descriu la construcció d'un mapa de transcripció de 2,2 Mb de l'HC21, a fi i efecte d'aïllar gens que puguin estar associats amb la DS. Amb aquest objectiu s'ha construït una genoteca de 502 cosmidis a partir del YAC de 3 regions de l'HC21, cobrint una regió total de 2,2 Mb. S'ha construit també una genoteca de cDNA a partir dels quals s'ha aplicat la tècnica de la selecció de cDNA per aïllar nous EST.Posteriorment, s'ha caracteritzat i localitzat en el nou mapa contigu de cosmidis cadascun dels 120 nous EST aïllats, integrant tota la informació en un mapa transcripcional.

En el segon estudi, "A human homologue of Drosophila minibrain (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region", s'ha caracteritzat un dels EST generats en el mapa de transcripció construït en l'estudi anterior. D'aquesta manera, es presenta la seqüència de la possible regió codificant d'un nou gen humà que l'hem anomenat Minibrain (MNBH), així com la comparació de la seva seqüència aminoacídica amb els seus gens homòlegs de Drosophila i rata. També es determina la seva precisa localització a la DSCR de l'HC21 i s'estudia el patró d'expressió del gen Mnbh en el cervell del ratolí mitjançant la tècnica d'hibridació in situ. Aquest mRNA es localitza de manera generalitzada per tot el SNC, però especialment en l'hipocamp, escorça cerebral, cerebel, el bulb olfactori i varis nuclis hipotalàmics.

En 1996, Banfi et al. identificaren varis EST humans amb un grau elevat de similitud amb gens de Drosophila. Un d'aquest EST mostrava similitud amb el gen minibrain de Drosophila, que fou localitzat al cromosoma 1 humà, mentre que a l'HC21 no hi trobaren cap gen. Així que la pregunta era obvia: ¿Hi ha més d'un gen MNBH en el genoma humà?. Aquesta qüestió es va desenvolupar en el treball presentat amb el títol "Minibrain (MNBH) is a single copy gene mapping to human chromosome 21q22.2", en el qual

es conclou que MNBH és un gen de còpia única i està localitzat en el cromosoma 21, per anàlisis de FISH i de Southern blot.

Finalment, el darrer treball científic sota el títol "Human Minibrain homologue (MNBH): characterization, alternative splicing, differential tissue expression, and overexpression in Down syndrome", caracteritza les seqüències del 5' i 3'-UTR del cDNA del gen de MNBH, i descriu la totalitat de la seqüència del cDNA o full lenght. El treball també descriu dos inicis de transcripció alternatius que originen els trànscrits MNBHa i MNBHb, amb llurs seqüències promotores putatives, així com el seu patró d'expressió diferencial en diferents teixits humans. Es caracteritza cadascún dels límits entre introns i exons, es dissenyen encebadors específics de cada exó per a la seva posterior anàlisi de mutacions. Es descriuen també diferents C-terminals de la proteïna inferida de MNBH que originen les variants MNBH-iso1, MNBH-iso2, MNBH-iso3 i MNBH-iso4. S'analitza també la possible sobreexpressió de MNBH en teixit de cervell de persones amb DS i en el ratolí trisòmic Ts16Dn mitjançant estudis de transferència de Northern blot. 4.1.1. Cosmid contig and transcriptional map of three regions of human chromosome 21q22: identification of 37 novel transcripts by direct selection

Guimerà J, Pucharcós C, Domènech A, Casas C, Solans A, Gallardo T, Ashley J, Lovett M, Estivill X and Pritchard M

Genomics, 45: 59-67 (1997)

El doctorand és l'autor del treball de la preparació, l'anàlisi i caracterització del DNA dels YAC 72H9, 336G11, 238B1, 221B9 i 552A3, de la construcció de les genoteques de cosmidis a partir de cada YAC, de la purificació de la maxiprep de DNA dels 502 cosmidis per gradient de CsCl, de l'aïllament i preparació del l'RNA de cervell fetal humà, de la construcció de la genoteca de cDNA, de l'aplicació de la tècnica de selecció de cDNA i de la comprovació de l'enriquiment, subclonatge dels 576 EST putatius, aïllament dels clons ribosomals i humans, localització precisa de cada EST a un determinat YAC i purificació del DNA de tots els EST per llur assignació a un determinat cosmidi. I ha col·laborat en la construcció del mapa contigu de cosmidis, la seqüenciació i caracterització de cada clon, i en la comprovació de la seva expressió. GENOMICS **45**, 59–67 (1997) ARTICLE NO. GE974861

Cosmid Contig and Transcriptional Map of Three Regions of Human Chromosome 21q22: Identification of 37 Novel Transcripts by Direct Selection

Jordi Guimera,*^{,1} Carles Pucharcós,*^{,1} Anna Domènech,* Caty Casas,* Asun Solans,* Teresa Gallardo,† Jennifer Ashley,† Michael Lovett,† Xavier Estivill,*^{,2} and Melanie Pritchard*^{,2}

*Molecular Genetics Department, Cancer Research Institute, Hospital Duran i Reynals, Avia. de Castelldefels Km 2,7, L'Hospitalet de Llobregat, 08907 Barcelona, Spain; and †Department of Otorhinolaryngology, Department of Molecular Biology, and Department of Oncology and the McDermott Center, University of Texas Southwestern

Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-8591

Received January 10, 1997; accepted June 17, 1997

Human chromosome 21 is associated with many disorders, including Down syndrome (DS). In an effort to identify genes involved in brain development or function and therefore implicated in the mental retardation associated with DS, we chose YACs from three regions of chromosome 21: a region within the socalled "Down syndrome critical region," a region proximal to it, and one distal to it. We made cosmid libraries from these YACs and generated high-resolution physical maps by constructing cosmid contigs. These are the first cosmid contigs on chromosome 21 outside the critical region. The cosmids were used for direct selection of cDNAs to isolate chromosome 21 expressed sequences. We have isolated 45 nonredundant partial cDNAs and mapped these back to the cosmid contigs. We isolated 3 nonoverlapping portions of DSCR1 and a part of GIRK2 and identified 3 nonoverlapping partial cDNAs with similarity to the rat Dyrk gene, which turned out to be the human homologue (MNB) of the Drosophila minibrain gene. Twelve sequences had matches with either STS or EST entries in the databases, including a chromosome 21 EST, a chromosome 21 STS, and 6 unmapped expressed sequence entries. Only 1 sequence resulted in a match with a protein entry. The remaining 25 sequences revealed no similarity to any database entry. All of these partial cDNAs are expressed as determined by Northern blotting or by RT-PCR. © 1997 Academic Press

INTRODUCTION

Numerous genetic diseases have been linked to human chromosome 21, including Alzheimer disease, amyotrophic lateral sclerosis, holoprosencephaly, autoimmune disorders, epilepsy, deafness, and susceptibility to bipolar affective disorder (St George-Hyslop et al., 1987; Siddique et al., 1991; Muenke et al., 1995; Aaltonen et al., 1994; Lehesjoki et al., 1991; Veske et al., 1996; Straub et al., 1994). Chromosome 21 is also associated with syndromic chromosomal aneuploidy, both monosomy and trisomy. Trisomy 21 or Down syndrome (DS) is one of the most frequent congenital defects, affecting 1 in 700 newborns, and is associated with mental retardation, heart disease, anomalies of the gastrointestinal tract, increased risk of childhood leukemia, and defects of the immune and endocrine systems, as well as dysmorphic physical features (Epstein, 1986; Hassold and Jacobs, 1984). With the aim of identifying candidate genes for some of these disorders, intensive efforts have focused on the generation of transcriptional maps (Cheng et al., 1994; Kao et al., 1994; Peterson et al., 1994; Lucente et al., 1995; Tassone et al., 1995; Yaspo et al., 1995; Chen et al., 1996). Highresolution physical maps have not kept pace with the transcriptional maps but it is imperative to develop these to enable expressed sequences to be matched more precisely with disease candidate regions and also to enable the coverage of the transcriptional maps to be assessed. Regions of cosmids or individual cosmids with no corresponding cDNAs can be systematically searched for expressed sequences. Cosmids provide readily manipulable material for the isolation of genes and are also the resource of choice for large-scale sequencing projects. High-resolution cosmid contig maps along with the positional information they provide also help in grouping cDNAs into transcriptional units. For instance, if nonoverlapping partial cDNAs map to the

Sequence data from this article have been deposited with the Gen-Bank Data Library under Accession Nos. U81187–U81231.

¹Both authors contributed equally to the work.

² To whom correspondence should be addressed at the Molecular Genetics Department, Cancer Research Institute, Hospital Duran i

Reynals, Avia. de Castelldefels Km 2,7, L'Hospitalet de Llobregat, 08907 Barcelona, Spain. Telephone: (34-3) 263 0039. Fax: (34-3) 263 2251. E-mail: mpritchard@iro.es or estivill@iro.es.

same sets of cosmids, then it is likely that they belong to the same gene (Gardiner *et al.*, 1995). Such relationships cannot be inferred by mapping at the level of YACs.

We are interested in DS and, in particular, the mental retardation associated with the syndrome. We are directing our efforts to the isolation and mapping of chromosome 21 expressed sequences in the expectation of identifying genes with relevance in brain development and function. By studying the phenotype-genotype relationships in people trisomic for only a portion of chromosome 21, a region called the Down syndrome chromosome region (DSCR), encompassing marker D21S55, was defined, which when present in three copies is responsible for at least a subset of the variable phenotypic features observed in DS, including mental retardation (Delabar et al., 1993). However, this concept has been challenged (Korenberg et al., 1994) by the observation of patients with a DS phenotype caused by trisomy of a portion of chromosome 21 proximal to the so-called critical region. Given this controversy, we chose three regions of the chromosome, one within the DSCR, one proximal, and one distal, from which to isolate expressed sequences. We constructed cosmid contigs from YACs in these regions and used pools of these cosmids for cDNA selection. We have isolated 45 nonredundant partial cDNAs and mapped these back to the cosmid contigs. Among these were parts of two known genes, DSCR1 (Fuentes et al., 1995) and GIRK2 (Ferrer et al., 1995; Patil et al., 1995), and a new chromosome 21 gene MNB (Guimerà et al., 1996). We also identified a previously isolated chromosome 21 EST, a chromosome 21 STS, and 25 novel sequences, i.e., sequences with no matches with entries in the public databases. The remainder had weak similarities with database entries or matched previously unmapped EST sequences. All of these cDNAs are expressed as determined by Northern blotting or by RT-PCR.

MATERIALS AND METHODS

YACs, cosmids, and markers. YACs were from CEPH/Généthon. Five human chromosome 21 YACs (72H9, 336G11, 238B1, 221B9, and 552A3) were tested for chimerism by FISH using inter-Alu PCR products generated from each YAC as a probe. YACs were grown on AHC selective medium and were encapsulated in agarose beads essentially as described (Overhauser and Radic, 1987). For restriction mapping, agarose beads were digested partially or completely with restriction enzymes and electrophoresed on 1% pulsed-field agarose gels in 0.5× TBE. DNA was transferred to Hybond N⁺ membranes in 0.4 N NaOH. For library construction total yeast DNA containing YACs was partially digested with MboI, dephosphorylated, and ligated into the BamHI site of SuperCos (Stratagene). Cosmids were packaged using Stratagene Gigapack II Plus packaging extracts. Clones containing human inserts were identified by screening with a radiolabeled total human DNA probe. Chromosome 21 markers were amplified by PCR using the primer sequences in GDB. Primers specific for the left and right ends of the YAC vector pYAC4 were left, HYAC-C (5'-GCTACTTGGAGCCACTATCGACTA-CGCGAT-3') and YACL-A (5'-GTGATAAATTAAAGTCTTGCGCCT-TAAACC-3'), and right, HYAC-D (5'-GGTGATGTCGGCGATATA-GGCGCCAGCAAC-3') and YACR-A (5'-TCCGTAATCTTGAGA-

TCGGGGT-3'). Marker 374HL was a YAC insert end fragment generated by "bubble" or "vectorette" PCR (Ausubel *et al.*, 1994). The primer pair for 374HL was 374HL.1 (5'-CCGGTTCAAGTAATTCTC-CTG-3') and 374HL.2 (5'-AGCCTGGCAAACATGGTGAA-3').

Generation of cosmid contigs. Cosmid contigs were generated using riboprobe and linear PCR strategies. For riboprobes, 200-400 ng of cosmid DNA was digested with RsaI or HaeIII followed by phenol extraction and ethanol precipitation. The digested DNA was added to a reaction mix containing transcription buffer (40 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 2 mM spermidine, 5 mM NaCl), 10 mM DTT, 0.5 mM ATP/CTP/GTP (Pharmacia), 12 µM UTP, 50 µCi of [a-³²PJUTP, 20 units of RNase inhibitor (Promega), and 15 units of T3/ T7 polymerase (Pharmacia). The reaction was incubated at 37°C for 1 h and added directly to the hybridization. For linear PCR cosmid DNA was digested with RsaI or HaeIII, and 40 ng was used as a template in a 25- μ l reaction containing 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.5 μ M dCTP, 80 μ M dATP/dGTP/dTTP, 20 μ Ci of [α -³²P]dCTP, 20 pmol of the primer (T3 or T7), and 1 unit of Taq polymerase (Perkin-Elmer). Conditions were 94°C for 3 min, then 94°C for 30 s, 50°C for 30 s, 74°C for 30 s for 35 cycles followed by a 5-min extension at 74°C. The riboprobes were hybridized to filters containing gridded cosmid DNAs, and the linear PCR probes were hybridized to filters containing gridded cosmid colonies. PAC clones (UK HGMP Resource Centre) were isolated by probing filters containing a PAC library with partial cDNAs.

Probes and hybridization. Except for cosmid end probes all probes were labeled with $[\alpha^{-32}P]dCTP$ by random priming. Hybridizations, unless stated otherwise, were in 7% SDS/0.5 *M* sodium phosphate buffer (pH 7.0) at 65°C.

cDNA selection. Two cDNA selection experiments were performed essentially as described (Lovett et al., 1991; Morgan et al., 1992). A pilot study using 4 cosmids was undertaken initially. The cDNA source was a commercial human fetal brain cDNA library in λ gt10 (Clontech), and the selected putative cDNAs were subcloned in pBluescript. Thereafter, a pool of 502 cosmid DNAs from the five YACs was used for selection. For this second experiment cosmids were grown individually then pooled in groups of 10 to extract the DNA. The DNAs were banded on CsCl₂ gradients. The cDNA source was a mixture of human fetal brain mRNAs that had been reverse transcribed using random hexamers and oligo(dT). Linkers were ligated, and the cDNAs were amplified and then hybridized with the pool of biotinylated cosmids. The cosmid DNA was biotinylated using the BioNick Labeling System (GIBCO BRL), and to allow monitoring of the cDNA selection process, a small quantity of radiolabeled dCTP was added to the reaction. Two rounds of hybridization were performed. The selected cDNAs were amplified using the linker primer (5'-CTCGAGAATTCTGGATCCTC-3') with a (CUA)₄ tail and subcloned in pAMP10 (GIBCO BRL). Colonies were gridded on a Biomek 1000 station (Beckman) at a high density in duplicate onto nylon filters. To eliminate clones containing human repetitive sequences and ribosomal sequences, the filter was hybridized with a total human DNA probe and a ribosomal probe. DNA from the remaining clones was sequenced with the M13R and M13D primers using fluorescent DyeDeoxy Terminators on an ABI373A automatic DNA sequencer (Applied Biosystems).

Mapping of cDNAs. The inserts of the clones negative for ribosomal and human repeat sequences were amplified with the linker primers and excised from low gelling temperature agarose gels for use as probes on dot blots containing human genomic DNA, yeast genomic DNA, and YAC DNA and on filters containing gridded cosmid colonies or cDNA clones.

Expression studies. For RT-PCR, 0.2 μ g each of fetal brain, liver, and kidney poly(A)⁺ mRNA (Clontech) was reverse transcribed using the Perkin–Elmer GeneAmp kit. Each batch of RNA was tested for genomic DNA contamination by preparing a reaction under the same conditions but without reverse transcriptase and by amplification of known genes that give different-sized products depending on whether RNA or genomic DNA is the template. Aliquots of these reactions were then used in PCR with a primer pair specific for each nonredundant potential cDNA. Northern blots (Clontech) containing

Resultats 131

37 NOVEL HUMAN CHROMOSOME 21-EXPRESSED SEQUENCES



FIG. 1. An idiogram of human chromosome 21 showing the locations of the YACs used in this study and the markers they contain.

 $\operatorname{poly}(A)^+$ mRNA from human fetal or adult tissues were hybridized according to the manufacturer's protocols.

RESULTS

Cosmid Contigs

We selected YACs from three different regions of chromosome 21: 72H9, a 450-kb YAC near AML1; 336G11, 238B1, and 221B9, spanning an approximately 1-Mb region of the DSCR near marker D21S55: and 552A3, a 375-kb YAC near marker D21S53 (Fig. 1). A cosmid library was constructed from each YAC and contigs were generated by hybridizing cosmid end probes to filters containing gridded cosmid DNA or colonies. The orientation of the YACs with respect to genomic DNA was determined by isolating cosmid clones containing the left and right ends of the YAC by hybridization with pYAC4 vector primers. These cosmids were verified by generating a left- or right-end YAC vector-specific PCR product. A map of each individual YAC was constructed using restriction enzymes that cut the DNA infrequently. Digestion of the cosmids with the same enzymes allowed alignment of the cosmids with the YACs and confirmed cosmid overlap. The order of the cosmids was confirmed by their marker content (Fig. 2). In some cases the marker order reported was at odds with the order of the cosmids. For example, in YAC 221B9 the order determined in this work was cen-D21S259-D21S341-D21S342-D21S-³³⁹-tel and not cen-D21S341-D21S342-D21S259D21S339-tel (Shimizu *et al.*, 1994). In some cases gaps in the cosmid contigs could not be filled by rescreening the libraries with end probes but we were able to fill some of these gaps with PAC clones (Fig. 2).

cDNA Selection and Mapping of the Putative cDNAs

Since the cosmid contig construction and the cDNA selection proceeded at the same time it was not possible to use a minimum number of cosmids for cDNA selection. Initially, a pilot study was conducted using the DNA from four YAC 221B9 cosmids to select cDNAs from a commercial fetal brain library. After selection, 150 colonies were analyzed. Only 1 of these clones (B5-1.1) mapped back to the cosmids used for selection. Sixty percent contained mitochondrial sequences, which were presumably selected by the bacterial chromosomal DNA present in the cosmid preparations, which were not purified on CsCl₂ gradients. The remaining clones were primer artifacts. For a second experiment, a pool of 502 human positive cosmid clones from the five YACs was used for the selection of cDNAs from a primary human fetal brain cDNA library. A total of 576 putative cDNA clones were picked and arrayed on nylon membranes. Clones containing human repetitive sequences and ribosomal sequences were eliminated by hybridization of the gridded array with total human DNA and ribosomal probes. The remaining clones were hybridized individually to dot blots containing human genomic DNA and YAC DNA (Fig. 3). In this way 120 putative cDNAs were mapped back to chromosome 21 and assigned to a particular YAC. Curiously, in this experiment no putative cDNAs mapped back to YAC 221B9, even though the full complement of YAC 221B9 cosmids was present. Each putative cDNA was then mapped to the relevant cosmid grid (Fig. 2) and at the same time hybridized to a filter containing the 576 original clones to determine redundancy. These data are summarized in Table 1. Clone redundancy does not enable us to draw any inferences about the abundance of certain cDNAs in the fetal brain library because a pool of cosmids was used for cDNA selection and the redundancy may simply reflect the overrepresentation of some regions in the cosmids or the disparate growth of some cosmids, despite growing all the cultures individually.

Sequencing and Similarity Searches

Each putative cDNA was sequenced, and the sequence was compared in the databases using the BLASTN and BLASTX programs. We isolated three nonoverlapping portions of DSCR1 (Fuentes *et al.*, 1995), a part of GIRK2 (Ferrer *et al.*, 1995; Patil *et al.*, 1995), and three nonoverlapping partial cDNAs with similarity to the rat Dyrk gene, which turned out to be parts of the human homologue (MNB) of the Drosophilaminibrain gene (Guimerà *et al.*, 1996). All the results of the similarity searches are presented in Table 1. In summary, we had 37 putative new chromosome 21 GUIMERA ET AL.



62
А 72H9 336G11 23881 259H11 YEAST 221B9 552A3 265H12 HUMAN H₂O F В RIBOSOMAL H5-34 С G E6-32 F11-36 D н A10-32 F2-36 E I G1-36 D1-34

FIG. 3. Dot blot analysis used to determine the chromosome 21/ YAC origin of the partial cDNAs. (B-I) Examples of dot blot hybridizations with the cDNAs indicated. (A) A key for interpretating the dot blot results. (B) A clone containing a ribosomal sequence. YACs 552A3 and 265H12 overlap, as do 336G11 and 238B1.

expressed sequences. The sequences have been deposited with GenBank under Accession Nos. U81187 to U81231 in the order shown in the table.

Expression Studies

To prove that the putative cDNAs represented expressed sequences, we performed RT-PCR. Primers were synthesized from 40 of the nonredundant cDNAs (with the exception of 1 of the DSCR1 gene fragments, A7-34, we did not perform RT-PCR on fragments of known genes). Through the use of $poly(A)^+$ RNA from a mixture of fetal tissues (brain, lung, liver, and kidney), all these cDNAs were shown to be expressed. The results are presented in Fig. 4.

Northern analysis allowed some evaluation of transcriptional units. For instance, cDNAs F11-36, C4-32, A5-31, and D5-35 mapped to the same series of cosmids as MNB, and yet their sequences did not match any of the minibrain-like entries in the databases. On a Northern blot, MNB detected an approximately 6-kb transcript while these other cDNAs all detected a major transcript of about 4.5 kb (Fig. 5). Thus, we were able to conclude that F11-36, C4-32, A5-31, and D5-35 are probably partial cDNAs from the same gene and, furthermore, are encoded by a gene different from MNB.

DISCUSSION

The construction of a high-resolution physical map is a prerequisite for the advancement of molecular studies of human chromosome 21. Specifically, it provides immediate access to reagents for the construction of a detailed transcriptional map, which is an essential step toward the identification of genes important in the etiology of diseases linked to chromosome 21. Also, since the genome project is now shifting toward the sequencing of entire human chromosomes, with a global chromosome 21 initiative now beginning, cosmid contigs constitute the sequence-ready materials.

Chromosome 21 is very well studied, and several cosmid maps have been reported. The map of Nizetic et al. (1994) covers the entire long arm of chromosome 21, but the cosmids are grouped or binned; no attempt was made to establish a contig. The cosmid maps of Patil et al. (1994), Osoegawa et al. (1996), and Lucente et al. (1995) all focus on the DSCR, a region between the genes CBR and ETS2. Lucente et al. have binned cosmids based on exon content, and although both Patil et al. and Osoegawa et al. have sets of overlapping cosmid clones, in the region corresponding to one studied by us between marker D21S335 and ERG, both contigs have six gaps, with Osoegawa et al. bridging all but one of their gaps with PAC clones. When each tried to compare their maps with those of the other, numerous discrepencies were found both in the order of the groups of contigs (for instance, the order given by Patil et al. is group O, P, Q, R, S, T, and U and by Osoegawa et al. is O, Q, R, T, S, and U) and in their assignment to chromosome 21 (no cosmids in the Patil et al. group P could be integrated into the map of Osoegawa et al.). The cosmids used in all the studies mentioned above were obtained by hybridization of YACs to filters containing gridded cosmids, which probably accounts for the problem of misassignment because false-positive signals frequently occur when YAC probes are used. Indeed, it is estimated that as many as 30% of the



37 NOVEL HUMAN CHROMOSOME 21-EXPRESSED SEQUENCES

FIG. 2. YAC, cosmid, and cDNA map of three regions of human chromosome 21q22. Cosmid contigs are drawn below the YACs, and the cosmids are named with the numerals preceding the dash referring to the YAC from which they originated. Only the minimal tiling path of cosmids is shown. P represents the PAC clones used to fill gaps in the cosmid contigs. The representative cDNAs are shown below the cosmid groups to which they hybridize. Gene symbols are boxed and placed above the corresponding partial cDNAs isolated in this work, with the exception of AMLI, which was mapped by hybridization to the cosmids and is included for reference only. The direction of transcription is indicated with that for GIRK2 taken from Ohira et al. (1997) and AML1 from Y. Groner (pers. comm.). The STSs verified in the YACs and cosmids by PCR are denoted by red circles. We are unable to amplify marker D21S55 and have placed it approximately, relative to D21S337, according to the map position of Ohira et al. (1997). The YACs are drawn to scale, and restriction enzyme sites are indicated by blue arrowheads: NotI (No), NruI (N), BssHII (B), SacII (Sa), SalI (S), MluI (M), and EagI (E). The gap is represented by a shaded arrow.

GUIMERA ET AL.

TABLE 1

Summary of the Mapping and Database Search Results of the Chromosome 21 Partial cDNAs

	No. indep. isolatesª	YAC	Cosmids ^b	BLASTN		BLASTX	
Representative cDNA				Similarity	P(N)	Similarity	P(N)
C2-31	1	72h9	340-345-361-387-398	* None		None	_
A7-34	ī	72h9	340-346-361-363-387-398	Human DSCR1 (3'UTR) gb:U28833	1.7e-98	None	-
E6-32 G3-34	4 1	72h9 72h9	340-346-361-363-387-398 340-346-361-387-398	Human DSCR1 gb:U28833 Human DSCR1 (3'UTR) gb:U28833	6.4e-116 7.8e-98	Human DSCR1 sp:P53805 None	1.1e-55 —
A3-34	17	72h9	340-346-361-363-387	None		None	_
D8-31	1	72h9	340-346-361-363-387-398	None		None	_
F6-35	ī	336G11	48-66-67-87-94	Human EST yc08c11.s1 gb:T63265	3.6e-16	None	-
F11-36	1	336G11	48-66-67-81-87-94	None	-	None	-
C4-32	2	336G11	48-66-67	Human EST HUM21ES111 gb:L25473	2.8e-22	None	_
A5-31	1	336G11	93	None		None	_
C7-32	1	336G11	93	None		None	—
D5-35	1	336G11	45-57-60	None	_	None	_
D1-34	1	336G11	45-60-76-93	Human MNB gb:U52373	1.2e-62	Human NMB gb:U52373	1.3e-30
D8-34	11	336G11-238B1	49-57-50-99 10-43-44	Human MNB gb:U52373	1.2e-71	Human MNB gb:U52373	8.4e-45
D2-34	3	336G11-238B1	49-57-60-99 10-43-44	Human MNB gb:U52373	1.5e-59	Human MNB gb:U52373	3.0e-31
B11-34	2	336G11-238B1	45-49-57-60 10-43-44	None	_	None	-
A4-34	4	336G11-238B1	49-57-60-78-99 10-22-37-43-44	Human STS SHGC-5680 gb:G07613	2.2e-21	None	-
A10-32	2	336G11-238B1	49-57-58-78-85-96-99 10-37-43-44	Human hereditary hemochromatosis region gb:U91328	7.0e-07	None	-
C10-32	1	336G11-238B1	95 30-41	None		None	-
G2-31	2	336G11-238B1	95 30-41	None	-	None	-
B3-33	1	336G11-238B1	95	Human EST HSC30C101 embl:F11685	3.8e-66	None	-
C5-32	1	336G11-238B1	79 30-41-42	None	-	Sacchoromyces cerevisae Centractin sp:P38696	1.3e-00
A4-33	1	336G11-238B1	79 3-23-36-42	Human EST yb67h06.r1 gb:T59992	1.5e-24	Lupinus luteus Early nodulin 75 protein sp:Q06841	8.36-11
F2-36	1	238B1	1-7-15-24-26-29-31-33-40	None		None	-
D2-36	2	238B1	1-7-15-24-26-29-31	None	—	None	-
E7-34	1	238B1	7-24-31-38	None		None	-
D9-32	1	238B1	12-20	Human DNA sequence from cosmid N117B5 gb:Z75744	2.2e-08	None	-
D6-36	1	238B1	13-19	None	<u>*</u>	None	-
H11-32	1	238B1	6-8-14-27-28	Human EST gb:W27025	4.8e-20	None	
E12-34	1	238B1	6-8-14-27-28	None		None	
F11-34	1	238B1	6-8-14-27-28	Human EST HSCZTB032 gb:F04620	2.0e-05	None	-
H3-36	5	238B1	6-8-14-27-28	None		None	
G6-33 E3-32	3 1	238B1 238B1	6-8-14-27-28 6-8-13-14-19	None Human chromosome BAC		Nope None	-
B7-32	12	238B1	6-8-13-14-19	Human cosmid 26hf gb:U09822	5.7e-08	None	-
H5-34	5	238B1	6-8-13-14-19	None	23 <u></u> 13	None	-
B7-36	2	238B1	6-8-13-14-19	None	· · · · · ·	None	
A7-31	5	238B1	6-8-13-19	None	_	None	
E12-35	5	238B1	13-19	Human inward rectifying potassium channel (hiGIRK2) sh:U24660	2.9e-49	None	-
B5-1.1	1	221B9	5-8-30-43-44-47	Human cDNA clone J7194F gb:N56107	3.8e-19	None	-
F4-31	2	552A3	276-279-303-311-325	None		None	-
H5-32	2	552A3	273-278-288-309-312	None	_	None	-
G1-34	3	552A3	282-286-308	None		None	-
G1-36	1	552A3	273-288-312-320	None	—	None	
A9-32	1	552A3	271-288-312-320	None	6 	None	

Note. Only matches with a probability of less than P(N) = 1.0e-06 for BLASTN and P(N) = 1.0e-05 for BLASTX have been reported. ^a The number of independent isolates was determined by hybridization, so these may represent either an identical clone or an overlapping cDNA.

1

^b For simplicity the YAC identifier number has been omitted from the cosmid names.

Resultats 135

37 NOVEL HUMAN CHROMOSOME 21-EXPRESSED SEQUENCES



FIG. 4. RT-PCR of 40 nonredundant clones. To prove the cDNAs represented expressed sequences, RT-PCR was performed on $poly(A)^+$ RNA from a mixture of fetal tissues using primer pairs derived from each partial cDNA clone. The lanes are as indicated. Control reactions without reverse transcriptase and reactions containing water instead of clone DNA were performed for every sample, but here we show five examples. All control lanes were negative.

cosmids mapped by this approach are misassigned (Osoegawa *et al.*, 1996). Given the problems inherent in YAC hybridization, we decided on an alternative approach. We constructed cosmid libraries directly from nonchimeric YACs so all our cosmids are chromosome 21 specific; however, if the YACs contain small deletions, we would not be aware of this using this approach.

Associating a specific DS trait with a particular region of chromosome 21 is extremely difficult given the incomplete penetrance of most DS traits. Furthermore, the study of a low, but steadily increasing, number of rare individuals with partial trisomy 21 provides evidence for a significant contribution to features of the DS phenotype (including the facies, microcephaly, short stature, hypotonia, abnormal dermatoglyphics, and mental retardation) by genes outside the DSCR (Korenberg *et al.*, 1994). So, since the concept of a minimal region for DS has been questioned, in addition to a contig in the DSCR, we also provide the first cosmid contigs outside this region. Three criteria were used to verify the contigs: (i) marker content, (ii) the presence of rare-cutter restriction enzyme recognition sites, and (iii) cDNA content. Overlap in our cosmids was determined by end probe hybridization and by the presence of common DNA restriction fragments. Within the regions reported here, spanning more than 2 Mb, there is one gap.

It has been estimated that 600-1000 genes reside on human chromosome 21, and although many transcriptional maps of this chromosome have been reported, efforts such as ours are still identifying a large number of novel transcripts, even from regions as intensively studied as the DSCR. In this respect, it is noteworthy that only 1 of our partial cDNAs matched with 1 of the 559 potential exons reported by Chen et al. (1996). Clone E12-35 matched an exon with Accession No. X88211 but this was not the best match. As reported in Table 1, the greatest similarity was with human G protein (GIRK). We report the isolation of 45 nonoverlapping partial cDNAs. Among these are 2 known genes (GIRK2 and DSCR1), a new gene (MNB), and an EST previously mapped to chromosome 21. The remaining 37 include cDNAs identified by large-scale cDNA sequencing initiatives but previously unmapped (6 cDNAs), cDNAs with weak matches to human genomic sequences (4 cDNAs), 1 chromosome 21 STS, 1 cDNA with weak similarity to a protein in the databases, and sequences with no similarity to any known sequence (25 cDNAs). The significance of the protein similarity to yeast centractin is presently unclear, but it may give us an indication of the type of sequence we have isolated. Centractin in yeast is an actin-like protein required for proper orientation of the mitotic spindle (Clark and Meyer, 1994). The reason for isolating so few other cDNAs with similarity to proteins may be a consequence of the cDNA selection technique, as the method preferentially favors the 3' ends of genes as they tend to have longer exons (Lovett, 1994).

The reason for selecting only one cDNA from YAC 221B9 is unknown. It may be that this region is poor in genes or that the genes encoded in this region may not be expressed in fetal brain and, thus, not present in the RNA source used for selection. Either of these possibilities could have been the cause of the isolation of artifacts in the pilot study. It is noteworthy that when Osoegawa *et al.* (1996) placed the chromosome 21 ESTs isolated by Cheng *et al.* (1994) and Peterson *et al.* (1994) on their physical map, no cDNAs mapped between marker D21S341 and ERG, the approximate region encompassed by YAC 221B9. Furthermore, like our partial cDNAs, the ESTs from both these studies were selected from fetal brain libraries.

By RT-PCR or by Northern blots, we have shown that all the cDNAs are *bone fide* expressed sequences. With the exception of Tassone *et al.* (1995), other investigators assumed that because they were able to isolate known genes, all their trapped exons or partial cDNAs were expressed sequences and consequently provided no expression data to support this claim.

The task remains now to determine the number of genes these 45 nonoverlapping cDNAs represent. Some



FIG. 5. Northern blots hybridized with five partial cDNAs. Nylon membranes containing 2 μ g of poly(A)⁺ RNAs were hybridized with the partial cDNA clones indicated. D1-34 is a part of *MNB*. (A) Human adult tissues. (B) Human fetal tissues. S/muscle is skeletal muscle.

indication of this may be gained from the results of mapping the cDNAs to the cosmids. If nonoverlapping partial cDNAs map to the same series of cosmids, then it is likely that they belong to the same transcriptional unit. However, such conclusions require the support of more experimental data because both D5-35 and B11-34 map to the same group of cosmids as the *MNB* partial cDNA clones, but from Northern analysis and RT-PCR using primer combinations from neighboring clones, we have demonstrated that neither D5-35 nor B11-34 (data not shown) belongs to the same transcriptional unit as *MNB*.

The chromosome 21 mapping efforts of all the various groups cannot yet be integrated, and so the redundancy of gene isolations cannot be estimated. There may be considerable overlap among the groups and also it is highly probable that there exist multiple fragments from the same gene within each group. However, some overlap is desirable to increase productivity and monitor accuracy. Through these efforts the identification and mapping of every single gene on chromosome 21 is becoming a reality. Once the genes are known, the elucidation of the function of each gene and its possible involvement in disease etiology is tenable.

ACKNOWLEDGMENTS

Although it is only possible to credit the first two authors with an equal contribution to this work, the main body of the work was shared equally among the first five authors. We thank M. Nadal for FISH, A. Puig for technical assistance, J-J. Fuentes for the cosmid library of YAC 72H9, Y. Groner for the AML1 probe, and M. Lynch for discussions and useful comments on the manuscript. We also

thank the UK HGMP Resource Centre for PAC filters and clones. This work was supported by the European Union (CEC/BIOMED2 BMH4-CT96-0554), the Dirección General de Enseñanza Superior (PM95-0106-C02-01), the Fundació Catalana Síndrome de Down/ Marató de TV3-1993, and the National Center for Human Genome Research (HG00368) to M.L. J.G. was granted a HUGO Travel Award.

REFERENCES

- Aaltonen, J., Bjorses, P., Sandkuijl, L., Perheentupa, J., and Peltonen, L. (1994). An autosomal locus causing autoimmune disease: Autoimmune polyglandular disease type I assigned to chromosome 21. Nature Genet. 8: 83–87.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994). "Current Protocols in Molecular Biology," Suppl. 20, p. 6.10, Wiley, New York.
- Chen, H., Chrast, R., Rossier, C., Morris, M. A., Lalioti, M. D., and Antonarakis, S. E. (1996). Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping. *Genome Res.* 6: 747-760.
- Cheng, J-F., Boyartchuk, V., and Zhu, Y. (1994). Isolation and mapping of human chromosome 21 cDNA: Progress in constructing a chromosome 21 expression map. *Genomics* 23: 75-84.
- Clark, S. W., and Meyer, D. I. (1994). ACT3: A putative centractin homologue in S. cerevisiae is required for proper orientation of the mitotic spindle. J. Cell Biol. 127: 129-138.
- Delabar, J. M., Théophile, D., Rahmani, Z., Chettouh, Z., Blouin, J. L., Prieur, M., Nöel, B., and Sinet, P. M. (1993). Molecular mapping of twenty-four features of Down syndrome on chromosome 21. Eur. J. Hum. Genet. 1: 114-124.
- Epstein, C. J. (1986). "The Consequences of Chromosome Imbalance: Principles, Mechanisms and Models," Cambridge Univ. Press, New York.
- Ferrer, J., Nichols, C. G., Makhina, E. N., Salkoff, L., Bernstein, J., Gerhard, D., Wasson, J., Ramanadham, S., and Permutt, A. (1995).

66

Pancreatic islet cells express a family of inwardly rectifying K^+ channel subunits which interact to form G-protein-activated channels. J. Biol. Chem. **270**: 26086-26091.

- Fuentes, J-J., Pritchard, M. A., Planas, A. M., Bosch, A., Ferrer, I., and Estivil, X. (1995). A new human gene from the Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart. *Hum. Mol. Genet.* 4: 1935-1944.
- Gardiner, K., Ichikawa, H., Ohki, M., Patterson, D., and Cheng, J-F. (1995). Localization of cDNAs to a region poorly represented in the CEPH chromosome 21 YAC contig: Candidate genes for genetic diseases mapped to 21q22.3. Genomics 30: 376-379.
- Guimerà, J., Casas, C., Pucharcós, C., Solans, A., Domènech, A., Planas, A. M., Ashley, J., Lovett, M., Estivill, X., and Pritchard, M. A. (1996). A human homologue of *Drosophila minibrain (MNB)* is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. *Hum. Mol. Genet.* 5: 1305-1310.
- Hassold, T., and Jacobs, P. (1984). Trisomy in man. Annu. Rev. Genet. 18: 69–97.
- Kao, F. T., Yu, J., Tong, S., Qi, J., Patanjali, S. R., Weissman, S. M., and Patterson, D. (1994). Isolation and refined regional mapping of expressed sequences from human chromosome 21. *Genomics* 23: 700-703.
- Korenberg, J. R., Chen, X. N., Schipper, R., Sun, Z., Gonsky, R., Gerwehr, S., Carpenter, N., Daumer, C., Dignan, P., Disteche, C., Graham, J. M., Hugdins, L., McGillivray, B., Miyazaki, D., Ogasawara, N., Park, J. P., Pagon, R., Pueschel, S., Sack, G., Say, B., Suchuffenhauer, S., Soukup, S., and Ymanaka, T. (1994). Down syndrome phenotypes: The consequences of chromosomal imbalance. *Proc. Natl. Acad. Sci. USA* **91**: 4997–5001.
- Lehesjoki, A-E., Koskiniemi, M., Sistonen, P., Miao, J., Hastbacka, J., Norio, R., and De La Chapelle, A. (1991). Localization of a gene for progressive myoclonus epilepsy to chromosome 21q22. Proc. Natl. Acad. Sci. USA 88: 3696-3699.
- Lovett, M. (1994). Fishing for complements: Finding genes by direct selection. Trends Genet. 10: 352–357.
- Lovett, M., Kere, J., and Hinton, L. M. (1991). Direct selection: A method for the selection of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci. USA* 88: 9628-9632.
- Lucente, D., Chen, H. M., Shea, D., Samec, S. N., Rutter, M., Chrast, R., Rossier, C., Buckler, A., Antonarakis, S. E., and McCormick, M. K. (1995). Localization of 102 exons to a 2.5 Mb region involved in Down syndrome. *Hum. Mol. Genet.* 4: 1305-1311.
- Morgan, J. G., Dolganov, G. M., Robbins, S. E., Hinton, L. M., and Lovett, M. (1992). The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. *Nucleic Acids Res.* 20: 5173-5179.
- Muenke, M., Bone, L. J., Mitchell, H. F., Hart, I., Walton, K., Hall-Johnson, K., Ippel, E. F., Dietz-Band, J., Kvaloy, K., Fan, C-M., Tessier-Lavigne, M., and Patterson, D. (1995). Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of *SIM2* as a candidate gene for holoprosencephaly, and mapping of *SIM2* to a region of chromosome 21 important for Down syndrome. *Am. J. Hum. Genet.* 57: 1074-1079.
- Nizetic, D., Gellen, L., Hamvas, R., Mott, R., Grigoriev, A., Vatcheva, R., Zehetner, G., Yaspo, M. L., Dutriaux, A., Lopes, C., Delabar, J-M., Van Broeckhoven, C., Potier, M. C., and Lehrach, H. (1994).
 An integrated YAC-overlap and 'cosmid-pocket' map of the human chromosome 21. *Hum. Mol. Genet.* 3: 759-770.
- Ohira, M., Seki, N., Nagase, T., Suzuki, E., Nomura, N., Ohara, O.,

Hattori, M., Sakaki, Y., Eki, T., Murakami, Y., Saito, T., Ichikawa, H., and Ohki, M. (1997). Gene identification in a 1.6-Mb region of the Down syndrome region on chromosome 21. *Genome Res.* 7: 47–58.

- Osoegawa, K., Susukida, R., Okano, S., Kudoh, J., Minoshima, S., Shimizu, N., De Jong, P. J., Groet, J., Ives, J., Lehrach, H., Nizetic, D., and Soeda, E. (1996). An integrated map with cosmid/PAC contigs of a 4-Mb Down syndrome critical region. *Genomics* **32**: 375-387.
- Overhauser, J., and Radic, M. Z. (1987). Encapsulation of cells in agarose beads for use with pulsed-field gel electrophoresis. *Focus* **9:** 8-9.
- Patil, N., Peterson, A., Rothman, A., de Jong, P. J., Myers, R. M., and Cox, D. (1994) A high resolution physical map of 2.5 Mbp of the Down syndrome region on chromosome 21. *Hum. Mol. Genet.* 3: 1811-1817.
- Patil, N., Cox, D. R., Bhat, D., Faham, M., Myers, R. M., and Peterson, A. S. (1995). A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nature Genet.* 11: 126-129.
- Peterson, A., Patil, N., Robbins, C., Wang, L., Cox, D. R., and Myers, R. M. (1994). A transcript map of the Down syndrome critical region on chromosome 21. *Hum. Mol. Genet.* 10: 1735-1742.
- Shimizu, N., Antonarakis, S. E., Van Broeckhoven, C., Patterson, D., Gardiner, K., Nizetic, D., Créau, N., Delabar, J-M., Korenberg, J., Reeves, R., Doering, J., Chakravati, A., Minoshima, S., Ritter, O., and Cuticchia, J. (1994). Report of the Fifth International Workshop on Human Chromosome 21 Mapping 1994. Cytogenet. Cell Genet. 70: 148-164.
- Siddique, T., Figlewicz, D. A., Pericak-Vance, M. A., Haines, J. L., Rouleau, G., Jeffers, A. J., Sapp, P., Hung, W-Y., Bebout, J., Mc-Kenna-Yasek, D., Deng, G., Horvitz, H. R., Gusella, J. F., Brown, R. H., and Roses, A. D. (1991). Linkage of a gene causing amyotrophic lateral sclerosis to chromosome 21 and evidence of a genetic locus heterogeneity. N. Engl. J. Med. **324**: 1281-1384.
- St George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Haines, J. L., Nee, L., Watkins, P. C., Myers, R. H., Feldman, R. G., Pollen, D., Drachman, D., Growden, J., Bruni, A., Foncin, J-F., Salmon, D., Frommelt, P., Amaducci, L., Sorbi, S., Piacentini, S., Stewart, G. D., Hobbs, W. J., Conneally, P. M., and Gusella, J. F. (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science 235: 885-890.
- Straub, R. E., Lehner, T., Luo, Y., Loth, J. E., Shao, W., Sharpe, L., Alexander, J. R., Das, K., Simon, R., Fieve, R. R., Lerer, B., Endicott, J., Ott, J., Gilliam, T. C., and Baron, M. (1994). A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. Nature Genet. 8: 291-296.
- Tassone, F., Xu, H., Burkin, H., Weissman, S., and Gardiner, K. (1995). cDNA selection from 10 Mb of chromosome 21 DNA: Efficiency in transcriptional mapping and reflections of genome organization. Hum. Mol. Genet. 4: 1509-1518.
- Veske, A., Oehlmann, R., Younus, F., Mohyuddin, A., Müller-Myhsok, B., Mehdi, S. Q., and Gal, A. (1996). Autosomal recessive non-syndromic deafness locus (DFNB8) maps on chromosome 21q22 in a large consanguineous kindred from Pakistan. *Hum. Mol. Genet.* 5: 165-168.
- Yaspo, M-L., Gellen, L., Mott, R., Korn, B., Nitzetic, D., Poutska, A., and Lehrach, H. (1995). Model for a transcript map of human chromosome 21: Isolation of new coding sequences from exon and enriched cDNA libraries. *Hum. Mol. Genet.* 4: 1291-1304.

4.1.2. A human homologue of *Drosophila minibrain* (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region

Guimerà J, Casas C, Pucharcós C, Solans A, Domènech A, Planas AM, Ashley J, Lovett M, Estivill X and Pritchard M

Human Molecular Genetics, 5: 1305-1310 (1996)

El doctorand és l'autor de l'aïllament, seqüenciació, localització i caracterització del clon D1-34 (clon que va resultar ésser un fragment del nou gen, que l'hem anomenat *MNBH*), aïllament, caracterització i seqüenciació de tota la regió codificant del gen *MNBH*, de la localització molecular de cadascun dels exons en el mapa transcripcional, i és corresponsable amb la C. Casas dels estudis d'expressió de *MNBH* mitjançant transferències de *Northern*. I és també coautor amb la Dra. Anna M. Planas dels treballs d'aïllament de l'RNA, preparació de sondes i dels estudis d'expressió d'hibridació *in situ* sobre cervells adults de ratolí. © 1996 Oxford University Press

A human homologue of *Drosophila minibrain (MNB)* is expressed in the neuronal regions affected in Down syndrome and maps to the critical region

Jordi Guimerá¹, Caty Casas¹, Carles Pucharcòs¹, Asun Solans¹, Anna Domènech¹, Anna M. Planas², Jennifer Ashley³, Michael Lovett³, Xavier Estivill^{1,*} and Melanie A. Pritchard^{1,*}

¹Molecular Genetics Department, Cancer Research Institute, Hospital Duran i Reynals, Avia. de Castelldefels Km 2.7, L'Hospitalet de Llobregat, 08907 Barcelona, Catalonia, Spain, ²Pharmacology and Toxicology Department, Centro de Investigación y Desarrolo, CSIC, Jordi Girona 18–26, 08034 Barcelona, Catalonia, Spain and ³Department of Biochemistry and the McDermott Centre, University of Texas Southwestern Medical Centre, 5323 Harry Hines Blvd., Dallas, Texas 75235–8591, USA

Received April 23, 1996; Revised and Accepted June 19, 1996

The minibrain (mnb) gene of Drosophila melanogaster encodes a serine-threonine protein kinase with an essential role in postembryonic neurogenesis. A corresponding human gene with similar function to mnb could provide important insights into both normal brain development and the abnormal brain development and mental retardation observed in many congenital disorders. Trisomy 21 or Down syndrome (DS) is the most frequent human birth defect. It is associated with mental retardation and a broad spectrum of physical abnormalities. A region on human chromosome 21 has been designated the Down syndrome critical region (DSCR) and when present in three copies, this is responsible for many of the characteristic features of DS, including mental retardation. We have isolated a human homologue of mnb from the DSCR. MNB encodes a 6.1 kb transcript which is expressed in foetal brain, lung, kidney and liver. Using a human probe, two major transcripts (6.1 and 3.1 kb) were identified in mouse and expression was detected in situ in several regions of the mouse brain, including the olfactory bulb, the cerebellum, the cerebral cortex, the pyramidal cell layer of the hippocampus and several hypothalamic nuclei. This expression pattern corresponds to the regions of the brain that are abnormal in individuals with DS and suggests that overexpression of MNB could have detrimental consequences in DS patients.

INTRODUCTION

Down syndrome (DS) is one of the most frequent congenital defects. There is a broad spectrum of physical abnormalities

associated with the syndrome, including anomalies of the gastrointestinal tract, increased risk of leukemia, defects of the immune and endocrine systems, early onset of Alzheimer's dementia and distinct facial and physical features, but perhaps the most debilitating is a rather severe mental retardation (1,2). In most cases DS is due to three full copies of human chromosome 21 which arise primarily during maternal non-disjunction, but occasionally DS occurs in people carrying unbalanced translocations, which result in the triplication of only a part of chromosome 21. By correlating phenotype with genotype in patients with partial trisomies a region has been defined, named the DSCR (Down syndrome critical region) which, when present in three copies, is responsible for many of the characteristic features of DS including mental retardation (3).

The phenotypic consequences of DS presumably result from the overexpression and subsequent interactions of a subset of chromosome 21 genes and the future challenge is to correlate overexpression of these genes, singly or in combination, with the presence of the DS phenotype. Expression maps of chromosome 21 are currently being developed (4-8) which will help in the identification of all the genes. The first step is to identify the genes in the DSCR and then assess their potential contributions to the pathophysiology of DS. Assigning a function to a gene, particularly in humans, is not simple. Investigators rely on finding clues to function by analysing the expression pattern of a gene, by looking for protein domains or motifs with known functions or by extrapolating from another species in which the function of the homologous gene is known. Any gene with a role in early neurogenesis is potentially important with respect to the abnormal brain development and mental retardation seen in DS. In Drosophila, the mnb gene appears to play an essential role during postembryonic neurogenesis in regulating the numbers of distinct types of neuronal cells (9). Mutant mnb flies are characterised by a marked reduction in size of the adult optic lobes and the central brain hemispheres. This is caused by the abnormal spacing of neuroblasts and hence a reduction in the production of neuronal

f

progeny (10). We have isolated a human homologue of *mnb* from the DSCR and we show, by *in situ* RNA hybridisation studies in mouse brains, that *Mnb* is normally expressed in regions of the brain which are abnormal in individuals with DS.

RESULTS

Isolation of a mnb homologue

In preparation for isolating human chromosome 21 expressed sequences we constructed contiguous cosmid sub-libraries from YACs from various regions of chromosome 21, including YACs from the DSCR. Pools of these cosmids were used for the isolation of partial cDNAs by direct selection (11,12) and for exon trapping experiments (13,14). A total of 576 clones were isolated and arrayed. Of these 576 selected clones, 107 mapped back to human chromosome 21. In total we have 41 non-redundant putative cDNAs, of which 24 are novel, i.e. do not match known genes or ESTs in the databases (manuscript in preparation). We have isolated two known chromosome 21 genes, DSC1 and GIRK2 and another 13 partial cDNAs have significant matches with entries in the databases. Using the blastx program, two non-overlapping partial cDNAs (D7-X4 and D1-34) showed significant similarity to a Drosophila melanogaster serine/threonine kinase (accession no. X70794) [D7-X4 P(N) = $3.0e^{-18}$; D1-34 P(N) = $2.5e^{-36}$]. This serine/threonine kinase is the product of the Drosophila minibrain (mnb) gene. Using the blastn program, D1-34, D7-X4 and a third partial cDNA (D2-34) had almost complete sequence identity $[P(N) = 5.3e^{-60}$ for D1-34] with the Dyrk (Dual specificity Yak1-related kinase) mRNA from Rattus norvegicus (accession no. X79769). In addition, D2-34 also matched a previously mapped chromosome 21 EST (L25452) (4) and a mouse EST (Z31282). Using the partial cDNAs as probes, we subsequently isolated five overlapping clones from a human foetal brain library (Clontech) The combined sequence of these cDNAs spanned 2568 bp and included a complete coding region of 763 amino acids. We have designated the new gene MNB, a human homologue of mnb. Figure 1 shows an alignment of the amino acid sequences of MNB, mnb and Dyrk.

Mapping of MNB

Southern blot analysis of total human genomic DNA digested with *Eco*RI and *Taq*I showed a single band when hybridised with probe D2–34 (data not shown). Therefore, *MNB* is a single copy gene. Hybrid cell line DNA, containing human chromosome 21 as its only human component, showed the same bands when probed with D2–34 but with an additional mouse band. The partial cDNAs (D7–X4 from exon trapping, D1–34 and D2–34 from cDNA selection) were mapped by hybridisation to the chromosome 21 YAC 336G11 and to cosmids shown in Figure 2. YAC 336G11 was shown by FISH to be non chimaeric. This places the *MNB* gene within the proximal half of the DSCR (21q22.2), between *D21S17* and *D21S55*. *MNB* spans the cosmid containing marker *D21S270*, is approximately 200 kb from marker *D21S55* and is proximal to marker *D21S337*.

Expression of minibrain

Northern blot analysis using D1-34 as a probe identified one transcript of 6.1 kb in human foetal brain, liver, lung and kidney

(Fig. 3a). There was no evidence of smaller transcripts in the human tissues with a prolonged exposure of the autoradiograph. In mouse there were two major transcripts of 6.1 kb and 3.1 kb (Fig. 3b), similar to the results of Kentrup *et al.* in rat (15). RNA *in situ* hybridisation studies of mouse brains were carried out using a 40-mer antisense oligonucleotide derived from cDNA D1–34. Expression of *Mnb* was evident in the olfactory bulb, the cerebellum, the cerebral cortex, the pyramidal cell layer of the hippocampus and several hypothalamic nuclei (Fig. 4).

DISCUSSION

During the course of this work Kentrup and co-workers published the identification and functional studies of Dyrk. They speculated that Dyrk is involved in cell cycle control and is the rat homologue of mnb (15). Futhermore, based on the high similarity of Dyrk with the human EST (L25452), Kentrup et al suggested that the human homologue of Dyrk maps to 21q22.2. This human homologue of Dyrk is MNB. Human MNB is expressed as a 6.1 kb transcript and there is no evidence of a smaller transcript analogous to the 3.1 kb transcript found in rat and mouse. The Drosophila mnb gene encodes three alternatively spliced transcripts (5.5, 4.4 and 4.2 kb) (9). It is unknown whether the two transcripts observed in rodents are alternative splicing products. The deduced amino acid sequence of our partial MNB cDNA exhibits structural features which are shared with Dyrk. The core domains which contain amino acids found in the catalytic sites of protein kinases (16) are identical with the exception of two residues and both proteins have a potential nuclear translocation signal (Fig. 1).

Genes which show temporal or high levels of expression during the development of the central nervous system may be of special importance in DS, especially in the pathogenesis of mental retardation. In Drosophila, the mnb gene appears to play an essential role during postembryonic neurogenesis in regulating the numbers of distinct types of neuronal cells (9). Mutant mnb flies are characterized by a marked size reduction of the adult optic lobes and the central brain hemispheres. This is caused by the abnormal spacing of neuroblasts and hence a reduction in the production of neuronal progeny (10). The mnb gene encodes a serine-threonine protein kinase which is expressed in distinct neuroblast proliferation centres during Drosophila postembryonic neurogenesis. The minibrain kinases (mnb, MNB, and Dyrk) share sequence similarity with the cyclin-dependent kinases, which are known to regulate cellular proliferation, suggesting a role for mnb in the correct mitosis of neuroblast progeny (9). Although the overall scheme of neuronal development is quite different in invertebrates and vertebrates, molecular studies on vertebrate neurogenesis have revealed a remarkable evolutionary conservation of the cellular mechanisms of neuronal development (17). Moreover, cyclin-dependent kinases are known to regulate cellular proliferation in various species, suggesting a more universal regulatory mechanism (18). It is conceivable that MNB has a role in the processes which generate neuronal cells in the brain during postembryonic development.

The detection of *MNB* in the DSCR on chromosome 21 suggests that it may be involved in the altered neuronal development observed in DS. Although in *Drosophila* the *mnb* phenotype was due to a reduction in the level of expression of the *mnb* gene, we expect that in DS, *MNB* is overexpressed. At a gross morphological level, DS brains are smaller than normal (19) and there is a decrease in the number of neurons. Neuronal number is

Human Molecular Genetics, 1996, Vol. 5, No. 9 1307

MNB Dyrk mnb	MHTGGETSACKPSSVRLAPSFSFHAAGLQMAGQMPHSHQYSDRRQPNISD MHTGGETSACKPSSVRLAPSFSFHAAGLQMAAQMPHSHQYSDRRQPNISD
MNB	QQVSALSYSDQIQQPLINQVMPDIVMLQRRMPQTFRDPATAPLRKLSVDL
Dyrk	QQVSALSYSDQIQQPLTNOVMPDIVMLQRRMPQTFRDPATAPLRKLSVDL
mnb	MHHHSSPSSSSEVRAMQARIPNHEREPASGPERKLSVDL
MNB	I K T Y K H I N E V Y Y A <mark>K K K R F</mark> H Q Q G C G - DD S S H K K E R K V Y N D G Y D D D N Y D Y I V
Dyrk	I K T Y K H I N E V Y Y A <mark>K K K R F</mark> H Q Q G Q G - DD S S H K K E R K V Y N D G Y D D D N Y D Y I V
mnb	I K T Y K H I N E V Y Y A K K K R F A Q Q T Q G D D D S S N K K E R K L Y N D G Y D D D N H D Y I I
MNB Dyrk mnb	+ KNGEKWMDRYEIDSLIGKGSFGQVVKAYDRVEQEWVAIKIIKNKKAFLNQ KNGEKWMDRYEIDSLIGKGSFGQVVKAYDRVEQEWVAIKIIKNKKAFLNQ KNGEKFLDRYEIDSLIGKGSFGQVVKAYDHEEQCHVAIKIIKNKKPFLNQ
MNB	AQIEVRLLELMNKHDTEMKYYIVHLKRHFMFRNHLCLVFEMLSYNLYDLL
Dyrk	AQIEVRLLELMNKHDTEMKYYIVHLKRHFMFRNHLCLVFEMLSYNLYDLL
mnb	AQIEVRLLEMMNRADAENKYYIVKLKRHFMWRNHLCLVFELLGYNLYDLL
MNB	RNTNFRGVSLNLTRKFAQQMCTALLFLATPELSITHCDLKPENILLCNPK
Dyrk	RNTNFRGVSLNLTRKFAQQMCTALLFLATPELSIIHCDLKFENILLCNPK
mnb	RNTNFRGVSLNLTRKFAQQLCTALLFLSTPELNIIHCDLKPENILLCNPK
MNB	R SAIKIVDFGSSCQLGQRIYQYIQSRFYRSPEVLLGMPYDLAIDMWSLGC
Dyrk	R SAIKIVDFGSSCQLGQRIYQYIQSRFYRSFEVLLGMPYDLAIDMWSLGC
mnb	R SAIKIVDFGSSCQLGQRIYHYIQSRFYRSFEVLLGIQYDLAIDMWSLGC
MNB	ILVEMHTGEPLFSGANEVDQMNKIVEVLGIPPAHILDQAPKARKFFEK - L
Dyrk	ILVEMHTGEPLFSGANEVDQMNKIVEVLGIPPAHILDQAPKARKFFEK - L
mnb	ILVEMHTGEPLFSGCNEVDQMNKIVEVLGMPPKYLLDQAHKTRKFFDKIV
MNB	P D G T W N L K K T K D G K R E F K P P G T R K L H N I L G V E T G G P G G R R A G E S G H T V A D
Dyrk	P D G T W S L K K T K D G K R E Y K P P G T R K L H N I L G V E T G G P G G R R A G E S G H T V A D
mnb	A D G S Y V L K K N Q N G - R K Y K P P G S R K L H D I L G V E T G G P G G R R L D E P G H S V S D
MNB	Y L K F K D L I L R M L D Y D P K T R I Q P Y Y A L Q H S F F K K T A D E G T N T S N S V F T S P A
Dyrk	Y L K F K D L I L R M L D Y D P K T R I Q P Y Y A L Q H S F F K K T A D E G T N T S N S V S T S P A
mnb	Y L K F K D L I L R M L D F D P K T R V T P Y Y A L Q H N F F K R T A D E A T N T S G A G A T A N A
MNB	– M E C S Q S S G T T S S T S S S S G G S S G T S N S G R A R S D F T H Q H R H S G C H F T A A V Q
Dyrk	– M E C S O S S G T T S S T S S S G G S S G T S N S G R A R S D F T H O H R H S G G H F A A A V Q
mnb	G A G G S G S S C A G G S S G G V G G G L G A S N S S S G A V S S S – – – – S A A A P T A A T A
MNB	A M D C E T H S P Q V R Q Q F P A P L G W S G T E A P T Q V T V E T H P V Q E T T F H V A P Q Q N A
Dyrk	A M D C E T H S P Q V R Q O F P A P L G W S G T E A P T Q V T V E T H P V Q E T T F H V A P Q Q N A
mnb	A A T A A G S S G S G S S V G G G S S A A Q Q Q Q A M P L P L P L P L P L P P P
MNB	L H H H H G N S S H H H H H H H H H H H H H H H G Q Q A L G N R T R P R V Y N S P T N S S S T Q D S M E
Dyrk	L H H H H G N S S H H H H H H H H H H H H H H H H G Q Q A L G N R T R P R V Y N S P T N S S S T Q D S M E
mnb	L A G P G G A S D G Q C H D D R R *
MNB Dyrk mnb	V G H S H H S M T S L S S S T T S S S T S S S T G N Q G N Q A Y H N R P V A A N T L D F G Q N G A V G H S H H S M T S L S S S T T S S S T S S S T G N Q G N Q A Y Q N R P V A A N T L D F G Q N G A
MNB Dyrk mnb	M D V N L T V Y S N P R Q E T G I A G H P T Y Q F S A N T G F A H Y M T E G H L T M R Q G A D K E E M D V N L T V Y S N P R Q E T A I A G H P T Y Q F S A N T G F A H Y M T E G H L T M R Q G A D R E E
MNB Dyrk mnb	SPMTGVCVQQSPVASS* SPMTGVCVQQSPVASS*

Figure 1. Alignment of minibrain kinase sequences generated with the Wisconsin Package programs Gap, Pileup and Prettybox. Only amino acid identities are shaded. *nmb* is the *Drosophila minibrain* (accession no. X70794); *Dyrk* is a rat *minibrain* homologue (accession no. X79769). *MNB* is a human *minibrain* homologue from chromosome 21 (accession no. U52373). The complete coding sequences of MNB, mnb and Dyrk are shown. A dash indicates spacing between amino acids to achieve best alignment. The blastx amino acid identity between the *MNB* sequence and *Dyrk* was 99% [P(N) = 0.0], and between *MNB* and *mnb* was 69% [P(N) = $3.3e^{-262}$]. The core domain which contains amino acids found in the catalytic sites of protein kinases (16) lies between the plus signs. In this region there are two amino acid differences between *MNB* and *Dyrk*. A potential nuclear translocation signal is indicated by black shading. An asterisk represents a stop codon.

.

1308 Human Molecular Genetics, 1996, Vol. 5, No. 9



Figure 2. Localisation of human *MNB* to the DSCR on human chromosome 21. Schematic representation of the 21q22.2 region, STS map and maps of YACs 336G11 and 238B1. Cosmids from the contig involving the *MNB* sequences are shown. The partial cDNAs, indicated by an arrow, are below the cosmids they map to. The direction of transcription of *MNB* is shown. The STS verified in the cosmids by PCR are indicated by a filled circle.



Figure 3. Northern blot analysis of *MNB*. Nylon membranes containing 2 μ g of poly(A)⁺ RNAs were hybridized with the human clone D1–34 which corresponds to *MNB*. (a) Northern analysis of human foetal mRNA. (b) Northern analysis of adult mouse mRNA. In mouse lung a third band of 2.5 kb was detected. The lower panel in (a) and (b) shows hybridization with a β -actin probe.

reduced in distinct regions, including the cochlear nuclei, cerebellum, hippocampus, the cholinergic neurons of the basal forebrain, the granular layers of the cerebral cortex, and in areas of the brain stem (20). These abnormalities occur in regions where the *Mnb* gene is normally expressed and are consistent with the view that altered expression is in some way detrimental. During the search for the gene which is responsible for the weaver phenotype in mice (*Girk2*), Patil *et al.* assessed a partial sequence, which they called *Mmb* (mouse minibrain), as a possible candidate (21). *Mnb* and *Mmb* are probably the same gene.

Although a critical region on chromosome 21 has been defined which, when present in three copies, is responsible for the main features of DS, including mental retardation (3), others have challenged this concept (22). In spite of this controversy, the major efforts to identify genes are being centred on this region (5,6). Several new genes with potential relevance in brain development and/or function have been isolated from the DSCR (*GIRK2*, *SIM2* and *MNB*) and from the region just proximal (*DSC1*) (21,23–25). The generation of transgenic mice that overexpress each of these genes, either singly or in combination, should permit an evaluation of their involvement in the pathophysiology of DS. In a recent report, transgenic mice were generated that overexpress *Ets2*, a transcription factor and proto-oncogene which is located on chromosome 21 within the distal boundary of the DSCR. These mice develop skeletal anomalies reminiscent of those observed in DS (26).

It is presumed that the structural alterations observed in the brain, together with the accompanying functional changes may account for the subsequent physiological and cognitive abnormalities associated with DS. Therefore, it is likely that genes involved in neurogenesis and which have altered expression in DS might account, at least partly, for the alterations that lead to mental retardation. The location of *MNB* in the DSCR together with its probable function in neurogenesis, supports *MNB* as a strong candidate gene to produce some of the neurological abnormalities present in DS patients. With the help of neuropathological, neurochemical and behavioural studies in transgenic animals, we may be able to dissect the components contributing to the mental retardation and to the complexity of the DS phenotype.

MATERIALS AND METHODS

YACs and cosmids

Selected human chromosome 21 YACs were tested for chimaerism by FISH. YACs were grown on AHC selective medium and were encapsulated in agarose beads using a modification of the method described in ref. 27. Total yeast DNA containing YACs was partially digested with *MboI* and ligated into the *Bam*HI site of *SuperCos* (Stratagene). Cosmids were packaged using Stratagene Gigapack II Plus packaging extracts. Clones containing human inserts were identified by screening with a radiolabelled total human DNA probe. Cosmid contigs were generated using riboprobe and linear PCR strategies.

Human Molecular Genetics, 1996, Vol. 5, No. 9 1309



Figure 4. Expression of *Mnb* mRNA in the brain of the adult mouse using a 40-mer antisense probe derived from the human *MNB* sequence. Saggital sections of the brain are shown. (C) and (F) are the same as (B) and (E) stained with crystal violet. Signals were evident in olfactory bulb (olf), cerebellum (ce), cortex (ctx) and piriform cortex (pir), hippocampus (hip) and hypothalamus (hyp). A 40-mer sense oligonucleotide was used as a control under the same conditions, giving no signal in the hybridization experiments performed.

Probes and hybridization

Probes were labelled with $[\alpha^{-32}P]dCTP$ by random priming. Colony and plaque hybridizations were in 7% SDS/0.5 M sodium phosphate buffer (pH 7.0). cDNA probes were hybridized to filters containing the hybrid cell line WAV17, containing human chromosome 21 as its only human component (28).

cDNA selection and exon trapping

A pool of 502 cosmid DNAs from the libraries made from seven human chromosome 21 YACs (manuscript in preparation) was used for cDNA selection, essentially as described (11,12). The cDNA source was human foetal brain mRNA which had been reverse transcribed using random hexamers and oligo dT. Linkers were ligated and these cDNAs were then hybridized with the pool of biotinylated cosmids. The selected cDNAs were amplified using the linker primer (5'-CTCGAGAATTCTGGATCCTC-3') with a (CUA)₄ tail and subcloned in pAMP10 (GIBCO-BRL). Exon trapping was as described (13) using a pool of 12 overlapping, non-redundant cosmids from YAC 336G11 subcloned in pSPL3 (14). Amplified exons were directionally subcloned in pAMP1 (GIBCO-BRL). Using a Biomek 1000 station (Beckman), colonies were gridded at a high density onto nylon filters. Clones were sequenced with the M13R and M13D primers using fluorescent DyeDeoxy Terminators on an ABI373A automatic DNA sequencer (Applied Biosystems).

RNA analyses

Northern blots (Clontech) containing $poly(A^+)$ mRNA from human foetal tissues or adult mouse tissues were hybridized with D1–34 according to the manufacturer's protocol. RNA *in situ* hybridization studies on sections of mouse brain were carried out with a 40-mer antisense oligonucleotide (5'-GGAATACCCA-GAACTTCCACTATTTTATTCATCTGATCTA-3'), essentially as described previously (25), except that following hybridization, sections were washed twice in 1 × SSC (150 mM sodium chloride and 15 mM sodium citrate, pH 7.0) at 52°C for 1 h each. A 40-mer sense oligonucleotide was used as a control under the same conditions, and gave no signals in the hybridization experiments performed.

ACKNOWLEDGEMENTS

We thank M. Dierssen for useful contributions to the manuscript, M. Lynch for useful comments, M. Nadal for FISH and M. L. Yaspo for advice on the application of exon trapping. This work 1310 Human Molecular Genetics, 1996, Vol. 5, No. 9

was supported by the European Union (Grants CEC/BIOMED GENE-CT93-0015, GENE-CT93-0037 and GENE-PL95-0554); Fundació Catalana Síndrome de Down / Marató de TV3 - 1993; the Fondo de Investigaciones Sanitarias de la Seguridad Social (94/1905E) (Spanish Government); National Center for Human Genome Research HG00368 to ML. JG was granted a HUGO Travel Award.

REFERENCES

- Hassold, T. and Jacobs, P. Trisomy in man. (1984) Annu. Rev. Genet., 18, 69–97.
- Epstein, C.J. (1986) The consequences of chromosome imbalance: principles, mechanisms and models. *Cambridge University Press, New York*.
- Delabar, J.M., Théophile, D., Rahmani, Z., Chettouh, Z., Blouin, J.L., Prieur, M., Nöel, B. and Sinet, P.M. (1993) Molecular mapping of twenty-four features of Down syndrome on chromosome 21. *Eur. J. Hum. Genet.*, 1, 114-124.
- Cheng, J-F., Boyartchuk, V. and Zhu, Y. (1994) Isolation and mapping of human chromosome 21 cDNA: progress in constructing a chromosome 21 expression map. *Genomics*, 23, 75–84.
- Peterson, A., Patil, N., Robbins, C., Wang, L., Cox, D.R. and Myers, R.M. (1994) A transcript map of the Down syndrome critical region on chromosome 21. *Hum. Mol. Genet.*, 10, 1735–1742.
- Lucente, D., Chen, HM., Shea, D., Samec, SN., Rutter, M., Chrast, R., Rossier, C., Buckler, A., Antonarakis, SE. and McCormick, MK. Localization of 102 exons to a 2.5 Mb region involved in Down syndrome. (1995) *Hum. Mol. Genet.*, 4, 1305–1311.
- Yaspo, M-L., Gellen, L., Mott, R., Korn, B., Nitzetic, D., Poutska, A. and Lehrach, H. (1995) Model for a transcript map of human chromosome 21: isolation of new coding sequences from exon and enriched cDNA libraries. *Hum. Mol. Genet.*, 4, 1291–1304.
- Kao, F.T., Yu, J., Tong, S., Qi, J., Patanjali, S.R., Weissman, S.M. and Patterson, D. (1994) Isolation and refined regional mapping of expressed sequences from human chromosome 21. *Genomics*, 23, 700–703.
- Tejedor, F., Zhu, X.R., Kaltenbach, E., Ackermann, A., Baumann, A., Canal, I., Heisenberg, M., Fischbach, K.F. and Pongs, O. (1995) Minibrain: a new protein kinase family involved in postembryonic neurogenesis in Drosophila. *Neuron*, 14, 287–301.
- Fischbach, K.F. and Heisenberg, M. (1984) Neurogenetics and behaviour in insects. J. Exp. Biol., 112, 65–93.
- Lovett, M., Kere, J. and Hinton, L.M. (1991) Direct selection: A method for the selection of cDNAs encoded by large genomic regions. *Proc. Natl Acad. Sci. USA*, 88, 9628–9632.
- Morgan, J.G., Dolganov, G.M., Robbins, S.E., Hinton, L.M. and Lovett, M. (1992) The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. *Nucleic Acids Res.*, 20, 5173–5179.
- Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., Sharp, P.A. and Housman D.E. (1991) Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl Acad. Sci. USA*, 88, 4005–4009.

- Church, D.M., Stotler, C.J., Rutter, J.L., Murrell, J.R., Trofatter, J.A. and Buckler, A.J. (1994) Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nature Genet.*, 6, 98–105.
- Kentrup; H., Becker, W., Heukelbach, J., Wilmes, A., Schuermann, A., Huppertz, C., Kainulainen, H. and Joost, H.G. (1996) Dyrk: A dual-specificity protein kinase with unique structural features whose activity is dependent on tyrosine residues between subdomains VII and VIII. J. Biol. Chem., 271, 3488–3495.
- Hanks, S.K., Quinn, A.M. and Hutter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science*, 241, 42–52.
- Purves, D. and Lichtman, J.W. (1992) Early events in neural development. In Principles of neuronal development. (ed Sinauer) pp. 3–72.
- Nigg, E.A. (1995) Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *Bioessays*, 17, 471–480.
- Kemper, T.L. Neuropathology of Down Syndrome. (1989) In *The Psychobiology of Down Syndrome* (ed Nadel, L.) The MIT Press pp.269–289.
- Becker, L., Mito, T., Takashima, S. and Onodera, K. (1991) Growth and development of the brain in Down Syndrome. In *The Morphogenesis of Down Syndrome* The Wiley-Liss NY. Prog. Clin. Biol. Res., 373, 133–153.
- Patil, N., Cox, D.R., Bhat, D., Faham, M., Myers, R.M. and Peterson, A.S. (1995) A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nature Genet.*, 11, 126–129.
- Korenberg, J.R., Chen, X.N., Schipper, R., Sun, Z., Gonsky, R., Gerwehr, S., Carpenter, N., Daumer, C., Dignan, P., Disteche, C., Graham, J.M., Hugdins, L., McGillivray, B., Miyazaki, D., Ogasawara, N., Park, J.P., Pagon, R., Pueschel, S., Sack, G., Say, B., Suchuffenhauer, S., Soukup, S. and Ymanaka, T. (1994) Down syndrome phenotypes: The consequences of chromosomal inbalance. *Proc. Natl Acad. Sci. USA*, **91**, 4997–5001.
- Chen, H., Chrast, R., Rossier, C., Gos, A., Antonarakis, S.E., Kudoh, J., Yamaki, A., Shindoh, N., Maeda, H., Minoshima, S. and Shimizu, N. (1995) Single-minded and Down syndrome? *Nature Genet.*, 10, 9-10.
- Dahmane, N., Charron, G., Lopes, C., Yaspo, M-L., Maunoury, C., Decorte, L., Sinet, P.M., Bloch, B. and Delabar, J-M. (1995) Down syndrome-critical region contains a gene homologous to *Drosophila sim* expressed during rat and human central nervous system development. *Proc. Natl Acad. Sci. USA*, 92, 9191–9195.
- Fuentes J-J., Pritchard M.A., Planas A.M., Bosch A., Ferrer I. and Estivil, X. (1995) A new human gene from the Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart. *Hum. Mol. Genet.*, 4, 1935–1944.
- Sumarsono, S.H., Wilson, T.J., Tymms, M.J., Venter, D.J., Corrick, C.M., Kola, R., Lahoud, M.H., Papas, T.S., Seth, A. and Kola, I. (1996) Down's syndrome-like skeletal abnormalities in *Ets2* transgenic mice. *Nature*, 379, 534–537.
- Overhauser, J. and Radic, M.Z. (1987) Encapsulation of cells in agarose beads for use with pulsed-field gel electrophoresis. *Focus*, 9, 8–9.
- Raziuddin, A., Sarkar, F.H., Dutkowski, R., Shulman, L., Ruddle, F.H. and Gupta, S.L. (1984) Receptors for human alpha and beta interferon but not gamma interferon are specified on human chromosome 21. *Proc. Natl Acad. Sci. USA*, 81, 5504–5508.

4.1.3. Minibrain (MNBH) is a single copy gene mapping to human chromosome 21q22.2

Guimera J, Pritchard M, Nadal M and Estivill X

Cytogenetics and Cell Genetics, 77: 182-184 (1997)

El doctorand és l'autor de la recerca de similituds entre les sequències de Drosophila, MNBH i l' EST R38268. És també responsable dels estudis de transferència de Southern, així com de l'obtenció de la sonda específica de DNA del gen MNBH per als estudis de FISH realitzats per la Marga Nadal. Cytogenet Cell Genet 77:182-184 (1997)

Cytogenetics and Cell Genetics

Minibrain (MNBH) is a single copy gene mapping to human chromosome 21q22.2

J. Guimera, M. Pritchard, M. Nadal, and X. Estivill

Molecular Genetics Department, Cancer Research Institute, L'Hospitalet de Llobregat, Barcelona, Catalonia (Spain)

Abstract. We have identified a human homologue of the Drosophila *mnb* gene (MNBH) on chromosome 21, while another study mapped an EST clone (R38268) with similarity to *minibrain* to human chromosome 1. This report describes the mapping of MNBH to a single locus on human chromo-

some 21q22.2 by FISH and Southern blotting. Comparison of the similarities between the two sequences and *mnb* demonstrates that MNBH on chromosome 21 is the true human homologue of Drosophila *mnb*.

The minibrain gene of Drosophila melanogaster (mnb) encodes a serine/threonine protein kinase with an essential role in postembryonic neurogenesis (Tejedor et al., 1995). Flies with mutations in mnb are characterized by a reduction in size of the central brain hemispheres due to a decrease in the production of neuronal progeny, caused by abnormal spacing of neuroblasts (Tejedor et al., 1995). We recently identified a human MNBH gene (Guimera et al., 1996), which because of its location on human chromosome 21 in the so called "Down Syndrome Critical Region" is likely to be implicated in Down syndrome (DS). At a gross morphological level, DS brains are smaller than normal and have a decreased number of neurons (Kemper et al., 1989). Using human MNBH cDNA as a probe on mouse brain sections, we showed that the murine counterpart (Mnbh) is expressed in regions of the mouse brain which correspond to the regions of the human brain affected in DS (Guimera et al., 1996). By analogy with its function and mutant phenotype in Drosophila and its expression pattern in mouse,

Received 13 January 1997; manuscript accepted 11 March 1997.

KARGER

E-mail karger@karger.ch Fax + 41 61 306 12 34 http://www.karger.ch © 1997 S. Karger AG, Basel 0301–0171/97/0774–0182\$12.00/0 This article is also accessible online at: http://BioMedNet.com/karger MNBH is hypothesized to play a role in the mental retardation observed in DS.

Banfi et al. (1996) have identified several human ESTs that show similarity to Drosophila genes. One of these ESTs showed similarity to Drosophila *mnb* and was mapped to human chromosome 1. So the question arises, is there more than one human MNBH gene? Here we demonstrate that the human *mnb* homologue is a single copy gene located on chromosome 21 at 21q22.2 and that the level of similarity between the chromosome 1 sequence and mnb is focused in the serine/threonine kinase catalytic domain of mnb.

Materials and methods

Sequence similarity searches

Similarity searches were performed using the BLASTX program. The sequences compared were the WashU-Merck EST (accession no. R38268), Drosophila *mnb* (accession no. X70794) and the chromosome 21 MNBH (accession no. U52373). Sequences were aligned using the GAP program from the GCG package.

1

Southern analysis

For human chromosome 21, we grew the mouse-human hybrid cell line WAV17, which has chromosome 21 as its only human component (Raziuddin et al., 1984). 3 μ g of total human DNA and 15 μ g of WAV17 DNA were digested with *Eco*RI and *TaqI*, run on an agarose gel, blotted onto a Hybond-N+ membrane and hybridized with a partial MNBH cDNA (D1-34, Guimerà et al., 1996).

FISH studies

2 µg of a 2.8-kb MNBH cDNA generated by RT-PCR from human foetal brain poly A+ RNA were labeled with biotin 16-dUTP (Boehringer Mann-

Supported by the European Union (Grants CEC/BIOMED GENE-CT93-0015, GENE-CT93-0037 and GENE-PL95-0554); Fundació Catalana Síndrome de Down/Marató de TV3 - 1993; and the Fondo de Investigaciones Sanitarias de la Seguridad Social (94/1905E) (Spanish Government).

Request reprints from Dr Xavier Estivill, Molecular Genetics Department, Cancer Research Institute, Avia. Castelldefels Km. 2.7, 08907 L'Hospitalet de Llobregat, Barcelona, Catalonia (Spain); telephone: (34-3) 263 0039; fax: (34-3) 263 2251; e-mail: estivill@iro.es

heim) in a standard nick-translation reaction and purified by gel filtration. 400 ng of labeled probe were precipitated along with 1 μ g of Cot1 DNA and 1 μ g of salmon sperm DNA and the pellet resuspended in a hybridization mix containing 50% formamide and 50% (20% dextran sulfate in 12 × SSC). 10 μ l of the hybridization mix was applied to the slide and after heat denaturation, the slide was incubated overnight in a humid chamber at 37 °C. Posthybridization washes were performed in three changes of 50% formamide, 1 × SSC at 42 °C, followed by two changes of 2 × SSC at 42 °C. For detection of the signal, the slide was incubated at 37 °C with Avidin-FITC (Fluorescein isothiocyanate, Vector Laboratories) for 20 min, and then washed in two changes of 4 × SSC, 0.1% Tween 20 at 37 °C. The signal was amplified once using biotinylated Anti-Avidin D and Avidin-FITC (Vector Laboratories). Metaphases were studied under a fluorescence microscope equipped with the appropriate filter set and images were captured with the Cytovision software (Applied Imaging).

Results and discussion

Banfi et al. (1996) identified a human EST (R38268) with similarity to minibrain by searching the EST databases with Drosophila genes. A cDNA encompassing this EST was isolated and mapped to human chromosome 1. In an effort to evaluate which gene is the more likely mnb homologue, we compared the BLAST results obtained when each human sequence was aligned with Drosophila mnb. To make the comparison meaningful we used a fragment of the chromosome 21 MNBH gene equivalent in size (347 bp) to EST R38268 and at the same relative position in the Drosophila gene. The BLASTX output comparing the chromosome 21 MNBH sequence and EST R38268 sequence with Drosophila mnb resulted in values of $P = 3.0 \text{ e}{-16} \text{ vs} P = 8.0 \text{ e}{-53}$ and scores of 150 vs 432 for EST R38268 and the chromosome 21 MNBH gene, respectively. The alignment of the amino acid sequences is shown in Fig. 1. These results indicate that the gene mapped to chromosome 21 is more similar to Drosophila mnb than EST R38268. Furthermore, MNBH and mnb are similar at the nucleotide level, since by BLASTN the P value obtained for MNBH versus mnb was 2.4 e-54, while BLASTN revealed no similarity between EST R38268 and mnb when P was limited to 1.

The Drosophila *mnb* protein has a kinase core domain spanning residues 100–360. The region of similarity with EST R38268 resides in this domain. Indeed, a BLASTX search using this EST matches 195 entries in the non-redundant database and these are all protein kinases. The similarity to mnb observed with the chromosome 21 MNBH protein extends beyond this functional domain. Across the 763 amino acids of MNBH, there is 69% identity with a BLASTX *P* value of 3.3 e–262 and a score of 1640 (Guimera et al., 1996).

The genes identified by Banfi et al. (1996) and by Guimerà et al. (1996) both represent single copy genes. Banfi et al. mapped their gene unequivocally to 1q32 by FISH, whereas we observed hybridization only to chromosome 21q22.2 when using a 2.8-kb MNBH cDNA probe (Fig. 2). Of 100 metaphases, 26 showed a twin signal on both chromosome 21 homologues, 49 showed a twin signal on only one chromosome 21 and the remaining 25 did not show any signal. These hybridization scores are within the normal range observed when using small cDNA probes. There was no consistent signal on any other chromosome. We obtained a single band on a Southern blot

Percent Similari	ty: 93.043 Percent Identity: 83.478
Dro. mnb 7	<pre>4 KLYNDGYDDDNHDYIIKNGEKFLDRYEIDSLIGKGSFGQVVKAYDHEEQC</pre>
Human #21 MNB	KVYNDGYDDDNYDYIVKNGEKWMDRYEIDSLIGKGSFGQVVKAYDRVEQE
Dro. mnb	HVAIKIIKNKKPFLNQAQIEVKLLEMMNRADAENKYYIVKLKRHFMWRNH
Human #21 MNB	WVAIKIIKNKKAFLNQAQIEVRLLELMNKHDTEMKYYIVHLKRHFMFRNH
Dro. mnb Human #21 MNB	LCLVFELLSYNLYDL 188 : LCLVFEMLSYNLYDL
Percent Similari	ty: 49.123 Percent Identity: 24.561
Dro. mnb 74	.KLYNDGYDDDNHDYIIKNGEKFLDRYEIDSLIGKGSFGQVVKAYDHEEQ
EST R38268	PQ*MXGMXMQMGPIFMYLETI*LIDM.RC*KXLAKGSFGQVXRVYDHKLR
Dro. mnb	CHVAIKIIKNKKPFLNQAQIEVKLLEMMNRADAENKYYIVKLKRHFMWRN . : :::: . . . :. ::: ::.: :
EST R38268	QYVALKMVRNEKRFHRQAAEEIRILEHLKKQDKTGSMNVIHMLESFTFRN
Dro. mnb	HLCLVFELLSYNLYDL 188

Fig. 1. GAP alignment comparing the chromosome 21 MNBH and EST R38268 to Drosophila mnb. The amino-acid numbering is according to the conceptual open reading frame of the *mnb* cDNA (X70794).

HVSWPLNSEHRPL*AY

EST 838268



Fig. 2. Fluorescence in situ hybridization of the human MNBH 2.8-kb cDNA probe showing hybridization only to human chromosome 21 (21q22.2).

of human genomic DNA upon digestion with *Eco*RI and *Taq*I when a partial cDNA sequence (D1-34) of MNBH was used as a probe under non stringent washing conditions (Fig. 3). The 8.5-kb *Eco*RI fragment observed for hybrid WAV17 corresponds to the murine counterpart of MNBH as detected by the analysis of murine DNA (not shown).

In summary, from the information available, *mnb* is more similar to the chromosome 21 MNBH gene than to EST R38268. Furthermore, the similarity with MNBH and *mnb* extends over the entire coding region and therefore, it is more likely that MNBH is the true human homologue of the Drosophila *mnb* gene. It is probable that EST R38268 matches Drosophila *mnb* because it possesses a protein kinase domain.



Fig. 3. Southern blot showing a single human band upon digestion of genomic DNA with *Eco*RI and *TaqI* using a partial MNBH cDNA as a probe (D1-34). WAV17 shows the same pattern of bands, the extra band in the *Eco*RI digest is the murine counterpart of MNBH, as detected by the analysis of mouse genomic DNA (not shown).

Thus, caution must be exercised when assigning homology based on similarities within functionally conserved protein domains. The approach of Banfi et al. (1996) to identify human genes and assign them some functional significance by analogy with their orthologues in Drosophila is very elegant and should help the identification of genes involved in human disease. However, it is likely that several of the ESTs identified in the work, whilst having some sequence similarity to Drosophila genes, are not necessarily their human homologues. More sequence information is required before any conclusions may be drawn one way or the other about homology.

References

- Banfi S, Borsani G, Rossi E, Bernard L, Guffanti A, Rubboli F, Marchitiello A, Giglio A, Coluccia E, Zollo M, Zuffardi O, Ballabio A: Identification and mapping of human cDNAs homologous to Drosophila mutant genes through EST database searching. Nature Genet 13:167-174 (1996).
- Guimera J, Casas C, Pucharcòs C, Solans A, Domènech A, Planas A M, Ashley J, Lovett M, Estivill X, Pritchard M: A human homologue of *Drosophila minibrain* (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. Hum Mol Genet 5:1305–1310 (1996).
- Kemper TL: Neuropathology of Down Syndrome, in Nadel L(ed): The Psychobiology of Down Syndrome, pp 269-289 (The MIT Press Cambridge 1989).
- Raziuddin A, Sarkar FH, Dutkowski R, Shulman L, Ruddle FH, Gupta SL: Receptors for human alpha and beta interferon but not gamma interferon are specified on human chromosome 21. Proc natl Acad Sci, USA 81:5504-5508 (1984).
- Tejedor F, Zhu XR, Kaltenbach E, Ackermann A, Baumann A, Canal I, Heisenberg M, Fischbach KF, Pongs O: Minibrain: a new protein kinase family involved in postembryonic neurogenesis in Drosophila. Neuron 14:287–301 (1995).

4.1.4 Human Minibrain homologue (MNBH): characterization, alternative splicing, differential tissue expression, and overexpression in Down syndrome

Guimerà J, Casas C, Estivill X and Pritchard M

Sotmès a publicació

El doctorand és l'autor de la caracterització completa de la seqüència del gen *MNBH*, incloent els extrems 5'-UTR i 3'-UTR del cDNA de *MNBH*, de l'aïllament i caracterització dels diferents trànscrits de *MNBH* (MNBHa i MNBHb), les seves seqüències promotores putatives i llur patró d'expressió en estudis de transferència de *Northern*. És també el responsable de la identificació i determinació dels límits dels exons-introns, de la organització i localització genòmica de cadascun dels exons en el mapa transcripcional, de la caracterització dels possibles dominis estructurals de la proteïna i la seva comparació 6, de la identificació dels possibles dominis estructurals de la proteïna i la seva comparació amb la de rata i de *Drosophila*, i de la caracterització dels polimorfismes en la seqüència de DNA. És coautor amb la Dra M. Pritchard de l'aïllament de l'RNA i dels estudis de sobrexpressió del gen *MNBH* en teixit de cervell de persones amb la síndrome de Down, així com en el cervell del ratolí trisòmic Ts65Dn. I és coautor amb la Caty Casas de la caracterització de les variants de transcripció que afecten a la seqüència codificant del gen *MNBH*, les quals generen les isoformes putatives MNBH-iso1, MNBH-iso2, MNBHiso3 i MNBH-iso4. Human *Minibrain* homologue (MNBH): characterization, alternative splicing, differential tissue expression, and overexpression in Down syndrome

Jordi Guimera, Caty Casas, Xavier Estivill¹ and Melanie Pritchard^{1,2}

Medical and Molecular Genetics Center-IRO, Hospital Duran i Reynals, Barcelona, Catalonia, Spain

Running title: Characterization of the human Minibrain homologue

¹To whom correspondence should be addressed at:

Medical and Molecular Genetics Center-IRO, Hospital Duran i Reynals, Avia de Castelldefels, Km 2.7, L'Hospitalet de Llobregat, 08907 Barcelona, Catalonia, Spain. Telephone: 34-3-260.77.75. Fax: 34-3-260.77.76. E-mail: melanie.pritchard@med.monash.edu.au; estivill@iro.es

² Current address: IRD, Monash Medical Centre, Clayton, Victoria, Australia

ABSTRACT

The human homologue (MNBH) of the Drosophila minibrain gene maps on human chromosome 21 within the Down syndrome (DS) critical region, and is within the region minimally deleted in chromosome 21-linked microcephaly. As a first step in gaining insight into the role that MNBH may have in human neurogenesis, and as a lead up to the development of mouse models for MNBH overexpression, we have characterized the gene at the molecular level. We describe here the MNBH full-length transcript, alternative splicing, expression profile, and genomic organization. The full-length cDNA of MNBH is 5.2 kb and is composed of 17 exons spanning 150 kb. between markers D21S335 and D21S337. Transcripts MNBHa and MNBHb arise from the use of different first exons in the 5'-UTR, and are differentially expressed. MNBHa is expressed ubiquitiously in a broad spectrum of tissues and is apparently under the control of a CpG island. MNBHb is only expressed in heart and skeletal muscle, and is apparently under the control of a TATA-like box. Four alternative splicing events affecting the C-terminus of the protein yield at least 4 isoforms of MNBH (MNBH-iso1, MNBH-iso2, MNBH-iso3 and MNBH-iso4). A PEST sequence, potentially involved in the rapid degradation of the protein, is present in all the isoforms. A histidine repeat and a serine/threonine domain are only present in the largest form of the protein (MNBH-iso1). MNBH was overexpressed 1.5 fold in DS brains, and Dyrk (the murine counterpart of MNBH) about 2 fold in the brains of Ts65Dn mice. The information provided here should be valuable for MNBH mutation studies and aid in the development of DS animal models.

Keywords: Human Minibrain, Down syndrome, Chromosome 21, DSCR, Alternative splicing, Protein isoforms

INTRODUCTION

One of the fundamental goals of Down syndrome (DS) research is to identify the genes that take part in its pathophysiology. There is evidence, albeit controversial, that three copies of the chromosome 21 segment 21q22.2, (a relatively small region around marker *D21S55* of less than 3 Mb) are sufficient to cause many of the phenotypic features of DS, including the severe mental retardation (Rahmani *et al.*, 1989; Delabar *et al.*, 1993). This subchromosomal fragment called the Down syndrome critical region (DSCR) has been the subject of intense gene finding efforts. So far, 101 genes (GDB, 14 July 1998) have been identified on chromosome 21, but the exact biological roles of most of these genes and therefore their potential contributions to DS pathophysiology remain largely unknown.

Two of these genes, *MNBH* (human minibrain) (Guimera *et al.*, 1996) and *SIM2* (single minded 2) (Chen *et al.*, 1995) were cloned from the DSCR and have aroused intense interest as DS candidate genes, primarily due to the mutant phenotypes of their orthologues in *Drosophila melanogaster*. Both *mnb* and *sim* genes are essential for proper brain development in *Drosophila*. *MNBH* and *SIM2* are expressed in the developing human central nervous system (Guimera *et al.*, 1996; Dahmane *et al.*, 1995) and both genes (when overexpressed) have been proposed to play a role in causing the developmental delay or mental retardation seen in DS.

In *Drosophila*, the development of the adult central nervous system requires a precise and reproducible pattern of neuroblast proliferation during postembryonic neurogenesis. The *mnb* gene encodes a serine-threonine protein kinase, which appears to play an essential role during postembryonic neurogenesis in regulating the numbers of distinct types of neuronal cells in neuroblast proliferation centers. Mutations in the *Drosophila minibrain* gene cause an abnormal spacing of neuroblasts in the outer proliferation center of the larval brain, due to a decreased production of neuronal progeny. As a consequence, the adult *mnb* brain exhibits a specific and marked size reduction of the optic lobes and central brain hemispheres (Tejedor *et al.*, 1995). Similarly, in humans, at a gross morphological level, DS brains are smaller than normal and there is a decreased number of neurons in distinct regions, including the cerebellum, hippocampus, the cholinergic neurons of the basal forebrain and the granular layers of the cerebral cortex. In the mouse brain *Dyrk* is expressed in regions that correspond to those in the human brain that are affected in DS (Guimera *et al.*, 1996).

Dyrk, the murine counterpart of MNBH (96.6% nucleotide similarity and 99% amino acid identity) is a dual specificity protein kinase that catalyzes its autophosphorylation on both serine/threonine and tyrosine residues. The extensive structural similarity between the minibrain kinases (MNBH, Dyrk and mnb) and the cyclin-

dependent kinases, which are known to regulate cellular proliferation, could be an indication of functional relatedness.

Functional screening of 2 Mb of human chromosome 21q22 implicates *MNBH* in the learning defects associated with DS. Behavioral studies of mice harboring human transgenes, revealed that an extra copy of the genes contained on a 180 kb fragment of YAC 152F7 resulted in learning and memory deficits, which were reminiscent of several of the cognitive defects observed in DS (Smith *et al.*, 1997). So far, *MNBH* is the only known gene on this fragment. This revelation, in addition to the *Drosophila mnb* mutant phenotype, and the correlation observed between the expression pattern of *Dyrk* in the mouse brain and the areas affected in the human DS brain, support the hypothesis that the overexpression of *MNBH* has detrimental consequences in DS.

MNBH has also been proposed as a candidate gene for microcephaly, as it maps within the chromosomal segment (D21S11-D21S55) minimally deleted in monosomy 21 associated with microcephaly, intrauterine growth retardation and low-set ears (Matsumoto *et al.*, 1997).

In an initial effort towards elucidating the biological function of MNBH, we report the isolation of fulllength cDNA isoforms and the pattern of expression of *MNBH* transcripts in a broad range of human tissues, and we demonstrate that MNBH is overexpressed in the fetal DS brain. We also report the genomic organization of the *MNBH* gene, including the characterization of the exon-intron structure. We have also defined the transcription initiation sites and sequenced the putative promoter regions of *MNBH*, which will be important for designing constructs with homologous promoters for transgenic studies.

MATERIAL AND METHODS

cDNA library screening

Approximately 1 x 10^6 recombinant phages from a human fetal brain cDNA library in λ gt10 (Clontech), and a human fetal liver cDNA library in λ gt11 (Clontech), were screened with a 585 bp PCR product derived from the 3' part of the coding region of *MNBH* or a 174 bp PCR fragment from the 5' end (Guimerà *et al.*, 1996). The same libraries were screened a second time with 50-55 bp specific probes (mnb77D and mnb78D), derived from transcripts *MNBHa* and *MNBHb*, respectively. Hybridization was in 0.5 M sodium phosphate, pH 7.2/7% SDS, at 65°C overnight. Membranes were washed with 2x SSC/0.5% SDS for 15 min at 45°C, 1x SSC/0.5% SDS for 30 min at 65°C.

Rapid amplification of cDNA ends (RACE)

RACE was carried out using the Marathon cDNA Amplification kit (Clontech) from human fetal brain, adult cerebellum and whole fetus with primers mnb4D and nested primer mnb89D for the 3'-ends and primers mnb46R and nested primers mnb47R, mnb80R and mnb86R for the 5'-ends. The products were subcloned into pBluescript T-vector for sequencing. T-vector was prepared essentially as described (Marchuk *et al.*, 1991).

RACE was also performed using primer mnb4R in a first-strand cDNA synthesis reaction from human adult brain $poly(A)^+$ RNA. The cDNA products were purified twice through centricon columns, then ligated to the phosphorylated primer anchor1 in a standard T4 RNA ligase reaction. PCR was carried out with the anchor2 primer and the nested *MNBH*-specific primer mnb46R. A second round of PCR was performed with anchor2 and mnb16R. Products were subcloned in pBluescript T-vector and sequenced.

RT-PCR

Total RNA from human brain, liver and muscle was isolated using the RNeasy Kit (Qiagen). Poly(A)⁺ RNA was from Clontech. Reverse transcription at 60° C was done using 2 µg/100 ng of each total RNA/RNA poly(A)⁺ with *MNBH*-specific primers or random hexamers (following the manufacturer's recommendations) and the rtTh thermostable enzyme (Reverse transcriptase RNA PCR kit, GeneAmp[®], Perkin Elmer). One fiftieth of each cDNA first strand synthesis was submitted to cycle-amplification using gene specific primers by conventional PCR.

Exon identification, amplification and mapping

Intron/exon boundaries were identified and the precise lengths of the introns of the *MNBH* gene were determined by comparing the *MNBH* cDNA sequence to the genomic DNA sequences available at http://www-alis.jst.go.jp/gdv/kitasato.html#CH21. Intronic primers were designed and tested on genomic DNA and the

products sequenced with the same primers. PCR was performed in 25 μ l volumes with 200 ng human genomic DNA, 10 pmoles of each primer, 200 μ M dNTPs, 1.5 mM MgCl₂, 1 U Taq polymerase, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, and 0.01% (w/v) gelatin. PCR amplifications were performed in a Perkin-Elmer 9600 thermal cycler, using a hot start procedure. Initial denaturation of samples was at 95°C for 4 min followed by 32 cycles for all PCR primer pairs.

The genomic sequence containing the *MNBH* gene was scanned to uncover new putative exons by computer searching with the GENSCAN program (Kulp *et al.*, 1996). The predicted exons were analysed by the BLASTN and BLASTX programs. For mapping, each amplified exon was excised from a gel, purified through a Qiagen gel extraction column, then radiolabeled by random priming. Gridded filters containing cosmids from YAC 336G11 (Guimera *et al.*, 1997) were hybridized with each exon at 65°C.

Northern blotting

Total RNA from human fetal brains and adult Ts65Dn and control mouse brains was prepared as described above. Telencephalon from 2 DS brains or 2 non DS brains at 20 weeks gestation was mixed prior to RNA isolation. Ten μ g of each RNA and 3 μ g of RNA ladder (0.28-6.58 kb RNA Marker, Promega, Cat.# G3191) was electrophoresed on a 1% agarose/formaldehyde gel, and then transferred onto a nylon membrane in 20x SSC. Hybridization was performed at 65°C o/n in 0.5 M sodium phosphate buffer, pH 7.2/7% SDS with the gel-purified PCR product generated with primers mnbDI and mnbR1 from the *MNBH* cDNA. The membrane was washed in 2x SSC/0.1% SDS at 37°C, and then at 50°C in the same solution for 15 min, and exposed for 3 days.

The human adult multiple-tissue Northern blots (Clontech) contained 2 μ g of poly(A)⁺ RNA from the tissues indicated. Northern blots were hybridized with: i) a 4.5 kb cDNA fragment containing the whole *MNBH* coding region plus the 3'-UTR, and ii) oligonucleotides of 50-60 bp specific for transcripts *MNBHa* and *MNBHb*. Hybridization was performed at 65°C for 1.5 h in ExpressHyb hybridization solution (Clontech), according to the protocol provided. The membranes were washed twice in 3x SSC/0.1% SDS at 37°C for 15 min, and then in 2x SSC/0.1% SDS at 50°C for 10 min, 1x SSC/0.1% SDS at 65°C for 10 min, and exposed to an X-ray film with intensifiers for 2-3 days.

Probes and primers

Primers used for the generation of cDNA probes, RT-PCR, and RACE, were: mnb1D: ctcgggtgtattttggttg; mnb1R: cgttttccatctttggtcttc; mnb11D: gtggagcaagaatgggttg; mnb80R: acaccttctcggtttccagagtc; mnb51D: tcatctaaagggcattccgatgga; mnb86R: ggggcgctcgcgccgcttc; mnb16R gagtccagcggcaaaactataac; mnb4R:

mnb4D: caggccaagggtctacaattc; mnb89D: gtccagttttgtgtgctttactacacag; ctgggtcacggaaggtttg; mnb46R: gtttggctggcgacggtcactgtactg; mnb47R: ggggaagggagggagggagggagggagggaggtcagcg; mnb41D: gctgagactcaccggaggaagc; mnb93D: mnb81D: tatttgcccgaccactttttgacc; mnb83D: gagccgaggagagactgagcag; mnb53D: acgcagccagcacagcatac; gcgtgtgcgagggcgattgtg; mnb99R: aacgtcatgaacctcagggta; anchor1: cgcgcggccgcctatcgattctggaaccttc; anchor2: gaaggttccagaatcgatagg. Probes used for both the screening of phage cDNA libraries and the clones resulting from caggatattaataaagagttttgatggttaaggacaagcaggcgaattaa; probe77D: RACE probe78D: were: cgggagcccgaggctgagactcaccggaggaagcggcgcgagcgccccg

RESULTS

Identification of variant MNBH transcripts

Using a combination of cDNA library screening and RACE-PCR, two transcription start sites were identified. The transcripts were named MNBHa and MNBHb, according to their first exon (Fig. 1a). The only alternative splicing of an internal exon within the 5'-UTR of MNBH corresponds to exon 3, and is specific for transcript MNBHa. To confirm the existence of this alternatively spliced exon 3, we performed RT-PCR, using two primers (mnb41D from exon 2 and mnb46R from exon 5) flanking exons 3 and 4 (Fig. 1b). We detected two fragments on an agarose gel, corresponding to the splice variants described above. Both MNBHa variants were present in all tissues tested in both fetus and adult. Although the transcript without exon 3 seemed to be the predominant species, the splice variant containing this exon also accounted for a significant portion of total MNBHa mRNA. Contamination of the RNA preparation with genomic DNA could not be responsible for these amplifications, since introns 2a, 2b, 3 and 4 (42.9, 9.6, 53 and 52.3 kb, respectively) were intervening in the genomic sequence. We were never able to detect a product in any tissue, when using a primer from exon 1 (mnb51D) with a primer from exon 2 (mnb86R). Also, sequencing of at least 30 RT-PCR products isolated using primers mnb51D and mnb46R failed to identify a single clone containing exon 1 along with exon 2 and/or exon 3, yielding two different transcription start points. Exon 3 comprises part of an AluJ element. The genomic sequence adjacent to exon 3 has a $poly(A)_{30}$, as is expected for an Alu element (Schmid and Jelinek, 1982). We screened the human fetal brain and liver libraries using a MNBHa 5'end specific oligonucleotide probe (mnb77D), yielding 4 positive clones, of which only one contained exon 3.

The other alternative splicing events observed in *MNBH*, affect the open reading frame and produce at least five putative protein isoforms. The most abundant transcript in brain produces the largest protein isoform (MNBH-iso1), with a predicted length of 763 amino acids or a protein mass of 85.6 kDa (Fig. 1a). One variant of this is 9 amino acids shorter near the N-terminus in exon 6 and was previously reported by Shindoh *et al.*, (1996). Sequencing of 50 independent clones from RT-PCR experiments showed that this shorter form of the protein is predominant (98%). This splice variant is also conserved in rat (Kentrup *et al.*, 1996). An examination of DNA sequences flanking the skipped segment in exon 6 revealed a splice acceptor consensus sequence, indicating that this variant results from the alternative use of splice acceptor sites (Fig 1c).

The other splicing events result in the truncation of the carboxy-terminus of the protein. The skipping of exon 13a disrupts the MNBH translation unit introducing a -1 frameshift in the reading frame and a TGA stop codon at triplet 530 in exon 14 of MNBH-iso2. Alternative exon 13b was found using the GENSCAN program for the prediction of genes *in silico* on the genomic sequence containing the *MNBH* gene. The BLASTN results obtained upon analysis of this predicted exon revealed that it was part of an EST in the database (gb:AA342261) which contained the 3' UTR of the *MNBH* gene. The sequence of clones from RT-PCR experiments performed with primers specific to this exon (mnb99R) and primers upstream in the *MNBH* coding region (mnb1D and mnb11D) demonstrated that it was indeed part of the gene. When present, a TGA stop codon is introduced at triplet 541 in MNBH-iso3. We also detected the existence of a cDNA containing both exons (13a and 13b, named exon 13c), which also incorporated 47 nucleotides of the intronic sequence between exons 13a and 13b in the reading frame of MNBH-iso4, resulting in 15 extra amino acids and a new TGA stop codon at triplet 585.

Characterization of the 3'UTR of MNBH

Six independently isolated 3'-end cDNA clones were sequenced and were identical. The same 3'-end was also observed in clones resulting from 3' RACE. A BLASTN search with this 3'-UTR found several poly(A)-containing ESTs. A cannonical polyadenylation signal sequence AATAAA (Wickens and Stephenson, 1984) was present 436 bp upstream of the polyadenylation site. Moreover, computational analysis revealed three weak polyadenylation signals 49 bp upstream of the putative poly(A) addition site (Fig. 2). Seven ATTTA mRNA sequence destabilization motifs (Wilson and Treisman, 1988) were present in the 3'-UTR of *MNBH*.

Genomic organization of MNBH

The *MNBH* transcripts are distributed over 17 exons and the gene spans 149.5 kb of genomic DNA. The genomic organization of *MNBH* is shown in Fig. 3. Analysis of the DNA sequences at the intron/exon boundaries showed that they all adhere to the 5'gt/ag3' splice junction consensus rule for donor and acceptor splice sites (Breathnach and Chambon, 1981; Shapiro and Senapathy, 1987). The exons ranged in size from 69 bp (exon 2) to 3,263 bp (exon 14) and the sizes of the introns varied from 509 bp (intron 12a) to 54,465 bp (intron 1) (Table 1). Primers were designed to allow amplification of each exon were designed, and to facilitate their use in SSCA protocols for mutational analysis, the amplification products were designed to be less than 550 bp in length. The primer sequences, product sizes and annealing conditions are shown in Table 2.

For mapping, each of the 17 *MNBH* exons was amplified separately by PCR using the flanking intronic primers and each amplification product was used as a probe to hybridize a gridded filter containing the cosmid contig generated from YAC 336G11 (Guimera *et al.*, 1997). In this way, the 17 exons were mapped to the corresponding cosmids. *MNBH* is located between markers *D21S335* and *D21S337*, and is 186.5 kb centromeric to the reference marker *D21S55*. Intron 4 contains marker *D21S270* (Fig 3).

Transcription initiation sites and putative promoter of MNBH

A single start site was detected for *MNBHb* when the tissue source was whole human fetus. For transcript *MNBHa*, sequence analysis of 19 independent RACE clones indicated that transcription started within a stretch of 32 nucleotides (+1 to +32, Fig. 2, exon 2) with a single nucleotide used predominantly (G^{+9}). The initiation sites were at: G^{+1} , G^{+10} , G^{+18} and C^{+26} , each in two clones, G^{+5} , G^{+12} , C^{+13} , C^{+19} , C^{+21} and C^{+32} in one clone each, and G^{+9} in five clones. Additionally, four clones isolated from the human fetal brain and liver cDNA libraries using mnb77D as a probe, also identified an ambiguous putative transcription initiation site within this stretch of nucleotides at G^{+5} , G^{+9} , G^{+12} and C^{+13} . We designated the nucleotide from which the longest transcript was produced as the start (+1) of the transcript. There is a MED-1 motif (Multiple start site Element Downstream), 116 nucleotides from the transcription start of *MNBHa*. These regulatory sequences are associated with multiple transcription start sites (Ince and Scotto, 1995).

We were never able to obtain a product when primers 80-610 bp upstream of each transcription start (mnb93D and mn81D for *MNBHb* and mnb83D and mnb53D for *MNBHa*, Fig. 1b) were used in combination with primer mnb46R in exon 5 on an RNA template, therefore, we are confident that there is no RNA species containing these sequences 5' of our designated first exons.

Presumably, since *MNBHa* and *MNBHb* arise from alternative transcription start sites there are two promoters. Computer analysis of the region upstream of the transcription start site of transcript *MNBHa* failed to predict any canonical TATA or CAAT boxes but identified a GC-rich element. The overall G + C content of the 2,385 bp of DNA around the transcription initiation site of *MNBHa* was 75.1%, with a ratio of observed/expected CpG of 0.94 and a CpG/GpC ratio of 0.93 (Fig. 4). There were also recognition sites for numerous rare cutting enzymes and for methylation sensitive restriction enzymes within this region. The region immediately upstream of *MNBHb* had neither a perfect TATA-box nor the consensus sequence (YYCAYYYYY, where Y = C or T) for a strong initiator element, which is often found in TATA-box-less genes; however, a non-consensus AT-rich motif (AATAAAAAT), located -19 to -11 bp upstream of the mRNA cap site, and surrounded by a G + C rich sequence, may function in determining the start site. Moreover, a CAAT box (-50 to -47 bp relative to the start site of transcript *MNBHb*) was found upstream. A computer search of the 5'-flanking regions of *MNBHa* and *MNBHb* for cis-acting regulatory elements revealed a large number of them (Fig. 2).

Overexpression of MNBH in DS and in partial trisomy 16 mice

We compared the levels of expression of *MNBH* in DS and non DS brains. In each case RNA was pooled from the brains of two fetuses from interrupted pregnancies. The total amount of RNA in each lane was determined by densitometry of the gel photograph and the intensity of the signal of the autoradiograph was normalised against these values. In DS brains, *MNBH* was overexpressed 1.5 fold (Fig. 5). A similar result (2.1 fold increase) was obtained when comparing the level of expression of *Dyrk* in the brains of the Ts65Dn mice (which have three copies of *Dyrk*) and control littermates.

Differential expression of alternative MNBH transcripts

Human *MNBH* was shown to be widely expressed on Northern blots when probed with a 4.5 kb segment of the cDNA, which included the whole coding region and the 3'-UTR (Fig. 6a). Transcripts *MNBHa* and *MNBHb* are indistinguishible on Northern blots because of their similar sizes (5,182 and 5,287 bp respectively). When using probes specific for each transcript, *MNBHa* was found to be expressed in all the tissues tested (Fig 6b). In contrast, *MNBHb* was only expressed in heart and skeletal muscle, but not in brain or the other tissues (Fig. 6c). By RACE and RT-PCR we never detected *MNBHb* in fetal or adult brain nor in adult cerebellum.

Primary and Secondary Amino Acid Structure of MNBH

The most predominant isoform of MNBH (MNBH-iso1) is 754 amino acids in length and lacks exons 13b and 13c. The other splicing events (MNBH-iso1 with exon 6a, MNBH-iso2, MNBH-iso3 and MNBH-iso4) are relatively rare in brain, but their relative proportions in other tissues is unknown. During sequencing we identified two amino acid polymorphisms in the coding region of MNBH: aa 415 is Tyr (TAC) or Phe (TTC) and aa 681 is Gln (CAG) or His (CAC or CAT). Three other polymorphisms were detected but did not change the coded amino acid: His⁶⁰⁸ (CAT or CAC), His⁶⁰⁹ (CAC or CAT), and His⁶¹³ (CAC or CAT).

The hydrophobicity profile of MNBH-iso1 shows the protein to be extremely hydrophilic at both the Nand C-termini, with the hydrophilic peaks flanking a hydrophobic region, that corresponds to the kinase core domain (subdomains 1-IX) (Fig. 7a). Isoforms MNBH-iso2, iso3 and iso4 appear to be overall less hydrophilic, since they lack the C-terminal region. The kinase subdomains (I-XI) and the bipartite nuclear translocation signals are highly conserved between human, rat and fly as is a putative leucine zipper (Leu²⁷⁴-(X)₇-Leu²⁸¹-(X)₇-Leu²⁸⁸-(X)₇-Leu²⁹⁵) (Fig. 7b). The DNA region encoding the histidine repeat (not present in MNBH-iso2, iso3 and iso4) was assayed by PCR in 40 unrelated individuals and was found to be invariable. The MNBH proteins are predicted to be located in the nucleus by the PSORT program available on http://psort.nib.ac.jp.

DISCUSSION

The *MNBH* gene has two transcription initiation sites with exons 1 and 2 as alternative first exons. The use of these alternative transcription initiation sites yields two major transcripts, *MNBHa* and *MNBHb*. The 5'-flanking region of transcript *MNBHa* possesses the characteristics of a so-called CpG island (Bird, 1986; Gardiner-Garden and Frommer, 1987), i.e. is rich in G+C and has a CpG/GpC ratio greater than 0.6. There is also heterogeneity of transcription initiation, another common feature of TATA-less promoters. Many mammalian housekeeping genes possess this type of promoter and these types of genes are expressed at low levels in essentially all tissues (Dynan, 1986), consistent with the ubiquitous expression pattern of *MNBHa* in human tissues. Proximal to the transcription start site of *MNBHb* there is no perfect TATA-box but further upstream there is a CAAT box motif. The regions upstream of both putative start sites contain potential binding sites for many common transcription factors (Fig. 2). Notably, upstream of *MNBHa* there are 14 potential binding sites for Sp1. Exactly how all these factors converge to coordinate the expression of *MNBH* remains to be determined. In addition, since the 5' region of the *MNBH* gene has a high density of CpG dinucleotides, it is possible that methylation is one mechanism used to regulate *MNBH* expression.

Exon 3 is part of an *AluJ* element. The presence of the *AluJ* element may serve to lengthen the 5'-UTR of *MNBHa* and thereby promote efficient translation, since length seems to overcome the inhibitory effect on translation of secondary structures prevalent in GC-rich leader sequences (Kozak, 1987). An *AluJ* element was reported as an alternatively spliced sequence in the coding region of the *ADARB1* gene (Mittaz *et al.*, 1997) and this alternatively spliced product was conserved in rat and mouse, although obviously the inserted sequence was not an *AluJ* element. It will be of interest to study the 5'-region of *Dyrk* to determine if the longer isoform is also conserved in rat and mouse.

Exon 14 is common to all *MNBH* isoforms and contains several ATTTA motifs that are involved in destabilization of the message (Wilson and Treisman, 1988; Asson-Batres *et al.*, 1994). Such motifs are also found in the 3'-UTRs of mammalian mRNAs that' encode inducible proteins such as GM-CSF, IFN-b, c-myc and c-fos (Wellington *et al.*, 1993). These motifs are also conserved in *Drosophila mnb*. Also, there is a PEST sequence in all the MNBH isoforms which is involved in the rapid degradation of proteins. Therefore, the turnover of *MNBH* may be very rapid, as may be expected for genes involved in cell cycle. The N-terminal region of the protein is shared by all the MNBH isoforms but the carboxy-terminus is different in each case. A similar arrangement is observed in *Drosophila* where the three protein isoforms differ at their carboxy-terminii (Tejedor *et al.*, 1995).

A computer prediction indicated that MNBH is located in the nucleus. The alternative MNBH isoforms reported here all possess a bipartite nuclear translocation signal which appears in about half of nuclear proteins, but in less than 5% of non-nuclear proteins, and a leucine zipper which promotes DNA-protein interactions. These features suggest that MNBH functions in the nucleus. Note that residues His⁵⁹⁹ to His⁶¹⁹ and Ser⁶⁵⁹ to Thr⁶⁷² comprising the histidine repeat and the S/T domains, respectively, are only present in the largest form of the protein (MNBH-iso1), but not in the other isoforms. Histidine repeats have been reported in other protein kinases and transcription factors, but their exact functional significance is unknown. It is interesting that instead of a histidine repeat, *Drosophila* has an alanine repeat of similar length and in the same relative position in the protein (Fig. 7b).

In the coding region of *MNBH* we identified two polymorphisms which change the amino acid and three others that did not. These changes may be useful as allele-specific polymorphisms for use in linkage disequilibrium studies. It will be of interest to determine if either of the amino acid variants, in the homozygous or heterozygous state, results in phenotypic consequences.

We report a size of 5.2 kb for the full-length cDNA of *MNBH*. Although four groups have independently identified sequences of the *MNBH* gene (Guimera *et al.*, 1996; Shindoh *et al.*, 1996; Song *et al.*, 1996; Chen and Antonarakis, 1997), only the protein coding region and part of the 3'-untranslated region (in sum about 2.7 kb) has been reported. A longer *MNBH* sequence was deposited in the public domain (accession number D86550, Ohira *et al.*, 1997), extending the transcript at both extremes. However, from the data presented here we conclude that the 5'- end sequence deposited is genomic and is not represented in the full-length *MNBH* transcript.

MNBHa is ubiquitiously expressed from a CpG-rich promoter region, whereas transcript *MNBHb* is only expressed in heart and skeletal muscle. Although both promoters appear to be used in adult heart and in skeletal muscle, expression from the TATA-like promoter may be controlled in a tissue-specific manner in muscle. Thus, the expression pattern of *MNBHb* implicates the involvement of this transcript in muscle specific functions and suggests a possible role for *MNBH* in the pathophysiology of muscular hypotonia that affects almost all DS patients.

The human chromosome 21 sequencing initiative has proven very useful for determining the genomic structure of the *MNBH* gene. We describe a set of 15 primer pairs that can be used to individually amplify all the exons from genomic DNA. This information will provide an important resource for the discovery of mutations in *MNBH*, if the gene is associated with a particular pathology.

In conclusion, we have identified variants of the *MNBH* mRNA and reported their full nucleotide sequences in addition to the sequences corresponding to their putative promoter regions. *MNBH* is differentially expressed, with the transcript under the control of the TATA-like promoter only observed in skeletal muscle and heart. We have also demonstrated that *MNBH* is overexpressed in the brains of individuals with DS and in the mouse DS model, TS65Dn. This information should help towards understanding the possible involvement of *MNBH* in DS and aid in the development of animal models.

Aknowledgements

JG thanks MJ. Cuevas and J. Calonge for their encouragement and cooperation, M. Dierssen for fetal DS and Ts65Dn brain tissues, A. Aviñó for synthesis of PCR primers, C. Pucharcos for bioinformatics, M. Lynch and J. Fuentes for their constant help and F. Tejedor and M. Pérez for constructive criticisms. This work was supported by the European Union (Grants CEC/BIOMED2 BMH4-CT96-0554/BMH4-CT96-1364), the Dirección General de Enseñanza Superior (PM95-0106-C02-01), the Fundació Catalana Síndrome de Down/Marató TV3, and the Servei Català de la Salut.

REFERENCES

Asson-Batres, MA., Spurgeon, S.L., Diaz, J., DeLoughery, T.G. and Bagby, G.C. Jr. (1994). Evolutionary conservation of the AU-rich 3' untranslated region of messenger RNA. *Proc Natl Acad Sci U S A* 91: 1318-1322.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature 321: 209-213.

Breathnach, R. and Chambon, P. A. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* 50: 349-383.

Chen, H., Chrast, R., Rossier, C., Gos, A., Antonarakis, S.E., Kudoh, J., Yamaki, A., Shindoh, N., Maeda, H., Minoshima, S. and Shimizu, N. (1995). Single-minded and Down syndrome? *Nature Genet.* 10: 9-10.

Chen, H. and Antonarakis, S.E. (1997). Localisation of a human homologue of the Drosophila mnb and rat Dyrk genes to chromosome 21q22.2. *Hum Genet* 99: 262-265.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanatephenol-chloroform extraction. *Anal Biochem* 162: 156-159.

Dahmane, N., Charron, G., Lopes, C., Yaspo, M.L., Maunoury, C., Decorte, L., Sinet, P.M., Bloch, B. and Delabar, J.M. (1995). Down syndrome-critical region contains a gene homologous to Drosophila sim expressed during rat and human central nervous system development. *Proc Natl Acad Sci USA* 92: 9191-9195.

Delabar, J.M., Theophile, D., Rahmani, Z., Chettouh, Z., Blouin, J.L., Prieur, M., Noel, B. and Sinet, P.M. (1993). Molecular mapping of twenty-four features of Down syndrome on chromosome 21. *Eur J Hum Genet* 1: 114-124.

Dynan, W.S. (1986). Promoters for house-keeping genes. Trends Genetics 2: 196-197.

Gardiner-Garden, M. and Frommer, M. (1987). CpG islands in vertebrate genomes. J Mol Biol 196: 261-282.

Guimera, J., Casas, C., Pucharcos, C., Solans, A., Domenech, A., Planas, A.M., Ashley, J., Lovett, M., Estivill, X. and Pritchard, M.A. (1996). A human homologue of Drosophila minibrain (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. *Hum Mol Genet* 5: 1305-1310.

Guimera, J., Pucharcos, C., Domenech, A., Casas, C., Solans, A., Gallardo, T., Ashley, J., Lovett, M., Estivill, X. and Pritchard, M. (1997). Cosmid contig and transcriptional map of three regions of human chromosome 21q22: identification of 37 novel transcripts by direct selection. *Genomics* **45**: 59-67.

Hanks, S.K. and Quinn, A.M. (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* 200: 38-62

Ince, T. A. and Scotto, K.W. (1995). A conserved downstream element defines a new class of RNA polymerasa II promotors. J. Biol. Chem. 270: 30249-30252.

Kentrup, H., Becker, W., Heukelbach, J., Wilmes, A., Schurmann, A., Huppertz, C., Kainulainen, H. and Joost, H.G. (1996). Dyrk, a dual specificity protein kinase with unique structural features whose activity is dependent on tyrosine residues between subdomains VII and VIII. *J Biol Chem Feb* 271: 3488-3495.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15: 8125-8148.

Kulp, D., Haussler, D., Reese, M.G. and Eeckman, F.H., (1996) A generalized hidden Markov model for the recognition of human genes in DNA. *Ismb* 4: 134-142.

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying hydropathic character of a protein. J. Molec. Biol 157; 105-132.

Marchuk, D., Drumm, M., Saulino, A. and Collins. FS. (1991). Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res* 19: 1154.

Matsumoto, N., Ohashi, H., Tsukahara, M., Kim, K.C., Soeda, E. and Niikawa, N. (1997). Possible narrowed assignment of the loci of monosomy 21-associated microcephaly and intrauterine growth retardation to a 1.2-Mb segment at 21q22.2. *Am J Hum Genet* 60: 997-999.

Mittaz, L., Scott, H.S., Rossier, C., Seeburg, P.H., Higuchi, M. and Antonarakis, S.E. (1997). Cloning of a human RNA editing deaminase (ADARB1) of glutamate receptors that maps to chromosome 21q22.3. *Genomics* **41**: 210-217.

Ohira, M., Seki, N., Nagase, T., Suzuki, E., Nomura, N., Ohara, O., Hattori, M., Sakaki, Y., Eki, T., Murakami, Y., Saito, T., Ichikawa, H. and Ohki, M. (1997). Gene identification in the 1.6-Mb region of the Down syndrome region on chromosome 21. *Genome Res* 7: 47-58.

Rahmani, Z., Blouin, J.L., Creau-Goldberg, N., Watkins, P.C., Mattei, J.F., Poissonnier, M., Prieur, M., Chettouh, Z., Nicole, A., Aurias, A. *et al.* (1989). Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome. *Proc Natl Acad Sci U S A* 86: 5958-5962.

Shapiro, M.B. and Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 15: 7155-7174.

Shindoh, N., Kudoh, J., Maeda, H., Yamaki, A., Minoshima, S., Shimizu, Y. and Shimizu, N. (1996). Cloning of a human homolog of the Drosophila minibrain/rat Dyrk gene from "the Down syndrome critical region" of chromosome 21. *Biochem Biophys Res Commun* 225: 92-99.

Smith, D.J., Stevens, M.E., Sudanagunta, S.P., Bronson, R.T., Makhinson, M., Watabe, A.M., O'Dell, T.J., Fung, J., Weier, H.U., Cheng, J.F., Rubin, E.M. (1997). Functional screening of 2 Mb of human chromosome 21q22.2 in transgenic mice implicates minibrain in learning defects associated with Down syndrome. *Nat Genet* 16: 28-36.

Song, W.J., Sternberg, L.R., Kasten-Sportes, C., Keuren, M.L., Chung, S.H., Slack, A.C., Miller, D.E., Glover, T.W., Chiang, P.W., Lou, L. and Kurnit, D.M. (1996). Isolation of human and murine homologues of the Drosophila minibrain gene: human homologue maps to 21q22.2 in the Down syndrome "critical region". *Genomics* 38: 331-339.

Tejedor, F., Zhu, X.R., Kaltenbach, E., Ackermann, A., Baumann, A., Canal, I., Heisenberg, M., Fischbach, K,F. and Pongs, O. (1995). *minibrain*: a new protein kinase family involved in postembryonic neurogenesis in *Drosophila*. *Neuron* 14: 287-301.

Wellington, C.L., Greenberg, M.E. and Belasco, J.G. (1993). The destabilizing elements in the coding region of cfos mRNA are recognized as RNA. *Mol Cell Biol* 13: 5034-5042.

Wickens, M. and Stephenson P. (1984). Role of the conserved AAUAAA sequence: Four AAUAAA point mutants prevent messenger RNA 3'end formation. *Science* 226: 1045-1051.

Wilson, T. and Treisman, R. (1988). Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3' AU-rich sequences. *Nature* 336: 396-399.

Legends

Fig. 1

Alternative splicing of *MNBH*. a) Schematic representation of the differentially spliced *MNBH* transcripts. The exons and introns are numbered and drawn to scale. Open boxes are noncoding regions, black boxes represent the translated region. The asterisk indicates the alternative spliced exon 6. b) Primers used in the 5'-RACE and RT-PCR experiments. c) Splicing variant in exon 6. The invariant 3' splice AG is shown in bold and is underlined.

Fig. 2

Composite nucleotide and amino acid sequences of the human *MNBH* cDNAs. The 5' regions flanking *MNBHa* and *MNBHb* are shown in lowercase. Putative transcription regulatory elements are underlined, including the consensus binding sites for Sp1 and AP-1 (bold). Transcribed sequences are in capitals and the transcription initiation sites are indicated by +1. The nucleotides in large bold uppercase represent the first nucleotide of each exon. The amino acids are numbered on the right. The window of multiple initiation sites (+1 to +32) in exon 2 of *MNBHa* is double underlined. Both the single and bipartite nuclear translocation signals are boxed in exon 7. The leucines (L²⁷⁴, L²⁸¹, L²⁸⁸ and L²⁹⁵) of the zinc-finger are in linked boxes. The PEST region is boxed within exons 12 and 13a. The intronic sequence between exons 13a and 13b that becomes part of exon 13c is indicated with a bar above the nucleotide sequence. The discontinuous histidine repeat is in bold and double underlined. The S/T region is underlined. The catalytic region of the kinase domain is between brackets. A polyadenylation consensus signal AATAAA is double underlined and in bold, 436 bp upstream of the poly(A) tail. Other atypical polyadenylation signals are underlined. The seven ATTTA motifs involved in the desestabilization of RNA are in bold.

Fig. 3

Genomic structure of *MNBH*. The minimal tiling path of cosmids (thin bars) derived from the chromosome 21 YAC 336G11 (thick bar), and the restriction map is shown. *MNBH* exons are indicated by vertical lines on the cosmid clones. Circles on the bars represent the markers.

Fig. 4

G + C content around exons 1 and 2 of *MNBH*. The G + C content and the observed /expected CpG ratio are shown for the nucleotide sequence between the 5' proximal region of exon 1 and part of intron 2a/2b. The vertical axis indicates the G + C content (0-100%, line) and the CpG ratio (0.0-1.2 black area). The numbers below the horizontal axis indicate the nucleotide sequence length. White horizontal bars indicate the areas that satisfy the criteria for CpG islands (>50% G+C and >0.6 CpG ratio). Underneath the graph, the position of each CpG and GpC dinucleotide relative to a position in the sequence is indicated by a vertical line. The analysis was carried out using a program developed by Dr. M. Hirata at NIH of Japan. Boxes indicate the relative positions of exons 1 and 2.

Fig. 5

Overexpression of the minibrain transcripts in Down syndrome and in partial trisomy 16 mice (Ts65Dn). Lane 1 is Ts65Dn, lane 2 is a control litter mate of Ts65Dn (two copies of MMU16), lane 3 is DS, lane 4 is non DS.

Fig. 6

Tissue distribution of *MNBH* transcripts. a) The *MNBH* mRNA was detected as a single band in all tissues examined when a 4.5 kb cDNA (common in all the *MNBH* transcripts) was used as a probe. Longer exposures did not reveal other bands. To determine the tissue distribution of each transcript, the same Northern filters, were rehybridized with a probe specific for each transcript. b) Exon 2 (*MNBHa*) and c) Exon 1 (*MNBHb*).

Fig. 7

The mnb proteins. Hydrophilicity plots of the MNBH-iso1 protein, as calculated by the method of Kyte-Doolittle (1982). a) The only hydrophobic region of MNBH is confined to the catalytic domain, between subdomains I and IX. b) Schematic comparison of the structure of human (MNBH), rat (Dyrk) and *Drosophila* (mnb) minibrain proteins. Conserved motifs are labelled, with the black boxes and Roman numerals designating the highly conserved kinase subdomains according to the nomenclature of Hanks and Quinn (1991). The vertically striped box denotes the spliced domain conserved in human and in rat.

Table 1

The invariant ag/gt are in bold. Exon sequences are in uppercase, intron sequences are in lowercase. *MNBH* exon 6 is alternatively spliced (see text and Fig. 1c). 6a indicates the rare larger form of this exon, 6b indicates the shorter form of exon 6.


4

gaccttaatcacactagcagggcaattttgaaacaataatccgaaagaacttttgccagctacggcaacttttttatatcatggagatcagc -621
SOX-5. SKY UQXA
ctgaaagcaagctatttgccc <u>gaccacttttt</u> gacc <u>aatttatt</u> ttcggtttagtcaggtttcttcccaagtctgtgctaa <u>ttgatggtta</u> -529 RORalp CdxA
<u>ccttagtaattactttttgactaatttattttcgg</u> tttagtcggatttct <u>ttcaaagtctgtgctaatctg</u> aatggttaccttagta <u>tataa</u> -437 Oct-1 S8 CdxA C/EBP NF-Y CdxA <u>atgtcctagtaactgtagcaaggacatttttgctggaaatc<u>atttaaag</u>caggtgtat<u>ttgagtcat</u>gtagcatagacagcggcatttactg -345</u>
CdxA AP-1 cgagttttttacatc <u>taatattgttta</u> taagattttatctgttgtca <u>catttgtg</u> aa <u>atttatg</u> aaaaaaattggttacgaa <u>cactatcaca</u> -253
aggtggcaaaatcaagctt <u>agaaatcaattagtga</u> ctctgctgtaaaata <u>gtatata</u> tatacacacacactctttttaggtatac <u>attatt</u> -161
S8 AP-1 CdxA CdxA aaatgccattagagttggtacaaaatacaatttttagctggtggctctctaccgattacaacgcagccagc
EXON 1 gcctcgggagctcacccagcaatacatagctagggagaaggggtaggggagaataaaaatcagcgaaagcCAGGAT <u>ATTAATA</u> AAGAGTTTT CAAT-like box TATA-like box +1 CdxA, Pbx-1
GATGGTTAAGGACAAGCAGGCGAATTAAGACTCTGGAAACCGAGAAGGTGTATAGAAAAGAAAACCAGAGAGGCTGGACGACAAGGACAACAGT
AACAGCACAAGGGGATTCATCTAAAGGGCATTCCGATGGAGCAGGCAG
<pre>gcgcggcggtcgcggccgccgccgccgcgggggggggg</pre>
ggtgggettecceggg <u>teccegeceg</u> aaggeegegegegeeee <u>teeteete</u> eeeeteeeegee <u>ggggeageeege</u> teeeeeggegggg1246
MZF1, Sp1 MZF1
sp1 sp1, Egr-1, 2 and 3 ggcggggggggggggggggggggggggggggggggg
sp1, sp1 GATAI and GATA2 MZF1, sp1 taccctccctcggcctagccg <u>tccccctcctctccggcgaatcacgcgccctcctctgacactcgtagcggacccgagctccgaatt</u> -970
tatccttcacgtgacctgagggggttgggtggttgggtggtgggtg
acaggcccgccggctgggggggggggggggggggggggg
gcgaccgggtcggtcgtcgccatttgttggttggtttttttt
gtgaggcgggggtggcggcggcgacccgcggcgggggtatccgggggagactgctgctgctgctgctgctgctgctggcccaggtcggcctcag -602 Sp1 GATA1 and GATA2 GATA1 and GATA2 510
agageggacaccccgageggggggggggggggggggegegege
GATA1, GATA2 and GATA3
gugggeetgeetgeetgeetgeetgesggggeetggggeetggggeetggggegggggggg
YY1 ccccccccccccccccccccccccccccccccccc
sp1 gggggctgtcgcctccccggccccgggcgccgctggaaccgcgagccgaggagagactgagcaggctgcggcgggccaggagccggagcgc -92
sp1 gggggctgtcgcc <u>tccccggc</u> cccggggcgcgcgcgggaggaggaggaggaggaggaggag
Sp1 gggggctgtcgcc <u>tccccggc</u> cccggggcgccgctggaaccgcgagccgaggagagagaga
sp1 gggggctgtcgcc <u>tccccggc</u> cccggggcgccgctggaaccgcgggggggggg
sp1 gggggctgtcgcc <u>tccccggc</u> cccgggcgccgcgggaggggggggggggggg
Sp1 ggggggctgtcgccgggccgggccgggccggggcgggg
Sp1 ggggggctgtcgccccgggcccccgggcgcccgcgcgcg
Sp1 ggggggtgtcgtcgcccgggcgcccggggggggggggg
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c c c c c c c c c c c c c c c c c c c $
$\begin{array}{c} \text{sp1} \\ \hline \\ $
$\begin{array}{c} \text{sp1} \\ & \text{MEF1} \\ & \text{MEF1} \\ & \text{mass} \\ & $
$\begin{array}{c} \text{sp1}\\ gggggctgtcgccgtggccgtggcgcgcgcgcgcgcggggcgcggggcgcggggcgcgggg$
$ \begin{array}{c} \begin{array}{c} sp1 \\ gggggctgtcgcggcggcgcggcgggcgcggcgggcgcggcg$

.

Exon 12 ACC AAA GAT GGA AAA CGG ${\bf G}_{\rm AG}$ tac ama cca cca gga acc cgt ama ctt cat amc att CTT GGA GTG GAA т K D G K R E Y K P P G т R K L H N I L G V 430 E ACA GGA GGA CCT GGT GGG CGA CGT GCT GGG GAG TCA GGT CAT ACG GTC GCT GAC TAC TTG AAG TTC AAA т P G R A G E G Н т V D Y G G G R S A T. K K 453 TAT GAC CCC AAA ACT CGA Y D P K T R CTC ATT TTA AGG ATG CTT GAC GAT ATT CAA CCT TAT TAT GCT CTG CAG CAC I D D 0 Μ 0 H AGT TTC TTC AAG AAA ACA GCT GAT GAA GGT ACA AAT ACA AGT AAT AGT GTA TCT ACA AGC CCC GCC ATG K Т A D E G Т N Т S N S V S T S P A M 499 S K Exon 13a GAG CAG TCT CAG TCT TCG GGC ACC ACC TCC AGT ACA TCG TCA AGC TCA GGT GGC TCA TCG GGG ACA AGC E Q S Q S S G т т S S т S S S S G G S S G т s 522 G G S s G Т s 522 CAC CAG CAT CGG CAC AGT AAC AGT GGG AGA GCC CGG TCG GAT CCG ACG GGT GGG CAC TTC ACA GCT GCC H N S G R A R S D P T H Q Η R H S G G T A A 545 N S G R A R S D P T Η Q Н R H S G G H F T A A 545 GTG GTG CAG GCC ATG GAC TGC GAG ACA CAC AGT CCC CAG GTG AGC CAC GTT CAT TTG CTT TCA v D C H F М E т S Q 557 Q A v P S s Н V v н L v v Н S 0 L S 568 C 0 A M D E т 13b Exon GTG GAG CAG CAC TGG ATG CCA GGT GCC TTT AGA ATG ACC GTA TCA TTT ACC CTA CCT GCC ATT CTC AG V V E Q Н W Μ P G A F R М т S F Т L 533 R W S S т G C Q V Ρ L E 584 P Ι L A Exon 14 gga aat \mathbf{G} tg cgt cag caa ttt cct gct cct ggt tgg tca ggc CAT GAC GTT CCT GTC TAA GAG GTT V R Q Q F Ρ А Ρ L G W S 570 G G A S A I S C S S W T V R E V н D V P V 540 CAG GTC ACT GTT GAA ACT CAT CCT GTT CAA GAA ACA ACC TTT CAT GTA GCC CCT ACT GAA GCT CCT ACA T V T V T V E 593 E Q E H 0 H V A Н H H H G N S S H H H H H н H H H H 616 CAC CAT GGA CAA CAA GCC TTG GGT AAC CGG ACC AGG CCA AGG GTC TAC AAT TCT CCA ACG AAT AGC G N R H H H G 0 0 A L TCC TCT ACC CAA GAT TCT ATG GAG GTT GGC CAC AGT CAC CAC TCC ATG ACA TCC CTG TCT TCC TCA ACG Н H 662 S т Μ ACT TCT TCC TCG ACA TCT TCC TCC TCT ACT GGT AAC CAA GGC AAT CAG GCC TAC CAG AAT CGC CCA GTG V N 685 T S т G N Q G N Q A Q R P S S Т S S TCC AAT CCC GCT GCT AAT ACC TTG GAC TTT GGA CAG AAT GGA GCT ATG GAC GTT AAT TTG ACC GTC TAC т Q N G М D V N V Y S P 708 A N D G A L N CAA GAG ACT GGC ATA GCT GGA CAT CCA ACA TAC CAA TTT TCT CGC GCT AAT ACA GGT CCT GCA CAT TAC 0 731 G H R 0 E G Ι A ATG ACT GAA GGA CAT CTG ACA ATG AGG CAA GGG GCT GAT AGA GAA GAG TCC CCC ATG ACA GGA GTT TGT т т М R Q G A D R Е E S P M т G 754 M E G H L GTG CAA CAG AGT CCT GTA GCT AGC TCG TGA CTACATTGAAACTTGAGTTTGTTTCTTGTGTGTTTTTATAGAAGTGGTGTT 763 GCTCCCCCATTTAACTTGCCAACGCCACGCCCAGTCCACAGTGGGGGTTTTTTTGCCCTTCATCACAACCCACCTTAACAGCTCAGAGGTATCCCACGTTGGTCCTTGTGCCACAGCCACAGTGGTGGTCCTTCT CACGGAAATGCAGTACTGTAAGGAAGAGGGACCTCCACQTTCCACAAACACCATCTTCAGCTGTATGAAAGGGACGGTTGTGGGTGAAGTTT GTCAGGCACGGTAAGCATGCTGAGTGGCGGGGGATCACAACTCTCCCTATCTGAACCTACTGAGGAGCAAAGCAGCAATTACATGGGATCCTG TGGCTCTCCCGTTGCAGAGGCCACAGGAAGATACGATGGAACGTGACTGGTCTCCTAACCAAGGTGCACTGATAAGCAATCAACGGGTCGG TCGTGGCCAGTCCTGGGGAGGTCTGATGGTGGTCTTTGGGATAACCTTTGGCCTTATGGATTTGGACTCGAAATTAGAACAGCCTACCATT TCCACCATTTCTTCTGCACAAAGATGTCTTCTGTTCATCCTGAACATTTTTAAAAAATGCAGAATTTTATGTGACTGCTTTTTTTGCCCTCAC AATTATGCTGTGAATTTTACAAAA**ATTTA**TTTTTCTTTTTTGAAA**ATTTA**TTGTACCAAAGCTGTTTTTATAGCACATAGATGTCTGTAACC AATAATGTAGCAGTTCTGCACTTTGACACAAGGTGTAACTAGACCATTTTTAAATGTCAGTTGAAAATTATGGCTGTACTATTGCTTAAAC AAAACTGGAACTGTTGTTGAATCCATAGCCAATAC**ATTTA**CAGCAATCTGTGTACTGAACATAGTAGATTGACATCTAATTCAAGATTACA ACATCTGTTACATTCTAAGTGTGTTCAGGCTTCTGAAGGTAAAGGGACACTGGAACCCAGAAGCTATGGAACCAGCAGTAGATTCTTGTATT GTGTTCCCATTTTAAATTGACCAATTTGGGGTGTGACACTTTTGAGCGGTTGAATTGGGAGAATGAAGATAAGTA**ATTTA**CCTGTCCAGG ATCAAAAGAAGCCTAGAAAAGAAGCAGTAATCTACCTCTGCCGATAACCTGTTTAAGATGACTCAGCAGAACACCGCGTTTCATTCTATTG GTCATGCGGTCTTATGTATGATAAACAGTTGAATAATTTGTC(A),

CAG AGG ATA TAC CAG TAT ATT CAG AGT CGC TTT TAT CGG TCT CCA GAG GTG CTA CTG GGA ATG CCT TAT

GAC CTT GCC ATT GAT ATG TGG TCC CTC GGG TGT ATT TTG GTT GAA ATG CAC ACT GGA GAA CCT CTG TTC D L A I D M W S L G C I L V E M H T G E P L F

E к L

GGT GCC AAT GAG ${f G}$ ta gat cag atg aat aaa ata gtg gaa gtt ctg ggt att cca cct gct cat att

AAG TTC TTT GAG AAG TTG CCA GAT GGC ACT TGG

D G

Exon 10

CAA

GAC

D 0 A Q

GCA CCA AAA

Q S R F Y R S Ρ Е V L L G М

GCA AGA

K

Exon 11

K

Q

AGT S G А N Е v D Q М Ν К I V Е V L G

CTT

Resultats 177

338

361

407

P

W

AAC TTA AAG AAG

N

L K K

1

1

.







Kidney Liver SunJ Placenta Brain Heart



Skeletal muscle



Pancreas

MNBHa



MNBHb





- 7.5

4.4

- 2.4

kb







Subthalamic nucleus Substantia nigra Corpus callosum Caudate nucleus Amygdala





[T]
(and and
-
-
4

Splice Junction Sites at the Exon / Intron Boundaries of the Human *Minibrain* Gene (MNBH)

	Fvon				T t	
Exon	size	3' SPLICE ACCEPTOR	5' SPLICE DONOR	Intron	size	Tvpe
No.	(dq)			No.	(dq)	
1	177		GCGCCCCCGAAAG gtgggggggggggg	1	54,465	(noncoding)
2	69		GCCATCGTCCCG gtgagtgtccgg	2a/2b	42,935/52,664	(noncoding)
3	101	tttgtcacccag GCTGAAGTGCAG	CTGGGATTACAG gtaagtgccacc	e S	9,628	(noncoding)
4	86	ttcttcacacag TGTTATAGTTT	CG ATG CAT ACA G gtgactcaagtt M H T	4	52,299	(M ¹ -T ³)
ъ	197	ttcttcttgtag GA GGA GAG ACT T G G E T	TCT AAC AAC CAG gtaagttcatgg L T N Q	2	5,300	(G ⁴ -Q ⁶⁹)
ба	120	tgtctattta ag GTG ATG CCT GAT * • • V M P D	CAT ATT AAT ĜAG gt aagacttgat H I N E	6a	2,337	(V ⁷⁰ -E ¹⁰⁹)
6b	63	gtcatgttac ag AGG CGG ATG CCC R R M P	CAT ATT AAT GAG g taagacttgat H I N E	69	2,364	(R ⁷⁸ -E ¹⁰⁹)
7	189	ttctcatttcag GTT TAC TAT GCA V Y Y A	TCC TTT GGA CAG gt aatttaatgg S F G Q	2	5,640	(V ¹¹⁰ -Q ¹⁷²)
8	148	cttctttttag GTT GTA AAG GCA V V K A	AA TAC TAC ATA G gtaaacaacag K Y Y I	œ	3,560	(V ¹⁷³ -I ²²¹)
6	287	tttctctttc ag TG CAT TTG AAA C V H L K	TTG GGG CAG AGG gt aagtattatt L G Q R	6	2,555	(V ²²² -R ³¹⁷)
10	147	aatatattcag ATA TAC CAG TAT I Y Q Y	GGT GCC AAT GAG gt aaatgatgta G A N E	10	2,954	(I ³¹⁸ -E ³⁶⁶)
11	141	tttttaatac ag GTA GAT CAG ATG V D Q M	GAT GGA AAA CGG gt aaaataagga D G K R	11	9,025	(V ³⁶⁷ -R ⁴¹³)
12	307	tctctttacag GAG TAC AAA CCA E Y K P	CG TCA AGC TCA G gtctgtgctgct S S S S	12a/12b/12c	509/6,321/681	(E ^{414-S515})
13a	125	gtgatatttc ag GT GGC TCA TCG G G G S S	CAC AGT CCC CAG gt gagctcgcac H S P Q	13a	5,687	(G ⁵¹⁶ -Q ⁵⁵⁷) in MNBH-iso1 and MNBH-iso4
13b	84	tgccattctc ag GTG GAG CAG CAC V E Q H W S S T	GTC TAA GGA AAT gt aagtatttat V STOP	13b	5,557	(V ⁵⁰⁸ -V ⁵³²) in MNBH-iso3
13c	256	gtgatatttcag GT GGC TCA TCG G G G S S	GTC TAA GGA AAT gt aagtatttat 	13b	5,557	(G ⁵¹⁶ -E ⁵⁸⁴) in MNBH-iso4
14	3,263	ctgttctttc ag GTG CGT CAG CAA V R Q Q G A S A				(V ⁵⁵⁸ -S ⁷⁶³) in MNBH-iso1 (G ⁵¹⁶ -H ⁶⁰⁸) in MNBH-iso2
Consensi	SI	(Py) c ₇₈ a ₁₀₀ g ₁₀₀ /G ₅₅	$(C_{38}/A_{35}) A_{62} G_{77}/g_{100}t_{100}(a_{60}/g_{100})$	38) a ₇₄ g ₈₄ t ₅₀		

Resultats 183

> splice signal Consensus

TABLE 2

,

!

PCR Primers Pairs for MNBH Mutation Analysis

-		-	-			-		-	-	T	-	r		1	1		
Exon	amplified	1	2	ε	4	s	6a,6b	1	∞	6	10	п	12	13a, 13b, 13c	14	14 ^a	14 ^a
Ш	(C)	58	58	58	58	58	58	58	58	52	52	52	52	52	52	52	52
Size	(dq)	300	910	338	362	469	358	433	470	542	540	458	499	520	3,603	450	426
REVERSE PRIMER NAME	sequence (5' to 3')	ATTAGGCTGGAGCCATCATG	GGTGGCTCCGAAGCGAGAGTG	CACTGTGACCTCCCAGAGTACTG	AATAACACCACTATAAATCACTG	CCATAATCCCACGTTGCATG	AGTITAGGGTAAATAGGTCAC	TATGGTCATCCAAATACTTGC	GGCACATTCCCCAAGATTTAG	CCTACACTGTCCTACCTACTTC	AAATAGGTATATTCAAACTGACC	CTGGTCCATACCTGATAGTTC	TTCCCAAGCCTAATCTAACC	CTGCAAGAATGTGCTTCTCTATG	CTGATGCCTATGTCAATGTATC	TGGCCAACCTCCATAGAATC	ACTCAAGTTTCAATGTAGTC
		JG92R	JG54R	JG85R	JG56R	JG58R	JG60R	JG62R	JG64R	JG66R	JG68R	JG70R	JG72R	JG74R	JG76R	JG95R	JG9R
FORWARD PRIMER NAME	sequence (5' to 3')	ACGCAGCCAGCACAGCATAC	GAGCCGAGGAGAGAGACTGAGCAG	GGGTTAAGATGGTGTCTTACAGG	TTATGTTAGATATTCCTCAGTTG	GTGJTGAGTAACATATACCTG	GTGGTTGTCTTTATTGGAATC	TGTCAACTGTAGAAAATAGGTG	ATGGCATCTCTTCTACTTAGG	TCAGCAGAAGAGCAACTTAGA	CTAGGTAGGATTAGACTAAATG	CTGTATGCTGGATGTCTAGAG	TGGGATAGTAATGTGTGAAAAG	CCTACCITAACCAGACITC	TTGAATGTATTTGGGGATTTTGTG	TIGAATGTATTTIGGGATTTTIGTG	CAGGCCAAGGGTCTACAATTC
		JG93F	JG83F	JG84F	JG55F	JG57F	JG59F	JG61F	JG63F	JG65F	JG67F	JG69F	JG71F	JG73F	JG75F	JG75F	JG4D

^a Exon 14 is too large for SSCA analysis. Thus, primers were designed to amplify shorter fragments covering the coding region.

4.2. TREBALLS COL·LABORATIUS

En aquesta secció s'inclouen set treballs col·laboratius dins la línia de recerca del cromosoma 21, en el qual vaig participar en l'inici de la meva estada al Centre de genètica mèdica i molecular. Aquests treballs són part de la tesi doctoral de la Dra. A. Bosch, (1994c). La meva participació a cada treball queda detallada.

4.2.1. Two dinucleotide repeat polymorphisms at 21q22.3 (D21S416 and D21S1235)

Bosch A, Wiemann S, Guimera J, Ansorge W, Patterson D, Estivill X

Hum Mol Genet 2:1744 (1993)

El doctorand és l'autor de la determinació de l'heterozigositat, a les famílies de referència CEPH, dels microsatèl·lits D21S416 i D21S1235.

Two dinucleotide repeat polymorphisms at 21q22.3 (D21S416 and D21S1235)

Assumpció Bosch, Stefan Wiemann¹, Jordi Guimerà, Wilhelm Ansorge¹, David Patterson² and Xavier Estivill^{*} Molecular Genetics Department, Cancer Research Institute (IRO), Hospital Duran i Reynals, 08907 L'Hospitalet de Llobregat, Barcelona, Catalunya, Spain, ¹EMBL, Meyerholstraße 1, D-6900 Heidelberg, Germany and ²Eleanor Roosevelt Institute, Denver, CO 80206, USA

Source/Description: To isolate human chromosome 21 (HC21) clones containing CA-repeat polymorphisms, an *Eco*RI HC21 phage library (LA21NS01) was screened using a (GT)₁₀ oligonucleotide (Bosch *et al.*, 1993). Positive clones were subcloned into Bluescribe and regions flanking the repeat were sequenced by walking primer, using a prototype DNA sequencer. Primers flanking the CA-repeats were designed for two clones: ABM-C19 (D21S416) and ABM-C75 (D21S1235) with a repeat motif of (CA)₁₄ (TA)₅ and (CA)₁₉ respectively (EMBL accession numbers Z22787 and Z22801).

Primer Sequences and PCR Conditions: Asymmetric PCR was performed where one of the two primers for each dinucleotide repeat was end labelled with $\gamma^{-32}P$ ATP and its concentration was limiting at 1.5 mM to 10 mM of the unlabelled primer, 1.8 mM MgCl₂, 60 mM KCl, 200 mM of each dNTP (final concentrations) and 100 ng of DNA in 25 μ l. Initial denaturation was at 90°C for 5 min. Amplification was for 28 cycles of 95°C for 20 s, annealing for 20 s at 56°C for ABM-C19 and at 60°C for ABM-C75, extension at 74°C for 20 s, followed by 5 min at 74°C. PCR products were analyzed on 6% ureapolyacrylamide sequencing gels.

 PCR produ 	ct
e length (bp)	Primer sequences
I-C19 99-129	ABM-C19D1
	GAAGTGTTATCAAGTGTTCTGA
	ABM-C19R1
	GGCTAATTAAGCACCATGATTT
I-C75 112-144	ABM-C75D1
	CTTTCATGTGTGTCTACGGAT
	ABM-C75R1
	GGCTACTCTCTGCCCAGAT
	PCR produ length (bp) -C19 99-129

Frequency: Allele frequencies were estimated from the 80 unrelated CEPH parents (160 chromosomes). ABM-C19: Observed heterozygosity of 0.79 (Fig. 1).

Allele	bp	Frequency	Allele	bp	Frequency
Al	129	0.01	A5	121	0.18
A2	127	0.04	A6	119	0.17
A3	125	0.34	A7	117	0.04
A4	123	0.14	A8	99	0.01
ABM-C	275: Obse	rved heteroz	vgosity	of 0.80 (Fig.	2).
Allele	bp	Frequency	Allele	bp	Frequency
AI	144	0.04	A7	122	0.14
AZ	120	0.01		100	
10	130	0.01	A8	120	0.01
A3	130	0.33	A8 A9	116	0.01 0.06
A3 A4	130 128	0.33	A8 A9 A10	116 112	0.01 0.06 0.19
A3 A4 A5	130 128 126	0.01 0.33 0.06 0.04	A8 A9 A10 A11	120 116 112 104	0.01 0.06 0.19 0.01

Chromosomal Localization: We confirmed the HC21 origin of these clones by amplifying the HC21 specific cell line WA17. Subchromosomal localization was determined using a HC21 somatic cell hybrid panel (Gardiner *et al.*, 1988; 1990). We assigned both clones to the 21q22.3 band, beeing ABM-C19 distal from ABM-C75. Mendelian Inheritance: Codominant segregation were observed for ABM-C19 and ABM-C75 in the three generation CEPH families.

Acknowledgements: Fondo de Investigaciones Sanitarias de la Seguridad Social (92/0885E), Comisión Interministerial Científica y Técnica (SAL91-0795-CE), Institut Català de la Salut, European Community Grant CEC PL910014, Eurogem Project, and by NICHHD (HD17449) and NCHGR (HG00716). A.B. was supported by a grant of the Ministerio de Educación y Ciencia, (Spain).

References: 1) Bosch et al. (1993) Genomics, in press. 2) Gardiner et al. (1988) Som. Cell and Molec. Genet. 4, 623–638. 3) Gardiner et al. (1990) EMBO J. 9, 25–34.



Figure 1. Autoradiograph showing the segregation of ABM-C19 (D21S416) through the CEPH pedigree 1322.



Figure 2. Segregation of ABM-C75 (D21S1235) in the CEPH family 1344.

* To whom correspondence should be addressed

4.2.2. Identification of two highly polymorphic CA-repeats (D21S1224 and D21S1261) on human chromosome 21q22.3

Bosch A, Guimera J, Wiemann S, Ansorge W, Patterson D and Estivill X

Human Genetics, 95: 367-369 (1995)

El doctorand és responsable de la determinació de l'heterozigositat, a les famílies de referència CEPH, dels microsatèl·lits D21S1224 i D21S1261.

Hum Genet (1995) 95:367-369

DNAIVARIANTS

Assumpció Bosch · Jordi Guimerà · Stefan Wiemann Wilhelm Ansorge · David Patterson · Xavier Estivill

Identification of two highly polymorphic CA-repeats (D21S1224 and D21S1261) on human chromosome 21q22.3

Received: 30 August 1994 / Revised: 20 September 1994

Abstract Two highly polymorphic CA-repeat microsatellites (D21S1224 and D21S1261) are reported. The clones containing these CA-repeats (ABM-C60 and ABM-C29) have been isolated from a human chromosome-21-specific library (LA21NS01) and have been localised to the q22.3 band using a specific chromosome 21 somatic cell hybrid panel. Both polymorphisms showed heterozygosities of 0.83 in the unrelated reference parents from the Centre d'Etude du Polymorphisme Humain. These new markers should improve the map of the 21q22.3 region, which is believed to contain a large number of genes.

Sequence information. Clones ABM-C60 and ABM-C29, containing CA-repeat sequences, were isolated from a human chromosome 21 (HC21) phage library (LA21NS01) using a (GT)₁₀ oligonucleotide as the probe. They were characterised by subcloning in pBluescript SK (Stratagene) and sequenced by primer walking, using a prototype DNA sequencer, or by directly sequencing the polymerase chain reaction (PCR) products obtained using (CA)_n/(GT)_n and vector primers (Bosch et al. 1993). Primers flanking the CA-repeats were designed for ABM-C60 and ABM-C29, which belong to different loci (D21S1224 and D21S1261) and which have repeat motifs of (CA)₂₂(AT)₅ and (CA)₂₆, respectively (EMBL accession numbers Z34973 and Z34970).

A. Bosch · J. Guimerà · X. Estivill (⊠)

Molecular Genetics Department, Cancer Research Institute (IRO), Hospital Duran i Reynals, E-08907 L'Hospitalet de Llobregat, Barcelona, Catalunya, Spain

S. Wiemann · W. Ansorge EMBL, Meyerhofstraße 1, D-69117 Heidelberg, Germany

D. Patterson Eleanor Roosevelt Institute, Denver, CO 80206, USA
 Table 1
 Primer sequences and allele lengths for the reference

 CEPH individual 134702 are detailed for the two highly polymorphic markers D21S1224 and D21S1261.

Locus	Clone name	Refer- ence allele lengths (bp)	Primer sequences
D21S1224	ABM-C60	148–148	ABM-C60D1 (CA strand) GAAGGAGCTATACAAGGA- CTC ABM-C60R1 (GT strand) GCTAGTAGCTACCATATT- GGA
D21S1261	ABM-C29	191–207	ABM-C29D1 (GT strand) ACACTAGAATCTGACC- CTTGA ABM-C29R1 (CA strand) CCTGGGCAACATGGCGAAA

Primer sequences and PCR conditions. Asymmetric PCR was performed with one of the two primers for each dinucleotide repeat, end-labelled with γ^{-32} [ATP] and present at a limiting concentration of 0.06 µM, whereas the unlabelled primer was at 0.4 µM. Table 1 shows primer sequences and PCR product lengths for each marker. PCR conditions were 1.8 mM MgCl₂, 60 mM KCl, 200 µM of each dNTP (final concentrations) and 80 ng template DNA in 25 µl. Initial denaturation was at 92°C for 5 min. Amplification was for 28 cycles of 95°C for 20 s, annealing for 20 s at 58°C for ABM-C60 and at 60°C for ABM-C29, and extension at 74°C. PCR products were analysed on 6% urea-polyacrylamide sequencing gels and visualised after autoradiography for 2–16 h.

Allele frequencies. Alleles were studied from 40 unrelated Centre d'Etude Polymorphisme Humain (CEPH) parents (80 chromosomes). Seven alleles were observed for ABM-C60. Their sizes, in base pairs (bp), and frequencies were

© Springer-Verlag 1995

Alleles

Fig. 1 Autoradiograph showing the polymorphic marker ABM-C29 (D21S1261) in 10 unrelated individuals. Seven different alleles are shown



Fig. 2 Autoradiograph showing the assignment of ABM-C60 (D21S1224) to band 21q22.3, between the ring 21 and the 21;22 breakpoints using the somatic hybrid cell lines specified *above*

as follows: 140 bp, 0.17; 142 bp, 0.02; 144 bp, 0.02; 146 bp, 0.15; 148 bp, 0.40; 150 bp, 0.23; and 154 bp, 0.02. The observed heterozygosity was 0.83 and the polymorphism information content (PIC) value, 0.74.

ABM-C29 showed 13 different alleles (Fig. 1), whose sizes and frequencies were the following: 185 bp, 0.13; 189 bp, 0.01; 191 bp, 0.46; 193 bp, 0.09; 195 bp, 0.08; 197 bp, 0.05; 199 bp, 0.13; 201 bp, 0.03; 203 bp, 0.04; 205 bp, 0.04; 207 bp, 0.04; 209 bp, 0.03; and 211 bp, 0.01. The observed heterozygosity and PIC value were 0.83 and 0.78, respectively.

Chromosomal localisation. We confirmed the HC21 origin of these clones by amplifying the HC21-specific hybrid cell line WA17. Subchromosomal localisation was determined using a HC21 somatic cell hybrid panel with the following cell lines: 153E7b, 2Fur-1, ACEM, 6;21 (R50-3), 4;21 (GA9-3), 1;21 (1881C-13b), 3;21 (9528c), 8;21 (8q⁻a and 21q⁺), 10;21 (9542c-5a), 21;22 (Raj5), R210 W (ring 21), 7;21(643c-13), 6918–8a1 and MRC2G (Gardiner et al. 1988, 1990).



Fig. 3 Idiogram showing the location of markers D21S1224 and D21S1261, based on the chromosome 21 somatic cell hybrid panel

We assigned both clones to band 21q22.3. ABM-C60 is located between the ring 21 (R210W) and the 21;22 breakpoints (Fig. 2) and ABM-C29 between the 21;22 and the 7;21 breakpoints, distal to ABM-C60 (Fig. 3).

Mendelian inheritance. Co-dominant segregation was observed for ABM-C60 and ABM-C29 in 20 three-generation CEPH families.

Comments. The 21q22.3 band is a high density gene region of chromosome 21. Several genes have previously been identified in this band, such as HMG14, MX1, MX2, BCE1, CBS, CRY, PFKL, COL6A1 and COL6A2, although the function of the coded proteins of many of them remains unresolved. These two highly polymorphic markers should improve the analysis of this region of human chromosome 21, which contains a large number of other genes that are still unidentified.

368

Acknowledgements We acknowledge the Fondo de Investigaciones Sanitarias de la Seguridad Social (94/1905-E), Institut Català de la Salut, European Community Grants CEC/BIOMED1, PL93-0037 and PL93-0107, NICHHD (HD17449) and NCHGR (HG00716).

References

.

- Bosch A, Nunes V, Patterson D, Estivill X (1993) Isolation and characterization of fourteen CA-repeat microsatellites from human chromosome 21. Genomics 18:151–155
- Gardiner K, Watkins P, Munke M, Drabkin H, Jones C, Patterson D (1988) Partial physical map of human chromosome 21. Somat Cell Mol Genet 4:623–638
- Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S, Patterson D (1990) Correlation of physical and cytogenetic maps; gene and CpG island distributions. EMBO J 9:25–34

10

.

4.2.3. Five new microsatellite polymorphisms at the q21 region of human chromosome 21

Bosch A, Wiemann S, Guimera J, Ansorge W, Patterson D and Estivill X

Human Genetics, 95: 119-122 (1995)

El doctorand és l'autor de la determinació de l'heterozigositat, a les famílies de referència CEPH, pel marcador *D21S1264*.

Resultats 199

© Springer-Verlag 1995

Hum Genet (1995) 95:119-122

SHORT, COMMUNICATION

Assumpció Bosch · Stefan Wiemann · Jordi Guimerà Wilhelm Ansorge · David Patterson · Xavier Estivill

Five new microsatellite polymorphisms at the q21 region of human chromosome 21

Received: 11 July 1994 / Revised: 26 July 1994

Abstract Five clones, containing polymorphic CA-repeat sequences, have been isolated from a specific human chromosome 21 phage library and have been localised to band q21 of chromosome 21 using a somatic cell hybrid panel. These highly repetitive sequences (D21S1263, D21S1264, D21S1415, D21S1417 and D21S1420) have been characterised in the CEPH reference parents and have heterozygosities ranging from 0.30 to 0.81 and an average polymorphism information content (PIC) of 0.62. The relative order of these markers, based on the somatic cell hybrid panel, is cen-D21S1417, D21S1420-D21S1263, D21S1415-D21S1264-tel. The most polymorphic marker (D21S1264) has been included in the chromosome 21 genetic map. They have also been localised in the CEPH/ Généthon YAC panel, providing a refined localisation of these polymorphic sequences. These five CA-repeat markers should provide a better characterisation of the q21 region of chromosome 21.

Introduction

Polymorphic DNA markers are excellent tools in human genome mapping, allowing the ordering of DNA fragments by linkage analysis, and the localisation of disease genes (Todd 1992). The genetic map of human chromosome 21 has been considerably improved during the last 2 years (McInnis et al. 1993). Two continuum maps of this chromosome have been published, using overlapping YAC and cosmid clones ordered along the long arm of the chro-

A. Bosch · J. Guimerà · X. Estivill (🖂)

S. Wiemann - W. Ansorge EMBL, Meyerhofstraße 1, D-69117 Heidelberg, Germany

D. Patterson Eleanor Roosevelt Institute, 1899 Gaylord Street, Denver, CO 80206, USA mosome (Chumakov et al. 1992; Nizetic et al. 1994). In addition, the first complete chromosome 21 *Not*I restriction map with its corresponding linking clones has been generated (Ichikawa et al. 1993). Saturation of markers on the chromosome 21 map will improve the level of resolution. For this purpose, the markers of choice are those that are both abundant and highly polymorphic, such as CArepeat polymorphisms, which seem to be uniformly spread throughout the human genome (Weber 1990).

To improve the genetic map of chromosome 21 using CA-repeat microsatellite markers we have isolated and characterised new clones from a human chromosome 21 phage library containing CA-repeat sequences (Bosch et al. 1993). We present here the characterisation of five further clones, which have been localised to band q21 of chromosome 21 using a specific somatic cell hybrid panel. Two of the new polymorphic markers have also been localised to some YACs from the CEPH/Généthon YAC panel.

Materials and methods

An *Eco*RI human chrmosome 21 phage library (LA21NS01) was screened using a $(GT)_{10}$ oligonucleotide (Bosch et al. 1993) and positives were subcloned in pBluescript (Stratagene). CA-repeat flanking regions of each clone were sequenced using T3, T7 and walking primers in a prototype DNA sequencer (EMBL). Primers flanking the CA-repeats were designed and the chromosome 21 origin was confirmed by amplifying DNA from the mouse-human hybrid cell line WA17, which contains chromosome 21 as the only human chromosome.

Sub-localisation of these clones was determined by amplifying DNA from a specific human chromosome 21 Chinese hamster-human somatic cell hybrid panel containing the following cell lines: 153E7b, 2Fur-1, ACEM, JC6, 6;21, 4;21, 1;21, 3;21, 8;21, 10;21, 21;22, R2-10W, 7:21, 6918–8a1 and MRC2G (Gardiner et al. 1988; 1990).

CEPH/Généthon YACs, mapping to the sub-region to which each clone was localised, were encapsulated in agarose beads using a modification of the method described by Overhauser and Radic (1987). These YACs were amplified using PCR with the primers flanking the CA-repeats. Heterozygosity for each marker was determined using a panel containing 40 CEPH (Centre d'Étude du Polymorphisme Humain) reference parents.

Molecular Genetics Department, Cancer Research Institute (IRO), Hospital Duran i Reynals. 08907 L'Hospitalet de Llobregat, Barcelona, Catalunya, Spain

Table 1 Five	new microsate	llite polymorph	nisms at human cl	hromosome 21q21							
D-number	Clone	EMBL	Repeat	Oligonucleotide sequence	MgCl ₂	KCI	Alleles			Ob-	PIC
	2	number	in origi- nal clone		(MILLIN	(14711)	Name	Size	Fre- quency	betero- zygosity	
D2151263	ABM-C20	Z34969	(CA) ₁₇	ABM-C20D1 GCAGCAGCATCAACATTC ABM-C20R1 GAAGAGATGTTGAACTTCACTG	2.4	80	- 9 6 4 9 9	188 190 192 196 198	0.08 0.11 0.05 0.05 0.01 0.08	0.30	0.53
D21S1264	ABM-C57	Z34972	(GT) ₂₂	ABM-C57D1 GGCATAGAAGCCAAACATTT ABM-C57R1 CCAAACGAAATAGGAACAGCT	2.4	80	- 7 6 7 5 5	119 121 123 125 127	0.07 0.26 0.12 0.28 0.05 0.05	0.81	0.78
D2151415	ABM-CI	Z34969	(GT) ₁₃	ABM-CIDI CTGTTAGTTTGACTTTTCTAGAT ABM-CIRI TGATCCCATTTGTTAAACATAC	8.	60	- 0 m 4 m c	81 85 93 101 101 105	0.24 0.04 0.58 0.09 0.03	0.58	0.61
1)2151417	ABM-CI6	Z.34968	(GT) ₁₈	ABM-CI6DI GCAGCAGCACATCAACATTC ABM-C20RI GAAGAGATGTTGAACTTCACTG		60	- 1 5 4 5 5 7 8	116 118 120 122 124 126 128 132	0.01 0.03 0.01 0.16 0.16 0.40 0.22	0.43	0.74
D21S1420	ABM-C43	Z34971	(GA) ₂ (CA), TA(CA) ₁₅	ABM-C43D1 AGGAGCAATAGGTAATGGTCT ABM-C43R1 CAAGTCCCAAAGACATACT	1.8	60	- 0 6 4	147 161 163	0.15 0.74 0.02 0.09	0.33	0.42

200 Resultats-

Asymmetric PCR, where one of the two primers for each dinucleotide repeat was end-labelled with γ -[³²P]-ATP and its concentration was limiting at 1.5 pmol, with 10 pmol of the unlabelled primer, 1.2–2.4 mM MgCl₂, 40–80 mM KCl, 200 μ M each dNTP, 50 ng of human genomic DNA and 0.5 units of *Taq* DNA polymerase (Boehringer Mannheim) in a 25- μ l volume was performed in a 9600 Perkin-Elmer thermal cycler. Initial denaturation was at 92°C for 5 min. Amplification was for 28 cycles of 95°C for 20s, annealing at 58°C for 20s, extension at 74°C for 20s and final extension at 74°C for 5 min. Primer sequences and specific PCR conditions for each marker are shown in Table 1. PCR products were analysed in 6% urea-polyacrylamide sequencing gels and autoradiographs were visualised after exposure for 2–16 h.

Results and discussion

Five clones (ABM-C1, ABM-C16, ABM-C20, ABM-C43 and ABM-C57) containing CA-repeats, isolated from an EcoRI-specific chromosome 21 phage library were characterised. Four of them contained simple repeats, ranging from 13 to 22 dinucleotides, while ABM-C43 contains a complex repeat (Table 1). Their chromosome 21 origin was confirmed using PCR with the WA17 cell line. Subchromosomal localisation to band 21g21 was determined with the specific somatic cell hybrid panel. D21S1417 and D21S1420 mapped between breakpoints 4;21 and 3;21, D21S1263 and D21S1415 between 3;21 and 1;21 and D21S1264 between 1;21 and ACEM (Fig. 1). When PCR was performed using the YACs located in these sub-regions as templates, the following YACs were found positive: 746B10, 90B3 and 757E11 for ABM-C16 (D21S1417) and 759C5 and 745C11 for marker ABM-C57 (D21S1264). All the YACs tested for the presence of ABM-C1 (D21S1415), ABM-C20 (D21S1263) and ABM-C43 (D21S1420) were negative, suggesting that the YAC panel has a gap in this region.

Based on the somatic cell hybrid panel the order of these markers on band q21 is *cen-D21S1417*, *D21S1420-D21S1263*, *D21S1415-D21S1264-tel* (Fig. 1). The locations of *D21S1417* and *D21S1264* have been confirmed on the CEPH/Généthon YAC map but we have not been able to localise markers *D21S1263*, *D21S1415* and *D21S1421* using YACs, so additional chromosome 21 YACs need to be developed to fill the gap which exists in this region of 21q21. These markers could be useful for this purpose. As the sequence tagged sites (STSs) used in this work are polymorphic microsatellite markers, by including them in the linkage map we should be able to confirm their relative locations.

Allele frequencies were estimated from 40 unrelated CEPH parents (80 chromosomes) and they are listed in Table 1. Heterozygosities ranged from 0.30 to 0.81, with an average of 0.50 and polymorphism information content (PIC) values ranged from 0.42 to 0.78, with an average of 0.62 (Table 1). Complete sequences of the polymorphic loci have been submitted to GDB and the EMBL data library. The most polymorphic marker reported in this paper, *D21S1264*, was genotyped in the CEPH reference families and is being included in a linkage map of chromosome 21 (Bosch et al. 1994). D-numbers, clone names,



Fig.1 Idiogram showing the relative location of five new microsatellite markers on band q21 of human chromosome 21 based on somatic cell hybrids

EMBL accession numbers, the repeat block in each original clone and oligonucleotide sequences, PCR conditions, allele sizes (in base pairs) and frequencies, observed heterozygosities and the PIC for each marker are shown in Table 1. These five new microsatellite polymorphisms should be useful for the accurate characterisation of band q21 of chromosome 21, by refining the YAC map, contributing to the linkage map, and providing useful tools for the screening of this region for diseases whose loci are unknown.

Acknowledgements We thank M. Pritchard for useful comments. This work was supported by Fondo de Investigaciones Sanitarias de la Seguridad Social (94/1905E). Institut Català de la Salut, European Community Grants CEC/BIOMED1 PL93-0037 and PL93-0107 and by NICHHD (HD17449) and NCHGR (HG00716).

References

- Bosch A, Nunes V, Patterson D, Estivill X (1993) Isolation and characterization of fourteen CA-repeat microsatellites from human chromosome 21. Genomics 18: 151–155
- Bosch A, Guimera J, Pereirade Souza A, Estivill X (1994) The EUROGEM map of human chromosome 21. Eur J Hum Genet 2: 244–245
- Chumakov I. Rigault P, Gillou S, Ougen P, Billaut A, Guasconi G, Gervy P. LeGall I, Soularue P. Grinas L, Bougueleret L, Bellanné-Chantelot C, Lacroix B, Barillot E, Gesnouin P, Pook S, Vayseix G, Frelat G, Schmitz A, Sambucy JL, Bosch A, Estivill X, Weissenbach J, Vignal A, Riethman H, Cox D, Patterson D, Gardiner K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Le Paslier D, Heilig R, Antonarakis S, Cohen D (1992) Continuum of overlapping clones spanning the entire human chromosome 21q. Nature 359: 380–386

121

122

- Gardiner K, Watkins P, Munke M, Drabkin H, Jones C, Patterson D (1988) Partial physical map of human chromosome 21. Somatic Cell Mol Genet 4: 623-638
- Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S, Patterson D (1990) Correlation of physical and cytogenetic maps; gene and CpG island distributions. EMBO J 9: 25-34
- Ichikawa H, Hosoda F, Arai Y, Shimizu K, Ohira M, Ohki M (1993) A Notl restriction map of the entire long arm of human chromosome 21. Nature Genet 4: 361-366
- McInnis M, Chakravarti A, Blaschak J, Petersen MB, Sharma V, Avramopoulos D, Blouin JL, König U, Brahe C, Cox T, Warren A, Conover C, Van Broeckhoven C, Litt M, Antonarakis SE (1993) A linkage map of human chromosome 21: 43 PCR markers at average intervals of 2.5 cM. Genomics 16: 562–571
- Nizetic D, Gellen L, Hamvas R, Mott R, Grigoriev A, Vatcheva R, Zehetner G, Yaspo ML, Dutriaux A, Lopes C, Delabar JM, Van Broeckhoven C, Potier MC, Lehrach H (1994) An integrated YAC-overlap and 'cosmid-pocket' map of the human chromosome 21. Hum Mol Genet 3: 759–770

Overhauser J, Radic MZ (1987) Encapsulation of cells in agarose beads for use with pulsed-field gel electrophoresis. Focus 9: 8-9

- Todd JA*(1992) La carte des microsatellites est arrivé! Hum Mol Genet 1: 663-666
- Weber JL (1990) Informativeness of human (dC-dA), · (dG-dT), polymorphisms. Genomics 7: 524-530

4.2.4. Characterisation of three microsatellite polymorphisms (D21S1262, D21S1419 and D21S1421) from band 21q22.1

Bosch A, Guimera J, Patterson D and Estivill X

Human Genetics, 95: 596-598 (1995)

El doctorand és responsable de la determinació de l'heterozigositat, a les famílies de referència CEPH, per als microsatèl·lits D21S1419 i D21S1262.

Resultats 205

© Springer-Verlag 1995

Hum Genet (1995) 95:596-598

DNA VARIANTS

Assumpció Bosch · Jordi Guimerà · David Patterson Xavier Estivill

Characterisation of three microsatellite polymorphisms (D21S1262, D21S1419 and D21S1421) from band 21q22.1

Received: 26 October 1994 / Revised: 18 November 1994

Abstract We have isolated three clones, containing highly polymorphic CA-repeat sequences, from a human chromosome 21 phage library (LA21NS01). These clones have been localised to band q22.1 by using a chromosome 21 somatic cell hybrid panel. D21S1262 is located between breakpoints 6918-8a1 and 3x2S, and D21S1419 and D21S1421 are localised between breakpoints JC6-A and MRC2G. Their observed heterozygosities range between 0.75 and 0.85 as shown by unrelated reference parents from the Centre d'Etude du Polymorphisme Humain. These highly polymorphic markers should be useful for improving the analysis of this region of chromosome 21, which contains important genes such as SOD1, GART and IFNAR.

Introduction

Human chromosome 21 has been intensively studied as it is the smallest human autosome. It contains several genes involved in some severe genetic diseases, such as familial amyotrophic lateral sclerosis, some early-onset forms of familial Alzheimer's disease and some forms of leukaemia (Antonarakis 1993). In addition, three copies of chromosome 21 result in Down syndrome, which is the most common cause of mental retardation, and for which the genes responsible are still unidentified. Based on CpG island density, most of the genes seem to be located on the distal part of chromosome 21, mainly in the q22 region (Gardiner et al. 1990), which is thus of special interest.

Polymorphic DNA markers are excellent tools for human genome mapping, allowing the ordering of DNA

A. Bosch · J. Guimerà · X. Estivill (⊠)

D. Patterson

fragments by linkage analysis and the localisation of disease genes (Todd 1992). Although the genetic map of human chromosome 21 has been considerably improved (McInnis et al. 1993), the saturation of highly informative markers on the chromosome 21 map should increase the level of resolution, thereby facilitating the localisation of important disease genes.

CA-repeat microsatellite polymorphisms seem to be the most suitable type of DNA markers, as they are both abundant and highly polymorphic, and are uniformly spread throughout the human genome (Weber 1990). To improve the genetic map of chromosome 21, we have isolated and characterised several clones, containing CA-repeat sequences (Bosch et al. 1993) from a human chromosome 21 phage library. We present here the characterisation of three further clones that have been localised to band 21q22.1 using a specific somatic cell hybrid panel.

Materials and methods

An *Eco*RI HC21 phage library (LA21NS01) was screened using a $(GT)_{10}$ oligonucleotide as probe. Positive phage clones were amplified by the polymerase chain reaction (PCR) using vector and $(CA)_n$ or $(GT)_n$ primers as described (Bosch et al. 1993). CA-repeat flanking regions of each clone were sequenced using lambda-vector primers or $(CA)_n$ primers with a degenerate 3' end, in an automatic DNA sequencer (Applied Biosystems, 373A). Primers flanking the CA-repeats were designed and their origin from chromosome 21 was confirmed by amplifying DNA from the mouse-human hybrid cell line WA17, which contains chromosome 21 as the only human chromosome (Raziuddin et al. 1984).

The localisation of these clones to specific regions on chromosome 21 was determined by amplifying DNA from a chromosome 21 / Chinese hamster-human somatic cell hybrid panel containing the following cell lines: 153E7b, 2Fur-1, ACEM, JC6, 6;21 4;21, 1;21, 3;21, 8;21, 10;21 21;22, R2-10W, 7;21, 6918-8a1 and MRC2G (Gardiner et al. 1988, 1990).

Asymmetric PCR (where one of the two primers for each dinucleotide repeat was end-labelled with $\gamma^{-32}P$ or $\gamma^{-33}P[ATP]$ and its concentration was limiting at 1.5 pmol), with 10 pmol of the unlabelled primer, 1.8–2.4 mM MgCl₂, 60–80 mM KCl, 200 μ M each dNTP, 50 ng human genomic DNA, 0.5 u Taq DNA polymerase (Boehringer Mannheim) in a 25- μ l volume, was performed in a 9600 Perkin-Elmer thermal cycler. Initial denaturation was at 92° C for 5 min. Amplification was for 28 cycles of 95° C for 20 s, an-

Molecular Genetics Department, Cancer Research Institute (IRO), Hospital Duran i Reynals, 08907 L'Hospitalet de Llobregat, Barcelona, Catalunya, Spain

Eleanor Roosevelt Institute, 1899 Gaylord Street, Denver, CO 80206, USA

D-number	Clone	EMBL	Repeat	Oligonucleotide sequence	MgCl ₂	KCI	Allele	S		Ob-	PIC
	name	acces- sion no.	in the original clone		()	(mM)	Name	Size	Fre- quency	served hetero- zygos- ity	
D21S1421	ABM-C23	Z46734	(CA)13	ABM-C23D1	1.8	60	1	234	0.21	0.75	0.77
	ē			GTTGAGTGTACACTGAGGTAG			2	236	0.25		
				ABM-C23R1			3	238	0.04		
				GAGAATAATTCTGATAGTGCAT			4	240	0.18		
							5	242	0.30		
							6	244	0.03		
D21S1419	ABM-C33	Z46735	(TG)23	ABM-C33D1	1.8	60	1	153	0.04	0.85	0.75
		,		GCAGCATTTGCCTTCATCAT			2	157	0.02		
				ABM-C33R1			3	159	0.02		
				GGCTGTTCACACTCAACAGA			4	161	0.10		
							5	163	0.10		
							6	165	0.19		
							7	167	0.42		
							8	169	0.08		
							9	171	0.02		
D21S1262	ABM-C82	Z46736	(CA)23	ABM-C82D1	2.4	80	1	167	0.01	0.78	0.79
				CATAATCCCACAGATCTTCTAT			2	175	0.01	2007/20	2002

Table 1 Three new microsatellite polymorphisms at human chromosome 21q22.1

Fig.1 Idiogram showing the location of microsatellite markers > D21S1421, D21S1419 and D21S1262 on band q22.1 of chromosome 21, based on somatic cell hybrids

ABM-C82R1

GGAAGATGGAAGGAAACATGA

nealing at 56-58°C for 20 s, extension at 74°C for 20 s and final extension at 74°C for 5 min. Primer sequences and specific PCR conditions for each marker are shown in Table 1. PCR products were analysed on 6% urea-polyacrylamide sequencing gels and autoradiographs were visualised after exposure for 2-16 h.

Results and discussion

Three clones (ABM-C23, ABM-C33 and ABM-C82) containing CA-repeats (D21S1421, D21S1419 and D21S1262, respectively), isolated from an EcoRI-specific human chromosome 21 phage library were characterised. All contained simple repeats, ranging from 13 to 23 dinucleotides (Table 1). Their HC21 origin was confirmed using PCR with DNA from cell line WA17. Sub-chromosomal localisation to band 21q22.1 was determined with the specific chromosome 21 somatic cell hybrid panel. D21S1421 and D21S1419 mapped between breakpoints JC6-A and MRC2G, and D21S1262 was localised more distally, between breakpoints 6918-8a1 and 3x2S (Fig. 1).

Allele frequencies were estimated using 40 unrelated Centre d'Etude du Polymorphisme Humain (CEPH) par-



3

4

5

6

7

8

9

179

181

183

185

187

189

191

0.18

0.20

0.31

0.16

0.11

0.01

0.01

597

598

ents (80 chromosomes) and are listed in Table 1. Heterozygosities ranged from 0.75 to 0.85, and polymorphism information content (PIC) values ranged from 0.75 to 0.79 (Table 1). Complete sequences of the polymorphic loci have been submitted to the Genome Data Base and the EMBL Data Library. These markers have been genotyped in the CEPH reference families and are included in a linkage map of human chromosome 21 (A. Bosch et al., unpublished). D-numbers, clone names, EMBL accession numbers, the repeat block sequenced in each original clone, oligonucleotide sequences, PCR conditions, allele sizes (in base pairs) and frequencies, observed heterozygosities and PIC values for each marker are shown in Table 1. These three new microsatellite polymorphisms should be useful for an accurate characterisation of band 21q22.1, by contributing to the linkage map. They could also aid the screening of this region for diseases whose loci are unknown or for genetic studies of the families affected by some of the inherited diseases whose loci are contained in this band, such as familial amyotrophic lateral sclerosis which is caused by mutations in the SOD1 gene (Rosen et al. 1993). In addition, they are powerful sequence tagged sites for the physical map of the region.

Acknowledgements We thank M. Pritchard for useful comments. This work was supported by Fondo de Investigaciones Sanitarias de la Seguridad Social (94/1905E), Institut Català de la Salut, European Community Grants CEC/BIOMED1 PL93-0037 and PL93-0107, and by NICHHD (HD17449) and NCHGR (HG00716).

References

- Antonarakis SE (1993) Human chromosome 21: genome mapping and exploration, circa 1993. Trends Genet 9:142-148
- Bosch A, Nunes V, Patterson D, Estivill X (1993) Isolation and characterization of fourteen CA-repeat microsatellites from human chromosome 21. Genomics 18:151–155
- Gardiner K, Watkins P, Munke M, Drabkin H, Jones C, Patterson D (1988) Partial physical map of human chromosome 21. Somat Cell Mol Genet 4:623–638
- Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S, Patterson D (1990) Correlation of physical and cytogenetic maps; gene and CpG island distributions. EMBO J 9:25-34
- McInnis M, Chakravarti A, Blaschak J, Petersen MB, Sharma V, Avramopoulos D, Blouin JL, König U, Brahe C, Cox T, Warren A, Conover C, Van Broeckhoven C, Litt M, Antonarakis SE (1993) A linkage map of human chromosome 21:43 PCR markers at average intervals of 2.5 cM. Genomics 16:562-571
- Raziuddin A, Sarkar FH, Dutkowski R, Shulman L, Ruddle FH, Gupta SL (1984) Receptors for human alpha and beta interferon but not gamma interferon are specified on human chromosome 21. Proc Natl Acad Sci USA 81:5504-5508
- Rosen DR, Siddique T, Patterson D, Fliglewicz DA, Sapp P, Hentati A, Donaldson D, Gote J, O'Regan JP, Deng HX, Rhamani Z, Krizus A, McKenna-Yasek K, Cayabyab A, Gaston SM, Berger R, Tanzi RE, Halperin JJ, Herzfeldt B, Van den Bergh R, Hung WY, Bird T, Deng G, Mulder DW, Smyth C, Laing NG, Soriano E, Pericak-Vance MA, Haines J, Rouleau GA, Gusella JS, Horvitz HR, Brown RH (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 362:59-62
- Todd JÅ (1992) La carte des microsatellites est arrivé!. Hum Mol Genet 1:663-666
- Weber JL (1990) Informativeness of human (dC-dA),(dG-dT), polymorphisms. Genomics 7:524-530

4.2.5. New alleles at microsatellite *loci* in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell lines

Banchs I, Bosch A, Guimera J, Lazaro C, Puig A, Estivill X

Hum Mutat., 3: 365-372 (1994)

El doctorand és responsable de la caracterització de les mutacions corresponents als microsatèl·lits D21S364, D21S365, D21S367, D21S368, D21S369, D21S406, D21S408, D21S409, D21S414, D21S415, D21S416, D21S1235, D21S1261 i D21S1262.

RESEARCH ARTICLE

New Alleles at Microsatellite Loci in CEPH Families Mainly Arise From Somatic Mutations in the Lymphoblastoid Cell Lines

I. Banchs, A. Bosch, J. Guimerà, C. Lázaro, A. Puig, and X. Estivill*

Molecular Genetics Department, Cancer Research Institute (I.R.O.) Campus Bellvitge, Hospitalet de Llobregat, 08907 Barcelona, Catalunya, Spain; Fax: 34-3-263-2251

Communicated by Robert Williamson

In the analysis of 40 CEPH families, under the EUROGEM project, with a total of 29 microsatellites (26 CA-repeats, a TCTA-repeat within the vWFII-3 gene, a TTA-repeat within the PLA-2 gene, and an AAAT-repeat intragenic to the NFI gene) from human chromosomes 12, 17, and 21, we have detected 21 cases of abnormal segregation of alleles in 16 pedigrees for a total of 14 markers (48%). In 11 cases, the abnormal transmissions were of somatic origin, 10 of which (91%) occurred in the lymphoblastoid cell lines. In 9 other cases, it was not possible to determine if the origin of the new alleles was somatic or germline, and in one case hemizygosity in several family members was observed, so its origin was germline. The 20 new mutations detected in the 22,852 meioses analysed represent a mutation frequency of 8.7×10^{-4} per locus per allele. The germline mutation rate could be as high as 3.9×10^{-4} per locus per gamete (from 0 to 3.9×10^{-4}), but the rate of somatic mutations detected in the study was much higher (4.8 \times 10⁻⁴ to 8.7 \times 10⁻⁴ per locus per allele). Individual mutation rates ranged from 0 to 3.8×10^{-3} . Among the markers analysed, all three that were tri- or tetranucleotide repeats showed one or two new alleles, compared to only 10 of the 26 (38%) CA-repeats showing mutations. Three CEPH families (102, 45 and 1333) each had several mutational events, and one individual (10210) had somatic mutations for two microsatellites from different chromosomes. The mutation rate at microsatellite loci within families, using DNA directly obtained from cells from the individual, is less than 1×10^{-4} (true germline mutation rate), which should not affect the use of these markers in diagnosis and linkage. However, these results and previous data suggest that for DNA obtained from cell lines, mutations are much more frequent (1 × 10^{-2} -1 × 10^{-3}). © 1994 Wiley-Liss, Inc.

KEY WORDS: Alleles, CEPH families, Lymphoblastoid cell lines

INTRODUCTION

Microsatellite markers, or short tandem repeat polymorphisms (STRPs), have proved to be excellent tools in genetic analysis. The large number of alleles detected at microsatellite loci, their high heterozygosity and the easy and uniform method of analysis, by means of the polymerase chain reaction (PCR), have made these markers extremely useful in human gene mapping, paternity testing, forensic analysis, diagnosis of genetic disorders by linkage, loss of heterozygosity studies in cancer, and hemizygosity detection of deletions in inherited genetic disorders (Todd, 1992).

Among the different microsatellite sequences, the CA-repeat class is the most abundant and uniformly spread throughout the human genome, with an average of one CA-repeat block every 40 kb (Weber, 1990). Another useful microsatellite class is that composed by AAAT-repeats, which are common at the 3' end of Alu-repeat sequences (Economou et al., 1990). Although tetranucleotide repeats are less represented in the human genome than CA-repeats, the interpretation of allele sizes is often easier with tetranucleotides, than with dinucleotide repeats. The generation of human genetic linkage maps is mainly being accomplished using microsatellite markers, mostly CArepeats, analysed in the large three-generation

© 1994 WILEY-LISS, INC.

Received September 28, 1993; accepted November 23, 1993. *To whom reprint requests/correspondence should be addressed.

families distributed by CEPH (Centre d'Etude du Polymorphisme Humain) (Weissenbach et al., 1992).

In the analysis of 40 CEPH families with a total of 29 microsatellite markers, we have detected 21 cases of abnormal segregation of alleles in 16 pedigrees. In 11 cases, the abnormal transmissions were of somatic origin, the majority (91%) originating in the lymphoblastoid cell lines; in 9 other cases, it was not possible to determine if the new alleles were of germline or somatic origin; and only in one case the abnormal transmission of alleles was due to hemizygosity, which was germline in origin.

MATERIALS AND METHODS Human DNA Samples

DNAs from 40 CEPH families were kindly provided by CEPH, Paris, under the EUROGEM programme. For those cases in which ambiguous results were obtained, the analysis was carried out independently by a second person. When a mutation event was confirmed, additional DNA samples of the individuals (2–8 samples per individual) were analyzed. The additional samples were from different cultures or different batches of the same culture, and were kindly provided by CEPH. The DNA samples analyzed were not only from individuals whose cell line had the mutation but also from their parents and grandparents, where possible.

Microsatellite Markers and PCR Analysis

The microsatellites analysed in the study were 11 CA-repeat markers from human chromosome 12 (HC12): (D12S43, D12S77, D12S78, D12S79, D12S82, D12S83, D12S84, D12S87, D12S94, D12S95, D12S97) (Weissenbach et al., 1992); 15 CA-repeats from human chromosome 21 (HC21): (D21S364, D21S365, D21S366, D21S367, D21S368, D21S369. D21S408, D21S406, D21S409. D21S414. D21S415. D21S416, D21S1235, D21S1261, D21S1262) (Bosch et al., 1993a,b; in press; and unpublished); an AAATrepeat intragenic to the NF-1 gene [Xu et al., 1991]; a TTA-repeat in the PLA-2 gene (Polymeropoulos et al., 1990); and a TCTA-repeat in the vWFII-3 gene (Van Amstel and Reistma, 1990). Primer sequences have been described (Weissenbach et al., 1992; Polimeropoulos et al., 1990; Van Amstel and Reistma, 1990; Bosch et al., 1993a). Oligonucleotide primers flanking the repeated sequences were designed, synthesized (Applied Biosystems 391 PCR-Mate DNA Synthe-

sizer) and used without purification in PCR reactions with human genomic DNA. Asymmetric PCR was performed where one of the two primers for each dinucleotide repeat was end labelled with γ -³²P[ATP] and its concentration was limiting (1.5/10 pmol of the unlabelled primer), with 1.2-2.4 mM MgCl₂, 40-80 mM KCl, 200 mM each dNTP," 50 ng of human genomic DNA and 0.5 units of Tag DNA polymerase (Boehringer Mannheim) in a 20-µl volume in a 9600 Perkin-Elmer thermalcycler. Initial denaturation was at 90°C for 5 min. Amplification was for 28 cycles of 95°C for 20 sec, annealing for 20 sec at 55-60°C, extension at 74°C for 20 sec; and final extension at 74°C for 5 min. PCR products were analyzed on 6% ureapolyacrylamide sequencing gels and autoradiographs were visualized after 2 hr to overnight exposure at -80°C.

RESULTS

We have analysed a total of 26 CA-repeat markers (HC12 and HC21), a TCTA-repeat within the vWFII-3 gene, a TTA-repeat within the PLA-2 gene (HC12), and an AAAT-repeat intragenic to the NF1 gene (HC17) in 40 CEPH families. Data from HC21 and HC12 are being used in the construction of high-resolution maps for these two chromosomes (Bosch et al., 1993a), whereas the AAAT-NF1 microsatellite is a useful tool in the detection of deletions and in indirect genetic analysis of neurofibromatosis type 1 patients (Lázaro et al., 1992, 1993).

We have detected 21 cases of abnormal segregation of microsatellite alleles in 16 of the 40 CEPH families analysed. 20 of these cases were de novo mutations, whereas in one case, a deletion was found to segregate in the members of the family. A total of 788 meioses for each of the 29 microsatellites were analysed (almost 23,000 meioses, or parent to child transmissions, in total).

Germline or Somatic Mutations

In 9 cases, it was not possible to determine whether the mutation was somatic or germline (Table 1). Individual 3507 had one allele for marker ABM-C13 (D21S368), not present in either of the parents, due to either an increase of one or three CA-repeats. In family 1331 the new allele in individual 133117, for the TCTA-repeat at the vWFII-3 gene, was either the addition of one repeat or the deletion of two repeats. Individual 133306 had a new allele for marker AFM026tb5 (D12S77), being the increase of one or two CArepeats (Table 1). In individual 10212, for marker

MICROSATELLITE SOMATIC MUTATIONS 367

Individual	Locus	Change	Origin	Туре
3507	D21S368	† 1/3 CA	P/M	G/S
133117	vWFII-3	↑1 TCTA/↓ 2 TCTA	м	G/S
133306	D12S77	† 1/2 CA	Р	G/S
10212	D12S84	1 2 CA	Р	G/S
88407	D21S415	† 1 CA	Р	G/S
134006	NF1	1 AAAT •	M	G/S
1329408	D12S79	12 CA	Р	G/S
134605	D12S78	L1 CA	Р	G/S
10213	D12S78	↓1 CA	Р	G/S
10210	D12S84	13 CA	м	S(L)
10210	D21S1235	1 CA	P	S(L)
6610	D12S84	↓ 1/3/ ↑ 2/3 CA	P/M	S(L)
4509	D12S95	† 1/6 ĊA	P/M	S(L)
4504	D12S95	15 CA	Р	S(L)
133302	D21S416	† 1 CA	M	S(L)
1701	D21S416	11 CA	P/M	S
142302	D21S406	del./ † 5 CA	Р	S(L)
142403	PLA-2	del./ † 1 TTA	Р	S(L)
135010	PLA-2	↑ 2/3/5/↓3 TTA	P/M	S(L)
134907	NF1	† 1 AAAT	P/M	S(L)
1329211	D21S369	del.	P/M	G

TABLE 1. Mutations Detected in the Analysis of 40 CEPH Families for 29 Microsatellite Loci

P, paternally contributed chromosome; M, maternally contributed chromosome; G, germline; S, somatic; L, lymphoblastoid; \uparrow , increase; \downarrow decrease; del., deletion.

AFM116 \times b8 (*D12S84*), the mutation was an increase of 2 CA-repeats on the paternally contributed chromosome (Fig. 1).

In family 884, a new allele for marker ABM-C27 (D21S415), due to the increase of one CArepeat on the paternally contributed chromosome, was detected in one of the offspring (88407). In family 1340, the mutation was in individual 134006 for the AAAT-repeat in the NFI gene, being the addition of one AAAT-repeat on the maternally contributed chromosome. In family 13294 a deletion of 2 CA-repeats was found in individual 1329408 for marker AFM067yc5 (D12S79). For marker AFM026tf3 (D12S78), a deletion of one CA-repeat on the paternally contributed chromosome was detected in individual 134605. In family 102, a loss of one CA-repeat on the paternally contributed chromosome was detected in individual 10213 (Table 1).

In these 9 cases, the mutations might be germline in origin, but the possibility of a somatic origin, either arising in the lymphoid line of the individual or in the establishment of the lymphoblastoid cell line, can not be excluded. These mutant alleles were detected in the youngest generation of the respective families and not in the parents; thus, it was not possible to characterize the new alleles with respect to transmission in the families. The analysis of new DNA samples provided by CEPH (2–8 new DNA samples for several members of each family) reproduced the abnormal





FIGURE 1. Three new mutations detected at locus *D12S84* for families 66 and 102. Mutation in family 66 originated in the cell line as detected in several DNA aliquots (not shown). In individual 10212 the mutation could be germline or somatic in origin, whereas in 10210 the mutation was somatic in origin. FF, father of father; FM, mother of father; F, father; CF, female child; and M, mother. Letters a-c indicate different aliquots. The original aliquots analysed were those indicated by a.

pattern detected in each case, which would support the mutation having arisen in the individual or in the germline.

Somatic Mutations

In 11 cases the mutations identified were of somatic origin. A mutation was found in family 102

¥ .

214 Resultats-

368 BANCHS ET AL.

for marker AFM116×b8 (D12S84) in individual 10210 (allele 9 arising from either allele 5 or 6 of the maternal chromosome, which is a loss of either 3 or 4 CA-repeats). Two new DNA samples from the same individual showed three alleles, 4, 6, and 9. (allele 9 had a faint signal; Fig. 1, sample 10c). Therefore, the mutation was the loss of 3 CArepeats on the maternally contributed chromosome which originated in the lymphoblastoid cell line. In the same individual (10210) three alleles were detected for marker ABM-C75 (D21S1235). The analysis of DNA from different DNA samples of the same cell line showed that the new allele was due to the increase of one CA-repeat, arising on the paternally contributed chromosome in the lymphoblastoid cell line (Table 1).

In family 66, individual 6601 has a new allele (allele 5), for marker AFM116 \times b8 (D12S84), not inherited from either of his parents, and which he has passed to three of his offspring. For individual 6610, the initial DNA sample analysed was heterozygous with alleles 3 and 6. The analysis of further DNA samples from other cell cultures detected three alleles (alleles 3, 5 and 6) in this individual. Therefore, the abnormal segregation of alleles in family 66 was due to a somatic mutation in the lymphoblastoid cell line of individual 6610 (Fig. 1).

In family 45, two different mutations were detected for marker AFM207ve1 (D12S95). In individual 4509 three alleles (3, 4, and 9) were found. In one of the new samples of DNA analysed, the de novo allele 3 was absent, confirming a mutation in the lymphoblastoid cell line, due to the increase of either 1 or 6 CA-repeats. Individual 4504 was heterozygous for alleles 3 and 9. The analysis of new DNA samples showed heterozygosity for alleles 3 and 4, therefore the mutation was of somatic origin in the lymphoblastoid cell line and due to the deletion of 5 CA-repeats on the paternally contributed chromosome (Fig. 2).

CEPH families 1333 and 17 presented new alleles due to increases in one CA-repeat unit for marker ABM-C19 (D21S416): from 13–14 repeats in individual 133302, and from 17 to 18 CA-repeats in individual 1701. As the mutations were not segregating in the families, they represent somatic mutations in either the lymphoblastoid cell line or the lymphoid line of these individuals. For the latter possibility, the mutation must have occurred very early in development with substitution of the original cells by those carrying the new allele. In individual 133302, the mutation arose on the maternally contributed chromosome, whereas CEPH Family 45



FIGURE 2. Two somatic mutations detected for marker D12S95 in family 45, both arising in the lymphoblastoid cell line. F, father; C/F, female child; C/M, male child; and M, mother. Letters a-d indicate different aliquots. The first aliquots analysed were those indicated by a.

it was not possible to determine the parental origin of the new allele in individual 1701, since he was the first member of the family from which DNA was available. The analysis of new DNA samples confirmed the origin in the lymphoblastoid cell line for the new allele in individual 133302, but not in 1701.

In family 1423 homozygosity for allele 6 (18 repeats) at locus D21S406 (ABM-C52) was detected in the mother (142302), whereas the grandfather was homozygous for allele 1 (13 repeats) and her kindred inherited alleles 6 and 1. Overexposure of the autoradiograph showed a faint band that corresponds to allele 1 in the mother. The analysis of a second DNA sample of this individual did not detect homozygosity for this locus, demonstrating that the change at D21S406 was due to a mutation of somatic origin in the lymphoblastoid ycell line (Fig. 3).

Two cases of de novo mutations were detected for the TTA-repeat in the PLA-2 locus on HC12 (Fig. 4). In family 1424 homozygosity for allele 2 was found in individual 142403, which could be due to the addition of one repeat or to the deletion of the whole locus, originating on the paternally contributed chromosome. A new DNA sample showed that 142403 was heterozygous for alleles 2 and 3, although the intensity of allele 3 was weaker than allele 2. This agrees with a somatic mutation from allele 3 to allele 2 or with the deletion of allele 3, occurring in the lymphoblastoid cell line.

MICROSATELLITE SOMATIC MUTATIONS 369



FIGURE 3. Mutation of somatic origin at the D21S406 locus in individual 142302. Arrows and letters (a, b) indicate the different DNA batches analysed.

In family 1350, individual 135001 has an apparent new allele (either allele 6 or 7) for the TTA-PLA-2 microsatellite, which he has passed to his offspring. This new allele might be a germline mutation in the paternally contributed chromosome, consisting of the loss of one or two TTA-repeats. The possibility of uniparental disomy for chromosome 12, generated at the first meiotic division was excluded with the correct segregation of alleles for 11 other HC12 markers in this family. The analysis of new DNA batches for members of this family showed that individual 135010 had three alleles (2, 5, and 7). Allele 7 was passed to his son, therefore, either allele 2 or allele 5 was generated during the growth of the lymphoblastoid cell line, leading in some DNA batches to the presence of the de novo allele 2 or 5 and to the loss of the original allele 7. Therefore, the mutation occurred in the lymphoblastoid cell line (Fig. 4).

In family 1349 a new allele (allele 1) at the AAAT-repeat within the NF1 gene was detected in individual 134907. The father was homozygous for allele 2 and the mother heterozygous for alleles 2 and 3. Although the mutation could be either germline or somatic, the analysis of a second DNA sample showed a weaker signal corresponding to the new allele, which supports a somatic mutation in the cell line (Table 1).

Hemizygosity at D21S369 of Germline Origin

In CEPH family 13292 the grandmother (1329211) was homozygous for allele 6 for marker



FIGURE 4. New mutations at the TTA-repeat of the PLA-2 gene in CEPH families 1424 and 1350. F, father; CM, male child; M, mother; FF, father of father; F, father; letters a-c indicate different alignots analysed, with a the first sample tested.

ABM-C15 (D21S369). However, the analysis of the other members of the family was consistent with her being hemizygous for locus D21S369. Therefore, individual 1329211 may carry a germline deletion, which has been transmitted to her son and to the whole kindred. We can not exclude the possibility of a point mutation at the 3' end of one of the primers used to amplify this locus, as has been described previously as a cause of the failure to amplify some microsatellite loci (Koorey et al., 1993) (Fig. 5).

DISCUSSION

Microsatellite markers are a universal resource of polymorphism for the study and diagnosis of genetic disorders, identification analysis in forensic medicine and paternity testing, studies of loss of heterozygosity in cancer, linkage disequilibrium analysis, the study of multifactorial disorders, and mapping of the human genome (Todd, 1992). Although the basis for the informativity of polymorphic markers is the variability that they exhibit, the reliability of the studies is based on their stability. The stability of microsatellite alleles is still not well studied. Based on the observation of mutational events through genotyping a large number of families, mutation rates at microsatellite loci has been estimated to be between 1.0×10^{-3} -1.0 x 10⁻⁴ in CEPH families (Kwiatkowski et al., 1992; Weissenbach et al., 1992; Bowcock et al., 1993;

370 BANCHS ET AL.



FIGURE 5. Hemizygosity at D21S369 of germline origin, in individuals 1329201, 03, 06, 08, and 09. Hemizygosity or homozygosity cannot be unambiguously established in family members 04, 05 and in the grandmother (11).

Mahtani and Willard, 1993; Petrukhin et al., 1993; Straub et al., 1993; Weber and Wong, 1993), but a higher frequency of mutation has been detected in tumour samples, about 3.0×10^{-3} (Jones et al., 1992).

We have detected 20 new mutations at microsatellite loci in the CEPH families, for a total of 29 markers analysed. 11 mutations have a somatic origin and for 9 it was not possible to determine their origin. Possible contamination of the DNA samples and nonpaternity were excluded using new DNA samples in the mutated cases and by the segregation analysis of the other 28 microsatellite markers in the families, in which the new alleles were detected. In several cases (6610 for marker D12S84; and 135010 for marker PLA-2), the individual in which the new allele arose was only unambiguously determined after the study of new samples from the index individual and samples from parents and grandparents.

Mutations were found for 13 markers (45%). For 7 of these markers (54%), the mutations affected a single individual, for 5 markers two new alleles were found (38%), and for one marker three independent mutation events were detected. The 20 mutations detected in the 22,852 meioses analysed represent a mutation frequency of 8.7×10^{-4} per locus per allele. None of the mutations detected can truly be regarded as germline in origin. If the mutations that could be either somatic or

germline are considered, the germline mutation rate could be as high as 3.9×10^{-4} per locus per gamete (from 0 to 3.9×10^{-4}). However, the rate of somatic mutations detected in the study is much higher than that for potential germline mutations, at 4.8×10^{-4} - 8.7×10^{-4} per locus per allele. The individual mutation rates for the microsatellites analysed range from 0 to 3.8×10^{-3} .

The mutation frequencies determined in our study are similar to those estimated in other studies in which CEPH families have been analysed (Kwiatkowski et al., 1992; Weissenbach et al., 1992; Bowcock et al., 1993; Mahtani and Willard, 1993; Petrukhin et al., 1993; Straub et al., 1993; Weber and Wong, 1993). However, we have been able to demonstrate that the majority of new alleles arising in the CEPH families were due to somatic mutations, with a mutation rate as high as 8.7×10^{-4} , similar to that reported recently by Weber and Wong (1993), who studied a similar number of parent–offspring transfers.

For the somatic mutations the abnormality could have arisen in the precursor cells of the circulating lymphocytes at a very early stage of development or in vitro, as a clonal expansion of a cell in the generation or growth of the cell line. For 10 of the 11 (91%) unambiguous somatic cases, it was possible to demonstrate that the mutations occurred in the cell lines, because we observed different typings with different DNA batches, even from the same cell culture but with DNA obtained from different flasks. Lymphoblastoid cell lines are thought to be very stably maintained in culture through cell generations (Neitzel, 1986). However, the data presented here suggest that mutations at microsatellite loci are much more common in these cell lines compared to mutation frequencies in individuals. The higher frequency of somatic mutations at microsatellite loci, as opposed to germline mutations, is in agreement with data obtained by Weber and Wong (1993), who demonstrated the somatic origin of the mutations in the CEPH families, using DNA from blood lymphocytes in a large proportion of cases. Allelic losses in the CEPH families have also been reported for minisatellite markers (Royle et al., 1993).

The majority of the markers analysed here were CA-repeats (26) and mutations were found for 10 of them (38%). Although most of the available data on new mutations comes from CA-repeat microsatellites, there are also several observations and a larger study on tri- and tetranucleotide repeat markers (Weber and Wong, 1993), which suggest higher mutation rates for these repeats than for CA-repeats (Kwiatkowski et al., 1992; Lázaro et al., 1993; Mahtani et al., 1993; Weber and Wong, 1993; S.E. Antonarakis, personal communication). In fact, among the markers analysed here, all three that were not CA-repeats (tri- or tetra-repeats) had one or two new alleles, compared to 10 of the 26 CA-repeats showing one or several mutations.

Three of the families studied had several mutational events (4 in family 102; 2 in family 45; and 2 in family 1333), 5 of which were somatic and the other 3 either germline or somatic in origin. It is interesting that one individual (10210) had somatic mutations for two microsatellites on different chromosomes, one on the paternal chromosome (D21S1235) and the other on the maternal chromosome (D12S84), both arising in the lymphoblastoid cell line. Another family (45) had two independent in vitro somatic mutation events in two different members of the family. These families are among those whose cell lines have been established for longer, so they may have accumulated a larger number of alterations.

Among the germinal or somatic cases, most mutations in CA-repeats (86%) originated on the paternally contributed chromosome, whereas 2 of the tetranucleotide repeat mutations arose from the maternally contributed alleles. Only in 6 of the 11 cases of somatic mutation was the chromosomal origin clear (4 paternal and 2 maternal). The slightly larger number of new alleles on the chromosome of paternal origin observed in the group of germinal or somatic cases, suggests that some of them may be germinal. However, our data do not allow us to define a preference for germinal mutations in males rather than in females, as has been proposed (Weber and Wong, 1993).

Abnormal segregation of alleles in CEPH families has also been shown to be due to deletions, leading to apparent homozygosity and to errors in linkage analysis (Royle et al., 1993). In our study, only in two cases (D21S406 in family 1423, and PLA-2 in family 1424) could the somatic mutations be the result of allelic losses, although a mutation within the repeat sequence can not be excluded. In a third case (D21S369 in family 13292), hemizygosity for several members of the family could be due to a true deletion (not generated in the analysed members). This suggests that allelic losses leading to hemizygosity may not be easily detected when diallelic markers or loci with low heterozygosities are analysed. Also, mutations at microsatellite loci in the offspring of CEPH famiResultats 217

lies might remain undetected if the parents share alleles and the mutation is the creation of one allele that is already present in the parents.

The data on de novo mutations at microsatellite loci obtained here, together with previously available data (Jones et al., 1992; Kwiatkowski et al., 1992; Weissenbach et al., 1992; Bowcock et al., 1993; Lázaro et al., 1993; Mahtani and Willard, 1993; Morral et al., 1993; Petrukhin et al., 1993: Straub et al., 1993; Weber and Wong, 1993) suggest that the mutation rate within families, using DNA directly obtained from cells of the individual, is of less than 1×10^{-4} (true germline mutation rate), which should not affect the use of these markers in diagnosis by linkage. However, care should be taken when analysing DNA obtained from cell lines with microsatellite markers, as mutations within the repeat sequences are much more frequent $(1 \times 10^{-2} - 1 \times 10^{-3})$ (somatic mutation rate). This higher mutation rate in cell line DNA may affect the maps of human chromosomes with resolutions of >2 centimorgans (cM), which are based on few recombinational events. Although this mutation rate should not compromise the reliability of the genetic maps that have so far been generated using microsatellite markers (Weissenbach et al., 1992), nor their application to disease linkage studies and diagnosis, but a high level of resolution will probably not be possible using DNA samples obtained from lymphoblastoid cell lines.

ACKNOWLEDGMENTS

We thank Agnes Marcadet for providing unlimited aliquots of the relevant members of the CEPH families in which de novo mutations were detected, and Antonia Gaona for excellent technical assistance. We acknowledge Melanie Pritchard and Helena Kruyer for help with the manuscript. This work was supported by the Fondo de Investigaciones Sanitarias de la Seguridad Social (92/ 0885E and 92/0532), the Comisión Interministerial Científica y Técnica (SAL91-0795-CE), the European Community (grant CEC PL910014, EUROGEM Project), and the Institut Català de la Salut (Generalitat de Catalunya). A.B. was supported by a grant of the Ministerio de Educación y Ciencia and C.L. by the Generalitat de Catalunya.

REFERENCES

Bowcock A, Osborne-Lawrence S, Barnes R, Chakravarti A, Washington S, Dunn C (1993) Microsatellite polymorphism linkage map of human chromosome 13q. Genomics 15:376– 386.

Bosch A, Nunes V, Patterson D, Estivill X (1993a) Isolation and

characterization of fourteen CA-repeat microsatellites from human chromosome 21. Genomics 18:151-155.

- Bosch A, Wiemann S, Guimerà J, Ansorge W, Patterson D, Estivill X (1993b) Two dinucleotide repeat polymorphisms at 21q22.3 (D21S416 and D21S1235) Hum Mol Genet 2:1744.
- Bosch A, Wiemann S, Ansorge W, Patterson D, Estivill X (1993c) Three CA/GT repeat polymorphisms from loci D21S414 and D21S1234 on Human Chromosome 21. Hum Genet (in press).
- Economou EP, Bergen AW, Warren AC, Antonarakis SE (1990) The polydeoxyadenylate tract of Alu repetitive elements is polymorphic in the human genome. Proc Natl Acad Sci USA 87:2951–2954.
- Jones MH, Nakamura Y (1992) Somatic mutations at CA-repeat loci. Hum Mutat 1:224-228.
- Koorey DJ, Bishop GA, McCaughan GW (1993) Allele non-amplification: a source of confusion in linkage studies employing microsatellite polymorphisms. Hum Mol Genet 2:289–291.
- Kwiatkowski DJ, Henske EP, Weimer K, Ozelius L, Gusella JF, Haines J (1992) Construction of a GT polymorphism map of human 9q. Genomics 12:229–240.
- Lázaro C, Ravella A, Casals T, Volpini V, Estivill X (1992) Prenatal diagnosis of sporadic neurofibromatosis 1. Lancet 339: 119-120.
- Lázaro C, Gaona A, Ravella A, Volpini V, Casals T, Fuentes J-J, Estivill X (1993) Novel alleles, hemizygosity and deletions at an Alu-repeat within the neurofibromatosis type 1 (NF1) gene. Hum Mol Genet 2:725–730.
- Mahtani MM, Willard HF (1993) A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. Hum Mol Genet 2:431-437.
- Morral N, Nunes V, Casals T, Chillón M, Giménez J, Bertranpetit J, Estivill X (1993) Microsatellite haplotypes for cystic fibrosis:

mutation frameworks and evolutionary tracers. Hum Mol Genet 2:1015–1022.

- Neitzel H (1986) A routine method for the establishment of permanent growing lymphoblastoid cell lines. Hum Genet 73: 320-326.
- Petrukhin KE, Speer MC, Cayanis E, De Fátima Bonaldo M, Tantravahi U, Soares MB, Fischer SG, Warburton D, Gilliam TC, Ott J (1993) A microsatellite genetic linkage map of human chromosome 13. Genomics 15:76–85.
- Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1990) Trinucleotide repeat polymorphism at the human pancreatic phospholipase A-2 gene (PLA-2). Nucleic Acids Res 18:7468.
- Royle NJ, Armour JA, Crosier M, Jeffreys AJ (1993) Abnormal 'segregation of alleles in CEPH pedigree DNAs arising from allele loss in lymphoblastoid DNA. Genomics 15:119–122.
- Straub RE, Speer MC, Luo Y, Rojas K, Overhauser J, Gilliam TC (1993) A microsatellite genetic linkage map of human chromosome 18. Genomics 15:48–56.
- Todd JA (1992) La carte des microsatellites est arrivé! Hum Mol Genet 1:663-666.
- Van Amstel HKP, Reitsma PH (1990) Tetranucleotide repeat polymorphism in the vWF gene. Nucleic Acids Res 18:4957.
- Weber JL (1990) Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. Genomics 7:524-530.
- Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128.
- Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794–801.
- Xu G, Nelson L, O'Connell P, White R (1991) An Alu polymorphism intragenic to the neurofibromatosis type 1 gene (NF1). Nucleic Acids Res 19:3764.

²¹⁸ Resultats–

³⁷² BANCHS ET AL.

4.2.6. The EUROGEM map of human chromosome 21

Bosch A, Guimera J, Pereira de Souza A, Estivill X.

European Journal of Human Genetics, 2: 244-245 (1994).

El doctorand és l'autor del genotipatge dels 532 individus de les famílies de referència CEPH per a cadascun dels microsatèl·lits D21S369, D21S408, D21S415, D21S409, D21S364, D21S406, D21S414, D21S368, D21S1367, D21S1264, D21S1365, D21S1262, D21S1235, D21S1224 i D21S1261 utilitzats en la construcció del mapa de Iligament genètic de l'HC21. És també l'autor de la comprovació del genotip de tots aquells individus que presentaven doble recombinació.
A. Bosch^a J. Guimerà^a A. Pereira de Souza^b X. Estivill^a

The EUROGEM Map of Human Chromosome 21

Molecular Genetics Department, Cancer Research Institute, Barcelona, Catalonia, Spain;
 CEPH/Généthon, Paris, France

Human chromosome 21 is the smallest human autosome, representing less than 2% of the human genome. The long arm has been widely studied and its estimated length is about 40 Mb.

The chromosome 21 linkage map has been considerably improved recently using 43 markers, with an average interval of 2.5 cM [54], but still many polymorphic markers have to be developed and added to the genetic map to reach the 1 cM resolution necessary to locate polygenic disorders. Microsatellite sequences are the best markers for building linkage maps as they are highly polymorphic, easy to analyse using PCR and they are uniformly spread throughout the human genome. To improve the linkage map of chromosome 21, and under the EUROGEM Project, 22 CA-repeat markers, with an average heterozygosity of 0.71, were genotyped using the CEPH reference families and a genetic map with 44 highly polymorphic markers was constructed using multilocus linkage analysis with the CRI-MAP program. The framework map was built by first constructing 3 different maps following the same criteria outlined in the construction of the map for chromosome 12 (AB et al., this publication).

Twenty-two index markers were chosen from available genotype databases (CEPH version 6 and CHLC version 2) with an average heterozygosity of 0.79. Together, the markers used in this map have an average heterozygosity of 0.75. In total, 39 dinucleotide repeats, 3 tetranucleotides, 1 VNTR and 1 RFLP marker were used. The linkage map contains 39 uniquely placed loci ordered with odds of at least 1000:1 and any alternative order was excluded by permutations of 4 neighbouring loci. Four additional markers were positioned on the map approximately. The framework map consists of 24 anchor points, 10 of them megaloci, containing from 2 to 4 markers. The sex-average length of the map is 68.9 cM, about 3.5 cM longer than a previous published map [54], due to higher male recombination rate at the telomere (male and female lengths are 57.8 and 80.5 cM, respectively). On the sex-average map the 23 intervals have an average length of 3.0 cM, with the two largest gaps having a length of 5.4 cM (between D21S11 and D21S414, and between D21S65 and D21S167). The largest gap in the previous genetic map was 6.3 cM located above D21S11. This gap has now been filled with 4 new markers (D21S406, D21S364, D21S1266 and D21S409) at distances of 1.4, 0.3, and 0.9 cM respectively. Of the 22 markers genotyped for this project, 18 fit uniquely in the framework map. Among the other 4 markers, D21S1234 was found to be the same marker as D21S415, D21S156 inflated considerably the length of the map, so it was only positioned approximately, as were D21S1419 and D21S1421, which could not be placed uniquely on the map. This work provides 22 markers that have not previously been localised in a linkage map [55]. In the present map, the position of these markers is consistent with their cytogenetic location, and the index markers included also have equivalent relative locations to the previously reported maps [7,2]. These 22 new markers will contribute greatly to the integrated genetic map, and also to the construction of a map, based only on highly polymorphic markers, with the desired 1 cM resolution. The inclusion of these markers in the YAC map [55] provides a more refined location, allowing the integration of the linkage, cytogenetic and physical maps of chromosome 21 [56].



Fig. 21. The EUROGEM Map of Human Chromosome 21. In the inset, IFNAR and D21S259 were not typed within EUROGEM.

4.2.7. Integration of 30 CA-repeat markers into the cytogenetic, genetic and YAC maps of human chromosome 21

Bosch A, Guimera J, Graw S, Gardiner K, Chumakov I, Patterson D and Estivill X

European Journal of Human Genetics, 4: 135-142 (1996)

El doctorand és l'autor del genotipatge dels 532 individus de les famílies de referència CEPH per a cadascun dels microsatèl·lits D21S364, D21S365, D21S367, D21S368, D21S369, D21S406, D21S408, D21S409, D21S414, D21S415, D21S416, D21S1235, D21S1261 i D21S1262, utilitzats en el mapa de lligament genètic. És també corresponsable amb la Dra. A. Bosch de la Fig. 2.

Original Paper

Eur J Hum Genet 1996;4:135-142

Integration of 30 CA-Repeat Markers into the Cytogenetic, Genetic and YAC Maps of Human Chromosome 21

Abstract

The number of polymorphic DNA markers developed for the whole human genome during the last 2 years has been vastly increased. For this reason, the genetic map is continuously improving, but the cytogenetic and physical maps are not progressing at the same speed. Therefore, there is a need to integrate genetic, cytogenetic and physical mapping data. We have developed and localized on the breakpoint map of human chromosome 21 thirty microsatellite markers. Twenty of them have been used in the construction of a genetic map of chromosome 21, which contains a total of 44 markers. This map has 39 uniquely placed loci at 23 anchor points, ordered with odds of at least 1,000:1. The sex average length of the map is 64.4 cM, with the male and female lengths being 49.4 and 79.2 cM, respectively. Twenty-six of these newly developed markers have been localised on the CEPH/Généthon and Joint YAC Screening Effort YACs. Although these microsatellites were found uniformly spread along chromosome 21, the detection of various markers in the same or adjacent YACs suggests that CA-repeat microsatellites are clustered in several regions. The localization of these markers on the cytogenetic, genetic and YAC maps has provided a refined location for them and is a step further towards the construction of an integrated map of HC21.

The usefulness of microsatellite markers in mapping the human genome has been well established since they were first described [1, 2]. Several linkage maps for the whole human genome have been constructed during the last 2 years using these highly polymorphic markers [3–6]. 810 sequence tagged sites (STSs), 50 of which are polymorphic, were used to construct the most complete YAC map of human chromosome 21 (HC21) [7]. This map provides a useful framework for additional STS mapping and

.......................

for gene localization. Despite all these contributions, newly developed DNA markers have been mapped only to the specific type of map for which they were developed. While the genetic map of HC21 is continuously improving due to the high number of polymorphic DNA markers developed for this chromosome [8, 9], the cytogenetic and physical maps are not progressing at the same speed. Therefore, there is a need to integrate the genetic, cytogenetic and physical mapping data for this chromosome.

In preparation for the integration of the different types of HC21 maps, we have localized 30 microsatellite mark-

KARGER E-Mail karger@karger.ch Fax + 41 61 306 12 34 http://www.karger.ch © 1996 S. Karger AG, Basel 1018-4813/96/0043-0135\$10.00/0 Dr. X. Estivill Molecular Genetics Department Hospital Duran i Reynals Avia. Castelldefels km 2.7 E-08907 Barcelona, Catalunya (Spain)

European Journal of Human Genetics

J. Guimerà^a S. Graw^b K. Gardiner^b I. Chumakov^c D. Patterson^b X. Estivill^a

A. Boscha

 ^a Molecular Genetics Department, Hospital Duran i Reynals, Barcelona, Spain;
 ^b Eleanor Roosevelt Institute,

Denver, Colo., USA;

° CEPH/Généthon, Paris, France

Key Words

Chromosome 21 CA repeats Linkage map Cytogenetics map YAC map

Introduction

ers, which we previously mapped in a HC21 somatic cell hybrid panel, to the genetic and YAC maps of this chromosome. Twenty of the most polymorphic markers were included in the linkage map and 26 were localised on the HC21 YAC map, providing a more accurate location of these markers. In addition, we have shown that, even though microsatellite markers seem to be uniformly spread throughout the human genome [10], for HC21 CArepeat microsatellites are clustered in several regions along the chromosome.

Materials and Methods

Human chromosome 21 clones containing CA repeats were isolated from a HC21 specific library (LA21NS01) using a $(GT)_{10}$ oligonucleotide as probe and the regions flanking the microsatellite were characterized [11]. Their HC21 origin was confirmed by amplifying DNA from the cell line WA17, which contains HC21 as the only human chromosome [12].

Mapping on the Breakpoint Map

Each clone was localized to its corresponding HC21 region using PCR to amplify the DNA from a HC21 somatic cell hybrid breakpoint panel, containing 21 intervals spread along both arms of chromosome 21 [9, 11, 13-17]. The panel contains the following cell lines: 153E7b, 2Fur-1, 1x18, JC6-A, ACEM2-10D, 3x2S, R50-3 (6;21), 6918-8a1, MRC2G, GA9-3 (4;21), 9528c (3;21), 1881c-13b (1;21), 8q-a, 21q+, 9542c-5a (10;21), R210W (ring21), Raj5 (21;22) and 643c-13 (7;21) [18, 19]. PCR conditions for each dinucleotide repeat were as previously reported [9, 11, 13-17]. Marker D21S1416 has not been described. Its primer sequences are the following: ABM-C10D1: 5'-CGTGTATGTTTGCAAATATATGT-3'; ABM-C10R1: 5'-ACTAAGCACATTATGTGTGT-3'. As for the rest of the microsatellites, asymmetric PCR was performed in a 9600 Perkin-Elmer thermal cycler, where one of the two primers for each dinucleotide repeat was end-labelled with γ -³²P or γ -³³P[ATP] and its concentration was limiting at 1.5 pmol, with 10 pmol of the unlabelled primer, 1.8 mM MgCl₂, 60 mM KCl, 200 µM each dNTP, 50 ng of human genomic DNA and 0.5 units of Tag DNA polymerase (Boehringer Mannheim) in a 25-µl volume. Initial denaturation was at 92 °C for 5 min. Amplification was for 28 cycles of 95°C for 20 s, annealing at 58°C for 20 s, extension at 74°C for 20 s and final extension at 74°C for 5 min. PCR products between 78 and 99 bp were analyzed on 6% urea-polyacrylamide sequencing gels and autoradiographs were visualized after exposure for 2-16 h.

Linkage Map

The 20 most polymorphic markers were analyzed through the CEPH reference families and the genotypes were introduced into the Sybase application called GENBASE [Lathrop and Sebaoun, unpubl.]. To build the linkage map, data was extracted in LINKAGE format and CRI-MAP input files were produced using the program LINK2CRI [Attwood, unpubl.]. Additional markers were selected from the CEPH version 6 and the CHLC version 2 databases to be used as reference markers in the new linkage map. Using version 2.4 of the CRI-MAP program [20, 21] all the markers were ranked in

order of informativeness and three independent preliminary maps were constructed. The first, starting with the most polymorphic pair, the second starting with the two most centromeric markers and the third starting with the most telomeric pair of markers. All of these starting pairs had an interlocus distance greater than 10 cM. Beginning with these pairs, all the other markers were subsequently added to the map at odds of at least 1,000:1. An initial framework map was obtained with the consensus map resulting from the combination of these three preliminary and concordant maps. This framework map was extended by including all possible remaining loci that could be placed at odds greater than 1,000:1, and, in particular, those markers that have never been included in a linkage map. When two markers showed a recombination fraction of 0.00 for both sexes with a lod score greater than 25, they were considered as part of the same megalocus but distances were only forced to zero when it was known that they were the same marker.

At different stages of the map building, the order was validated by consecutive permutations of 2 adjacent loci with the flips2 option of CRI-MAP, and only orders with a local support exceeding 1,000:1 were retained in the map. Once all possible loci were included in the map, flips4 option was run to discard any alternative order in a four-marker window with a likelihood support of 1,000:1.

When two or more crossovers were found for each individual meioses, using the chrompic option of CRI-MAP, possible genotyping errors were checked by retyping all the individuals involved in the possible recombination event and then rerunning chrompic and rechecking any new crossover. For marker D21S1421, family 102 was discarded as there was one crossover in the paternally contributed chromosome of several of the offspring, suggesting that individual 10201 carried a somatic mutation.

The accuracy of the final map was verified by recalculating the map length when each nonterminal marker was successively removed from the map. Markers considerably increasing the length of the map were rechecked for errors not detected previously by the chrompic option.

Markers which could not be included in the framework map at odds of at least 1,000:1 or which considerably increased the length of the map were placed approximately on the comprehensive map using the all option of CRI-MAP.

Localization of the Markers on the YAC Map

From the greater than 800 YAC clones contained in the published YAC map of HC21 [7], 70 clones were selected spanning the long arm of the chromosome. Each YAC was overlapped by two or three others. The relative locations for most of the YACs chosen had been previously validated [22]. Each YAC was encapsulated in agarose beads using a modification of the method described by Overhausen and Radic [23]. Following the localization of each marker to the breakpoint and genetic maps, YACs covering each related region were amplified using PCR with the primers flanking each CA-repeat.

As we were not able to localize all the markers on the CEPH/ Généthon YAC map, additional YAC clones, belonging to the chromosome 21 Joint YAC Screening Effort (JYSE), were used. Although the JYSE map is still not a continuum map of overlapping YAC clones, many of them have been validated for their relative locations and nonchimerism. Most of the clones contained in this map belong to the previously reported map [7] but several gaps were filled with new YAC clones developed in different laboratories. DNA from 100 individual clones was purified from 40 ml of a saturated culture in complete media (YEPD). After centrifugation, cells were resuspended in 3 ml of 0.9 M Sorbitol, 0.1 M EDTA, pH 7.5, and incubated at 37°C for 60 min with 25 µg/ml lyticase. Spheroplasts were centrifuged and resuspended in 5 ml 50 mM Tris pH 7.4, 20 mM EDTA. After addition of 0.5 ml 10% SDS, the suspension was incubated at 65°C for 30 min. Then, 1.5 ml 5 M potassium acetate was added and the tubes kept on ice for 60 min. After centrifugation, the supernatant was recovered and the DNA precipitated with ethanol. The pellet was resuspended in 3 ml 10 mM Tris pH 7.4, 1 mM EDTA. 0.45 µg/ml RNase was added and incubation performed at 37°C for 30 min. The DNA was extracted with phenol-chloroform, precipitated with isopropanol and resuspended in 10 mM Tris pH 7.4, 1 mM 7.4, 1 mM EDTA at 200 µg/ml.

For both sets of YACs, standard PCR conditions were used to amplify all the markers as follows: 10 pmol of each primer, 2.4 mM MgCl₂, 80 mM KCl, 200 μ M each dNTP, 0.5 μ l of YAC DNA and 2 units of *Taq* DNA polymerase (Boehringer Mannheim) in a 25- μ l volume. Initial denaturation was at 92 °C for 5 min. Amplification was for 35 cycles of 95 °C for 30 s, annealing at 56 °C for 30 s, extension at 74 °C for 30 s and final extension at 74 °C for 5 min in a 480 Perkin-Elmer Thermal-Cycler. PCR products were visualized in 2% agarose gels, stained with ethidium bromide.

Markers localized on more than one map have been checked for concordance and further analysis was performed when necessary.

Results

Thirty CA-repeat markers included in 29 clones, resulting from the screening of a HC21 phage library (LA21NS01) using a $(GT)_{10}$ oligonucleotide as probe, have been localized to a HC21 somatic cell hybrid panel (fig. 1a). Twenty-nine of them have been previously reported [9, 11, 13–17]. Marker D21S1416 has not been previously described. Its heterozygosity is 0.55, calculated using 80 unrelated chromosomes of the CEPH parents. The allele frequencies for D21S1416 are as follows: 78 nucleotides (nt), 0.15; 88 nt, 0.04; 90 nt, 0.75; 96 nt, 0.04; and 99 nt, 0.03.

Linkage Map

The 20 most polymorphic microsatellite markers (with an average heterozygosity of 0.71), analyzed in the CEPH reference families, have been included in a genetic map of the long arm of HC21 containing 44 highly polymorphic markers which are listed in table 1. The framework map contains 39 uniquely placed loci ordered with odds greater than 1,000:1 at 23 anchor points. Eleven of these anchor points are megaloci, containing from 2 to 4 markers. The sex average length of the map is 64.4 cM, with the female and male lengths being 79.2 and 49.4 cM, respectively. Figure 1b shows the sex average linkage map, which contains 22 intervals, with sizes ranging between 0.3 and 5.4 cM, and an average interval length of 2.9 cM. The two largest gaps, both 5.4 cM, are located between markers

D21S11 and D21S414 and between markers D21S65 and D21S167. The comprehensive map contains 5 additional markers placed approximately. The accuracy of the map was assured by regenotyping not only the individuals involved in any crossover event, but also, where possible, the parents and grandparents who contributed the recombinant chromosome, in order to detect any error or de novo mutation. However, the high mutation rate of microsatellite loci in lymphoblastoid cell line derived DNA may not be detected if the new allele follows Mendelian inheritance [24]. The map has a genotyping error rate of 0.1%. calculated from the method described by Lasher et al. [25]. Most of the newly developed markers were previously mapped on the framework of a genetic map which also included D21S112 [9]. The error rate of this previous map was 0.2% and the map lengths to 68.9, 80.5 and 57.8 cM (sex-average, female and male maps, respectively), suggesting that this marker contains some genotyping errors that were not detected by Southern analysis performed by others. For these reasons, we decided to remove D21S112 from the genetic map presented in this work.

The largest gap in previous genetic maps was of about 6 cM, located above D21S11 [6]. This distance has now been shortened with the addition of four new markers in this region (D21S406, D21S364, D21S1266 and D21S409), with a distance between D21S409 and D21S11 of 2.9 cM, and D21S364 and D21S406 mapping in between (fig. 1).

The two most distal markers on the map, D21S369 and D21DS1261, are markers developed in this work. D21S369 was described as the most centromeric marker specific to the long arm of HC21 [7]. However, amplification of this marker shows, in addition to the specific HC21 fragment, a polymorphic band from chromosome 18. Interestingly, on the YAC map, clone 330H1 is positive for this marker, but the amplified band presents the size expected for the chromosome 18 product and not for 21. All the other YACs that were positive for D21S369 had the band chromosome 21 specific band without the chromosome 18 fragment. Therefore, these results suggest that YAC 330H1 is chimeric for both chromosomes and that chimerism must have arisen by homologous recombination.

Among 12, 680 meioses analyzed for the 20 newly positioned markers, we detected 18 de novo mutations, giving an overall mutation rate of 1.4×10^{-3} per locus, per allele, which is consistent with previously published microsatellite data [3, 26]. This relatively high mutation rate is in part due to somatic mutations arising during culture of the lymphoblastoid cell lines [24].

D2151235 D2151224 D21S410 D21S1419 D21S1217 D21S369, D21S1233 D215364, D215364, D215368 D215368 D215367, D2151264 D215415, D2151234 D21S366, D21S1417 D21S1416 D21S1262 D21S1261 D21S416 D21S1421 D21S411 D21S408 D21S365 D21S407 V ۲ V V V Y ۷ ١ V 367B6 751CO1 897C10 D110G11 90B3 259H11 17C12 264P9 8208G3 830D5 52H5 145E3 238B1 746B10 841C4 821H8 124A7 193G12 818A6 D110G6 850C6 265H12 YAC MAP 185F1 937E12 781G5 752D12 745C11 449F6 A22A4 767D6 770B3 759D3 799G3 73F6 986G1 A228B4 814C1 849B9 615C9 552A3 826A8 330H1 243B6 759C5 430G1 757E11 830C7 599G6 232C7 881D2 849B10 c D21S1421 D215259 D21S171 I FNAR ¥ D21S367, D21S1264 D21S210, D21S265, D21S269, D21S365 D21S260 D215198, D2151235, D2151224 D215212 D21S120, D21S236, D21S408 D21S167, D21S267, D21S270 D21582, D2151419 D2151270 D215263 D21S409, D21S1266 D21S415, D21S1234 D21S414, D21S214 D21S368 D21S416, D21S168 D21S366, D21S11 LINKAGE MAP D21S1262 D21S1261 D21S406 D21S369 D21S364 D21S231 D21S65 . ₩ 6.0 1.7 0.9 3.5 +2.1 25 9. 19 0.3 1.4 1.5 2.7 2.4 2.8 4.8 5.4 5.4 4.7 4.1 3.9 4.4 2.2 ۹ D21S414, D21S1415, D21S1263 D21S416, D21S1224, D21S1235 D21S415, D21S1234, D21S1416 D21S367, D21S410, D21S1264 D21S364, D21S366, D21S406, D21S1417, D21S1420, D21S409 D21S369, D21S1233 D21S1419, D21S1421 CA-REPEATS D21S365, D21S368 D21S407, D21S1217 D21S1262 D21S1261 D21S411 D21S408 -⁵ D 0 11.2 11.1 22.2 22.3 1.1 2 11.2 21 22.1 BREAKPOINTS I I 1 21;22 1x18 MRC2G 4;21 10;21 ring 21 3x2S 3;21 1;21 JC6A 3x2S MRC2G 8;21 ACEM2-10D 153E7b ACEM2-10D 6918-8a1 6918-8a1, 6;21



1

Sex average length 64.4 cM

Localization of the Polymorphic Markers on the YAC Map

At least two positive YACs were identified for 26 loci out of the 30 analyzed. Figure 2 shows the YACs that were positive for these 26 markers (marked with an asterisk) and for their flanking STSs. Figure 1c shows the positions on the YACs of the markers analyzed, with their schematic distribution along the HC21. Four markers mapping to the 21q21 region (D21S414, D21S1263, D21S1415 and D21S1420) failed to amplify with all the YACs from the region, indicating that the YAC map has a gap in this band [17].

The correspondence between the positions of the markers located on more than one map is shown globally in figure 1. The three maps agree for all the markers except for D21S368 which is located at the same position on the genetic and the YAC map but proximal on the breakpoint map, as discussed below.

In the YAC map, D21S1261 is located between CD18 and D21S403, which is distal to D21S171 and D21S112. On the contrary, when D21S112 was added to the linkage map, it mapped 4.4 cM distal to D21S1261. As genotyping errors for D21S112 were suspected, we favor the position obtained on the YAC map, which places D21S1261 distal to D21S112. A similar situation than for D21S112 was found for D21S171, which was approximately located near D21S1261 on the genetic map (fig. 1b).

Discussion

Localizing polymorphic markers on the cytogenetic, genetic and YAC maps provides a refined location for these markers, allows for the correction of mapping errors and facilitates the construction of an integrated map of the chromosome. This integrated map will allow the localization of new markers, independently of the method of isolation.
 Table 1. Characteristics of the DNA markers included in the linkage map

Locus	Marker type	Heta	Ip	PK℃	Cytogenetic location
D21S369	DINUCL	0.60	376	189	q11.1
D21S408	DINUCL	0.56	371	194	q11.1-q11.2
D21S120	DINUCL	0.68	446	269	q11
D21S236	DINUCL	0.70	369	240	q11.1
D21S415	DINUCL	0.71	477	254	q11.2-q21
D21S1234	DINUCL	0.71	269	167	q11.2-q21
D21S409	DINUCL	0.71	480	337	q21
D21S1266	DINUCL	0.65	125	86	q21
D21S364	DINUCL	0.68	440	236	q21
D21S406	DINUCL	0.63	412	252	q21
D21S366	DINUCL	0.59	370	251	q21
D21S11	TETNUCL	0.92	538	346	q21
D21S414	DINUCL	0.74	456	326	q21
D21S214	DINUCL	0.84	408	249	q21
D21S368	DINUCL	0.55	359	229	q21-q22.1
D21S367	DINUCL	0.83	543	350	q21
D21S1264	DINUCL	0.81	294	255	q21
D21S210	DINUCL	0.86	312	220	q21
D21S265	DINUCL	0.85	138	124	q21-q22.1
D21S269	DINUCL	0.73	99	85	q21-q22.1
D21S365	DINUCL	0.48	298	74	q21-q22.1
D21S260	DINUCL	0.52	61	61	q22.1
D21S82	RFLP	0.58	367	227	q22.1-qter
D21S1419	DINUCL	0.85	189	154	q22.1
D21S1421	DINUCL	0.77	237	221	q22.1
D21S1270	TETNUCL	0.86	215	184	21q
D21S263	DINUCL	0.74	136	90	q22.1-q22.2
D21S1262	DINUCL	0.78	287	205	q22.1
IFNAR	TETNUCL	0.93	213	131	q22.1
D21S65	DINUCL	0.91	467	307	q11.2-q22.2
D21S167	DINUCL	0.80	269	197	q22.2-qter
D21S267	DINUCL	0.88	137	125	q22.1-q22.3
D21S270	DINUCL	0.85	154	106	q22.1-q22.3
D21S259	DINUCL	0.79	122	90	q22.3
D21S156	DINUCL	0.84	543	363	q22.3
D21S416	DINUCL	0.79	510	314	q22.3
D21S168	DINUCL	0.77	486	282	q22.3
D21S231	DINUCL	0.65	332	90	q22.3
D21S1235	DINUCL	0.76	281	222	q22.3
D21S1224	DINUCL	0.83	193	139	q22.3
D21S198	DINCUL	0.81	326	245	q22.3
D21S212	DINUCL	0.86	559	336	q22.3
D21S1261	DINUCL	0.83	294	258	q22.3-qter
D21S171	DINUCL	0.69	430	329	a22.3-ater

^a Observed heterozygosity in 80 nonrelated chromosomes of the CEPH families.

^b Informative meioses, as calculated by CRI-MAP.

 Informative meioses, phase known, as calculated by CRI-MAP. In bold are the markers included in other HC21 map in this work.

Fig. 1. Integration of CA-repeat markers on human chromosome 21. a Idiogram of HC21. Indicated on the left are the breakpoints defined by the somatic cell hybrid panel used in this analysis. On the right, the positions of the markers in the intervals of the HC21 breakpoint map are shown. b Sex-average linkage map of HC21 constructed with 44 markers. Genetic distances between markers (in cM) are indicated on the left. The sex-average length of the map is 64.4 cM. The framework map contains 39 uniquely placed loci at 23 anchor points. Five additional markers are positioned approximately (on the right). c Positions of 26 CA-repeat markers on the YAC map of HC21. On the left is the schematic representation of the YACs along the chromosome. On the right, the location of the markers in each YAC is shown by arrows. When several markers map to the same YAC or YACs, a line is used to show their positions.

Origin	CEPH	H I		Hand	CEPH	CEPH	CEPH	CEPH	CEPH	E	E		CEPH	CEPH	E S	E	Had	E Haso	CEPH	JYSE	JYSE	CEPH	E	CEPH	E B	EPH	EPH CEPH		- HSE	H	E B	5		CEPH	CEPH	CEPH	CEPH	CEPH	CEPH	CEPH CEPH	CEPH	CEPH	E	CEPH	CEPH CEPH			E H	JSL JSL	JSK.	EPH -	Hda	JYSE	CEPH	CEDI
YAC Name	52H5	78165	77/03	330H1	849810	185F1	897C10	84989	124A7	59966	75903	145E3	881D2	818A6	615C9	449F6	32805	746810	9083	986G1	937E12	264F9	193G12	160F2	759C5	745C11	79963	D110G11	D11066	830D5	430675	752012 841C4	4114	17012	23207	814C1	85006	751C01	103C4	37488	826A8	830C7	14A12	787C6	767D6	82H18	265H12	552A3	820863	AZCA4	ACOUL	34/15	A228R4	36786	20100
DNA Marker	÷	+	+	+	+	⊢	┢	+		H	+	╋	╀	Н	+	+	+	+	╋	╋	⊢	\vdash	-	Н	+	+	H	+	+	+	+	+	+	+	+	┝	-	Н	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	┢	$^+$
D215411 *	÷	+	+	+	+	t	t			H	+	+	+	H	+	-	+	+	╈	t	t		-	Η	+	+	H	+	+	+	+	+	+	╈	$^{+}$	+			-		-	+	+	+	+	+	+	+	$^{+}$	+	$^{+}$	$^{+}$	$^{+}$	t	t
D215215	H	1	1	+ +	+	+	t	T		H	+	+	t	H	1	+	+	+	╈	t	t			Η	+	1	H	1	+	1	+	+	$^{+}$	$^{+}$	t	t					-	+	+	+	1	+	$^{+}$	+	$^{+}$	$^{+}$	t	$^{+}$	t	T	t
D215369 *	Π		ŀ	+ +	+	+	+					1	T								T									,					T	Γ							T	T			T		T	T	T	T	T	T	T
D2151233*			ŀ	+ +	• +	+	+						Γ	Π			Τ		Τ		Γ						Π					T	Т										Τ								Ι		L		Τ
D21S308			ŀ	+ +	+	+	+												Ι	Γ	L													Τ	Г								Τ				Ι			T	I		Γ		Ι
D21S318				+	+	+		+	+																						_					L																			
D215408 *		-	+	۲		L	+	+	+	+	_	-													_	_		-		-	1	-		1						_		_	\downarrow	4	4	1	4	-	1	_	+	+	\downarrow	\perp	1
D21S120	\square	-	+	-	+	1	+	+	+	+	+			\square	4	4	4	1		+	1		_	\square	-	4	\square	+	_	-	+	+	+	+		┢				_	_	-	4	-	4	+	4	+	+	+	∔	+	+	₽	+
D2153/3	Н	+	+	+	⊢	⊢	⊢	+	Н	Н	*	+	+	-	+	-	+	+	╋	╇	╋	-	-		+	+	+	+	+	+	+	+	+	╉	+	┢	-	\square		-	-	+	+	+	+	+	╉	+	╋	╋	+	+	╀	╋	╀
D2151416*	Н	+	+	+	⊢	⊢	┝	H	\square	+	+	+++	+	+	+	+	+	+	╀	╀	╀	H	-		+	-	-+	+	+	+	+	+	+	+	+	┝	\vdash		-	-	-	+	+	+	+	+	+	+	┿	+	+	+	+	+	╀
0215372	Н	+	╉	+	+	⊢	┝		-	+	-ŀ	-	Ť	1 T	井	+	+	+	╋	╀	╀	Н	-	Н	+	+	H	+	+	+	+	+	+	+	+	┝	-	-	-	-	-	+	+	+	+	╉	+	+	+	+	+	+	+	╋	t
D215415 *	H	+	+	+	+	\vdash	┢	Н	Η	H	+	+	Ť	+	÷	+	+	+	+	+	+	+			+	+	H	+	+	+	+	+	+	+	+	+	\vdash	Η	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	$^{+}$
D2151234*	H	+	+	+	+	t	+	H	H	H	+	+	+	÷	÷1	+	+	+	+	+	+			Η	+	+	H	+	+	+	+	+	+	+	$^{+}$	\mathbf{t}	+	H	\square	+	+	+	+	+	+	+	+	+	+	$^{+}$	$^{+}$	+	$^{+}$	t	t
D21S172	H		1	1	T	F	T		Н	H	+	+	+	+	+	+	+	+	t	t	t			Η	+	1		+	\uparrow			+	+	\dagger	t	t		Η		+	1	+	+	+	+	1	+	+	+	+	t	+	t	t	t
D21S382	Π				Γ							T	+	+	+		+	+	T	T														T	T	Γ							1	1			T	T	T	T	T	T	T	Γ	t
D215366 *		T	T	T	L	Г	Γ					T	L				ŀ	+ +	- +	-								T			I	T		T	L	Γ							I		T		T	I	T	T	T	T	T	L	I
D21S1417*	1	-	1	1	1					H	1	1		Ц	1		-	+ +	+					Ц	1	1	\square	-	1			1	1	1						1	1		1	1	1	1	4	4	1	+	+	+	+	1	1
D21S59	\square	+	+	+	+	1	1			\square	_	+		\square	4	4	-	+ +	-	1	1				4	4		-	-	-	+	+	+	+	-	1				4	_	\downarrow	4	+	+	+	+	+	+	+	∔	+	+	╞	∔
D21S385	\square	+	+	+	⊢	⊢	+	-		\vdash	-	+		\square	4	4	-	+	+	+	+		_		-	-	-+	+	+	4	+	+	+	+	1	1	-		_	-	-	+	+	+	+	+	+	+	╇	+	+	+	╞	╞	ł
D215409 *	H	+	+	+	⊢	⊢	╞	-	H	\vdash	+	+	H	H	4	+	+	+	╀	+	+	+		-	+	-	-	+	+	+	+	+	╋	+	+	╞	-	-	-	-+	-	+	+	+	+	+	+	+	╇	╇	╋	+	┿	┝	╀
D215364 *	H	+	+	╈	╋	⊢	┝	H	\vdash	\vdash	+	+	H	+ +	+	+	+	+	╋	ť	+	+	H	Н	+	+	+	+	+	+	+	+	+	+	╋	⊢	-		-	-	-	+	+	+	+	+	+	+	+	+	+	+	╈	╋	╀
021518	H	+	+	+	+	⊢	+	-	H	+	+	+	+	+	+	+	+	╈	╋	+	f	Ŧ	+	Н	+	+	+	+	+	+	+	+	+	+	+	⊢		-	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	$^+$
D215222	H	+	+	+	+	⊢	t		Н	H	+	+	t	+	+	+	+	+	t	╀	┢		+	Н	+	+	+	+	+	+	+	+	+	+	+	┝	+	Η		-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	t
D215368 *	H	+	+	+	\mathbf{t}	t			Η	H	+	+	+	H	+	+	+	+	t	t	+	H	+	Η	1	1		+	+	+	+	+	+	+	t	t	t			1	1	+	+	+	+	+	$^{+}$	+	+	+	t	$^{+}$	+	+	t
D21S282			1	t	\mathbf{T}	t				H		+	t	Ħ	1	1	1	1	t	t	t		+	+	+	1		1	+	+	1	+	+	t	T	T						1	+	+		1	\dagger	+	+	+	t	T	t	t	t
D21S367 *											T		T	T			1		T	T	T				+	+				T	T		T	T	T	T							T	T	T	1	T	T	T	T	T	T	T	T	T
D21S1264*												T		Π											+	+																	T				T	T	T	T	T	T	Τ	T	T
D215234		1	1								-														+	+	+																									1			I
D21S287		+	+	+		1					-		_	Ц	_	-	_	-	1		1				-	+	+	+	-	4	4	-		1		L					_	\downarrow	\downarrow	_	_	\downarrow	4	-	+	+	+	\downarrow	1	\perp	1
D215365 *	\square	+	+	+	┢	┡	-			4	+	╇		Ц	4	-	-	+	+	+	1		_		-	+	+	+	+	+	+	+	+	+	+	╞				_	4	+	4	4	4	-	4	4	+	+	+	+	∔	╞	∔
0215289	\square	+	╉	+	+	┢	+-	H	Н	+	+	+	-	\square	-	-	+	+	╀	╀	┢	-	-	-	+	-	+	+	+	+	+	+	+	╇	+	┝	-	\vdash		+	-	+	+	+	+	+	+	+	+	╇	╋	+	╋	+	╀
D215410 *	H	+	┿	┿	┢	⊢	┝	Н	Н	H	+	┢	-	H	+	+	+	+	╀	┝	⊢	\vdash		Н	+	+	+	+	+	+	+	+	+	╋	+	┝	-	Н	-	-	-	+	+	+	+	╉	+	+	+	+	╋	+	╋	╋	ł
D215218	H	+	+	╈	┢─	⊢	┢	Н	Н	H	+	┢	H	H	+	+	+	+	+	+	┢	\vdash	-	Н	+	+	H	+	-	-	+	+	+	+	+	┢	+	Н	-	+	-	+	+	+	+	+	+	+	+	╋	+	+	╈	╋	t
D215294	H	+	+	╈	+	⊢	┢			H	+	+	t	H	┥	+	+	+	+	+	┢	\vdash		Н	+	+	H	+		+	+	+	+	+	+	┢	+	Н		+	1	+	+	+	+	+	+	+	+	+	+	+	+	+	t
D2151419*	H	+	$^{+}$	+	+	F	t	Н			+	+		H	+	+	+	+	t	t	t			Н	+	+	H	+			+	+	$^{+}$	+	t	t	t					+	+	+	+	+	+	+	$^{+}$	$^{+}$	$^{+}$	+	$^{+}$	t	t
D21S296		1			\top	Г				H	1		Г	H	1	1	1	1		t	t				1	1				1		+ +	+ 1 +	+	T	T	T						1	1	1	+	+		+	+	t	+	t	t	t
D21582		T	Τ	T										П					T	T	Γ											+ +	+	1	+	+							T	T			T	T	T	T	T	T	T		T
D2151421*												Γ								L	Γ										Ţ	+ +	+	1	+	+						_	\Box				I		T	T	T	T	Τ	L	I
D21S388	Ц	_									_	-			_	4									-			1	1		_			1	+	+						-			_					1	1	\perp	1		1
D21593	\square	-	+	+	+	⊢	-	1	4	Н	+	+	-	\square	-	-	+	-	+	╇	⊢		-	Ц	+	-	-	+	+	4	4	+	+	+	+	1	+			_	_	_	4	4	4	+	4	+	+	+	+	+	+	╞	+
02151262*	\square	+	+	+	⊢	⊢	-	4	Н	\vdash	+	+	╞	\square	4	-	+	+	╀	∔	╞			Н	+	-	\vdash	+	+	+	+	+	+	+	┢	╞	+	+	_	_	-	-	4	+	+	+	4	+	+	+	+	+	+	+	╀
D216241	H	+	+	+	┢	⊢	+	-	-	H	+	+	╀	H	+	-	+	+	╀	+	⊢	\vdash	-	Н	+	+	H	+	+	+	+	+	+	+	╋	┝	+	+	+	-		+	+	+	+	+	+	+	+	+	+	+	+	╋	+
D2151217*	H	+	+	+	┢	⊢	┢	H	-	H	+	+	+	H	+	+	+	+	╋	╀	┢	-	-	Н	+	+	H	+	+	+	+	+	╈	+	╋	⊢	+	Н	Η	+	-	-	+	+	+	+	+	+	+	+	+	+	+	┢	$^+$
ETS2	H	+	+	+	┢	⊢	t	H		H	+	╈	t	H	+	+	+	+	+	+	⊢	\vdash	-	Н	+	+	H	+	+	+	+	+	+	+	╈	⊢	⊢	Н	Η	+	Ť	-	#	+	+	+	+	+	+	+	+	+	+	+	t
D215346	H	+	+	+	+	t	t		-	H	+	+	t	H	+	+	+	+	╈	+	t	H		Η	+	1	H	+	+	+	+	+	$^{+}$	$^{+}$	t	t	t				-	+	+	Ŧ	ŧ	+	+	+	+	+	$^{+}$	$^{+}$	$^+$	$^{+}$	t
D215416 *	П	1	T	T	T	Γ				H	1	t				1	1		T	t	t					1		Ť	1	1	1	+	t	$^{+}$	t	t						+	+	+	+	+	+	+	$^{+}$	$^{+}$	t	t	t	T	t
D21S397	Π	Т	Τ	T	T								Γ	Π	Т	Т	T		T	T	Г			Π			П		T	T	T		T	T	Т	T						T	T	T		+	T	T	T	T	T	T	T	T	T
D21S354		T	Т	T	Γ	Γ	Г				T	T	Г	П			T	T	T	T	Г								1	1		T	T	T	T	T							T	T	1	1	+	T	T	T	T	T	T	Т	T
D21S1235*	Π	T	Т	T	Γ	Γ					T	Т	Г	П	1		T	T	T	T	T			Π					T			T	T	T	T	Т				1			1	1	1	1	+	+	T	T	T	T	T	Т	T
D21S53	Π							1						Π			T	T	T	Т	Γ							T	1			T	T	T	T	Г	T					1	1	1	1		+	+	+	T	T	T	T	T	T
D21S64		T									T	T	T	Π			T		T	T									T	T	T	T	T	T	Т	Т							T	T			T	+	+	+	T	T	T	T	T
D2151224*	Π	T	T	T		Γ							Γ	Π			T	T	Т	Т	Г										T	T	T	T	T	T							T	T			T	1.	+	+	T	Ţ	T	T	T
MX2	Π	T	T	Γ	Γ	Γ					T	Γ	Γ	Π	T		T	T	T	Γ												T		T	T	Γ	Γ						T				T	T	Ţ	+	T	T	T	Γ	T
PFKL	Π	T	T											Π			T	T	Γ	Г								T	T	T	T	T	T	T	T	Γ						T	T	T			T	T	T	1.	٠T	T	T	T	T
D21S112		T	T																	Γ								T	1				T	T	T								T	T			T	T	T	T	T	+	T	T	T
D21S171		T	T									I	Γ					T		T					1							T	T	T	T								T				T	T	T	T	T	ŀ	+	T	Τ
CD18	П	I	T	T								T								Γ									T		T	T	T	T													Ι	I	T	T	T	ŀ	+ +	+ +	·
D2151261*	\square	-	1	1		L				Ц	1			Ц	1			T		L									T	T	T			T	F						_	1	1	1		T	1	T			1	1	1.	++	-
0215403					1	1						1	1						1						1								1	1	1								1							1	1		1		1

2

The inclusion in the genetic and YAC maps of the 30 microsatellites reported here allowed us to detect errors in the breakpoint map locations for some of these markers. The position of D21S409, previously placed between breakpoints MRC2G and 4;21 [11], has been relocated to between the 4;21 and 3;21 breakpoints, the interval immediately distal. D21S416 maps to band 21g22.3, in the same breakpoint interval as D21S1235, but by its location on the YAC map, it is centromeric to D21S1235, their initial relative locations being reversed [13]. D21S366 was mapped between breakpoints 6918-8a1 and MRC2G [11], but data from both the genetic and the YAC maps have localized it proximal to D21D364. By designing additional primers we localized D21S366 between breakpoints 3;21 and 4;21, which agrees with the genetic and YAC maps. For one marker only, D21S368, we have not been able to integrate data from the three maps. On the genetic map, it is placed 4.7 cM more centromeric to D21S367, which is consistent with its location on the YAC map, but on the cytogenetic map it is located in a more distal position, between breakpoints ACEM and JC6 (fig. 1). As we have consistent results for the genetic and YAC maps we may conclude that the real position of D21S368 is proximal to D21S367. The nonconcordant breakpoint result may be due to sequence homology between the D21S368 primers and other chromosomes contained in some of the cell lines in the breakpoint panel used to characterize the cytogenetic location of this marker. However, we cannot exclude the possibility of small rearrangements in the cell lines during the construction of some of the hybrids.

Microsatellite markers have been described as being distributed randomly throughout the human genome [1]. As the markers analyzed in this work have been isolated from a chromosome 21 library, we would expect to find CA-repeat containing clones uniformly distributed along the chromosome. We localized CA-repeat markers to 13 of the 21 intervals contained in the somatic cell hybrid map, spanning from centromere to telomere. The interval between breakpoints 3;21 and 4;21, a region spanning about 3.5 Mb, which represents about 6% of the HC21, contains 20% of the clones isolated. So, the density of CA- repeats in this region may be significantly higher than for other regions of HC21. Surprisingly, when the polymorphic markers were mapped on the YACs, most of those in the same breakpoint interval were localized to the same YAC and very often between the same STSs. This is the case for D21S369 and D21S1233; D21S366 and D21S1417; D21S364, D21S409 and D21S406; and D21S367 and D21S1264 (fig. 1c). Since other microsatellites also mapped to contiguous YACs, most of the markers developed from this HC21 phage library are located in clusters, although these clusters seem to be uniformly distributed along the chromosome. The four markers localized to 21q21, mainly between breakpoints 3:21 and 1:21 - which failed to amplify YAC DNA - may also be clustered, suggesting that the gap in this region is not necessarily large and it could be filled with a few YACs. Further evidence for CA-repeat clustering was found when subcloning YACs into cosmids, with few subclones having CA-repeats, but those which did contained many repetitive nonadjacent blocks (data not shown). In addition, the high number of megaloci in the genetic map (11 of 23) may also support this hypothesis, as several markers are placed in the same locus.

The average observed heterozygosity for 27 of the 30 markers analyzed in this work was of 0.66, showing that the method used for their isolation was stringent enough to obtain useful polymorphic markers [11]. Furthermore, localizing these markers on the YAC map and, more precisely, between two adjacent STSs whose positions were previously reported [7] has provided more accurate location for these markers and makes them more useful for mapping studies. D21S1262 is located next to the SOD1 gene, and should therefore be useful for genetic analysis of families affected by familial amyotrophic lateral sclerosis [27]. Markers that could not be positioned on the YAC map should be useful for isolating new YACs to fill the existing gaps and to complete the continuum map of overlapping YAC clones on HC21. Finally, localizing markers in two or three of the maps as reported in this work is a step further towards the construction of an integrated map of HC21.

Acknowledgements

We thank J. Attwood for help with the linkage map and M. Pritchard for useful comments. This work was supported by the Fondo de Investigaciones Sanitarias de la Seguridad Social (94/1905E), Institut Català de la Salut, European Community Grants CEC/ BIOMED1 PL93-0037 and PL93-0107, by NICHHD (HD17449) and NCHGR (HG00716) and by the Fundació Catalana Síndrome de Down/Marató TV3.

Fig. 2. Distribution of the newly characterized markers on the YAC map. In the left column are the 26 markers positioned on the HC21 YACs in this work (with an asterisk), as well as their flanking STSs, previously located on the YAC map. Along the top are the YACs from the CEPH/Généthon and JYSE that were positive for the STS markers. Markers present in a clone are indicated with a +. Both, STSs and YACs, are ordered from centromere to telomere.

References

- Weber JL, May PE: Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 1989;44:388–396.
- 2 Litt M, Luty JA: A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within de cardiac muscle actin gene. Am J Hum Genet 1989;44:397–401.
- 3 Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau, Vaysseix G Lathrop M: A second-generation linkage map of the human genome. Nature 1992;359:794-801.
- 4 Buetow KH, Weber J, Ludwigsen S, Scherpbier-Heddema T, Duyk GM, Sheffield VC, Wang Z, Murray JC: Integrated human genome-wide maps constructed using the CEPH reference panel. Nature Genet 1994;6:391– 393.
- 5 Gyapay G, Morisette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J: The 1993-94 Généthon human genetic linkage map. Nature Genet 1994;7:2363-2367.
- 6 Matise TC, Perlin M, Chakravarti A: Automated construction of genetic linkage maps using an expert system (MultiMap): A human genome linkage map. Nature Genet 1994;6: 384–390.
- 7 Chumakov I, Rigault P, Guillou S, Ougen P, Billaut A, Guasconi G, Gervy P, LeGall I, Soularue P, Grinas L, Bougueleret L, Bellanné-Chantelot C, Lacroix B, Barillot E, Gesnouin P, Pook S, Vaysseix G, Frelat G, Schmitz A, Sambucy JL, Bosch A, Estivill X, Weissenbach J, Vignal A, Riethman H, Cox D, Patterson D, Gardiner K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Le Paslier D, Heilig R, Antonarakis S, Cohen D: Continuum of overlapping clones spanning the entire human chromosome 21q. Nature 1992;359:380–387.
- 8 McInnis MG, Chakravarti A, Blaschak J, Petersen MB, Sharma V, Avramopoulos D, Blouin JL, König U, Brahe C, Matise TC, Warren A, Talbot CC, Van Broeckhoven C, Litt M, Antonarakis SE: A linkage map of human chromosome 21: 43 PCR markers at average intervals of 2.5 cM. Genomics 1993;16:562-571.

9 Bosch A, Guimerà J, Pereira de Souza A, Estivill X: The EUROGEM genetic map of human chromosome 21. Eur J Hum Genet 1994;2: 244-245.

......

- Weber JL: Informativeness of human (dCdA)_n · (dG-dT)_n polymorphisms. Genomics 1990;7:524–530.
- 11 Bosch A, Nunes V, Patterson D, Estivill X: Isolation and characterization of 14 CA-repeat microsatellites from human chromosome 21. Genomics 1993;18:151–155.
- 12 Raziuddin A, Sarkar FH, Dutkowski R, Shulman L, Ruddle FH, Gupta SL: Receptors for human alpha and beta interferon but not gamma interferon are specified on human chromosome 21. Proc Natl Acad Sci USA 1984;81: 5504–5508.
- 13 Bosch A, Wiemann S, Guimerà J, Ansorge W, Patterson D, Estivill X: Two dinucleotide repeat polymorphisms at 21q22.3 (D21S416 and D21S1235). Hum Molec Genet 1993;2:1744.
- 14 Bosch A, Wiemann S, Ansorge W, Patterson D, Estivill X: Three CA/GT repeat polymorphisms from loci D21S414 and D21S1234 on human chromosome 21. Human Genet 1994; 93:359–360.
- 15 Bosch A, Guimerà J, Wiemann S, Ansorge W, Patterson D, Estivill X: Identification of two highly polymorphic CA-repeats (D21S1224 and D21S1261) on human chromosome 21q22.3. Human Genet 1995;95:367–369.
- 16 Bosch A, Wiemann S, Guimerà J, Ansorge W, Patterson D, Estivill X: Five new microsatellite polymorphisms at the q21 region of human chromosome 21. Human Genet 1995;95:119– 122.
- 17 Bosch A, Guimerà J, Patterson D, Estivill X: Characterisation of three microsatellite polymorphisms (D21S1262, D21S1419 and D21S1421) at the 21q22.1 band. Human Genet 1995;95:596-598.
- 18 Gardiner K, Watkins P, Münke M, Drabkin H, Hones C, Patterson D: Partial physical map of human chromosome 21. Somat Cell Mol Genet 1988;14:623–638.

- 19 Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S, Patterson D: Analysis of human chromosome 21: Correlation of physical and cytogenetic maps, gene and CpG island distributions. EMBO J 1990;9:25–34.
- 20 Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, Bowden DW, Smith DR, Lander ES, et al: A genetic linkage map of the human genome. Cell 1987;51:319-337.
- 21 Lander ES, Green P: Cosntruction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 1987;84:2363–2367.
- 22 Delabar JM, Créau N, Sinet PM, Ritter O, Antonarakis SE, Burmeister M, Chakravarti A, Nizetic D, Ohki M, Patterson D, Petersen MB, Reeves RH, Van Broeckhoven C: Report of the fourth international workshop on human chromosome 21. Genomics 1993;18:735–745.
- 23 Overhauser J, Radic MZ: Encapsulation of cells in agarose beads for use with pulsed-field gel electrophoresis. Focus 1987;9:8–9.
- 24 Banchs I, Bosch A, Guimerà J, Lázaro C, Puig A, Estivill X: New alleles at microsatellite loci in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell lines. Human Mutation 1994;3:365–372.
- 25 Lasher L, Reffer J, Chakravarti A: Effects of genotyping errors on the estimation of chromosome map length. Am J Hum Genet 1991;49 (suppl): abstr 2076.
- 26 Bowcock A, Osborne-Lawrence S, Barnes R, Chakravarti A, Washington S, Dun C: Microsatellite polymorphism linkage map of human chromosome 13q. Genomics 1993;15:376– 386.
- 27 Rosen DR, Siddique T, Patterson D, Fliglewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX, Rhamani Z, Krizus A, McKenna-Yasek K, Cayabyab A, Gaston SM, Berger R, Tanzi RE, Halperin JJ, Herzfeldt B, Van den Bergh R, Hung WY, Bird T, Deng G, Mulder DW, Smyth C, Laing NG, Soriano E, Pericak-Vance MA, Haines J, Rouleau GA, Gusella JS, Horvitz HR, Brown RH: Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 1993;362:59–62.

•

•1

Per tal d'identificar gens involucrats en el desenvolupament del cervell o en la seva funció, i possiblement implicats en la discapacitat intel·lectual observada en la DS, s'ha construit una genoteca de 502 cosmidis a partir de tres regions de l'HC21, que inclou la DSCR. L'aplicació de la tècnica de la selecció de cDNA a partir d'aquests cosmidis, i emprant una genoteca de cDNA de cervell fetal humà, ha permés l'aïllament i identificació de 120 cDNA parcials o EST, que han estat sequenciats i localitzats a l'HC21 (Taula 6). Entre aquests 120 cDNA, 45 són no redundants, dels quals 37 són EST inèdits. S'ha comprovat l'expressió de cadascún d'ells mitjançant transferències de Northern i anàlisis de RT-PCR amb encebadors específics. Cada EST s'ha localitzat en mapes físics de cosmidis, que integren tota la informació funcional i física en un sol mapa transcripcional (Fig. 12). La Fig. 13 resumeix el disseny de l'estrategia que s'ha seguit i que s'ha explicat en l'apartat de metodologia. La comparació de la sequència de DNA de cada EST amb les sequències del banc de dades, mostra que el 37% dels EST aïllats tenen similituds significatives, o sia $p(N) > 10^{-4}$ (Taula 6). Aquests resultats estan en consonància amb altres estudis realitzats en aquesta regió (Patil et al., 1994; Lucente et al., 1995). Entre les sequències d'aquests EST inèdits, n'hi havia una (la sequència del clon D1-34) que compartia una elevada similitud amb la sequència del gen mnb de Drosophila.

La completa caracterització del cDNA de l'EST D1-34 va resultar en la identificació d'un nou gen humà, que per la seva altíssima similitud a nivell nucleotídic i protèic amb el gen mnb de Drosophila l'hem anomenat MNBH (Minibrain homolog). El gen MNBH és un gen de còpia única (Guimera et al., 1997b) que es localitza en la DSCR de l'HC21, entre els marcadors D21S335 i D21S337, a 186 kb del marcador de referència D21S55 en direcció centromèrica. MNBH està distribuit en 17 exons i cobreix una regió genòmica de 150 kb. Segons quin sigui l'inici de transcripció hi ha dues variants transcripcionals: 1) MNBHa, té l'inici de la transcripció a l'exó 2 i la seva expressió pot estar controlada per una illa CpG, i 2) MNBHb, té l'inici de transcripció a l'exó 1, i la seva expressió pot estar sota la influència d'una caixa TATA-like.

Diferents splicings alternatius de la regió codificant de MNBH donen lloc a 4 variants proteiques, anomenades MNBH-iso1, MNBH-iso2, MNBH-iso3 i MNBH-iso4. De totes elles, MNBH-iso1 és la majoritària, i la resta de variants aïllades són espècies de RNA molt poc representades. Els dos senyals de translocació nuclear, la cremallera de leucines i els vuit subdominis del nucli catalític de la quinasa serina-treonina estan presents en les quatre variants de la proteïna, mentre que el domini PEST es troba

236 Resum i discussió –

TAULA 6 Resum dels resultats de mapatge i recerca de similituds als bancs de dades dels cDNA parcials aïllats

Nom del	Nom del Nombre YA de clons indep.		Cosmids	BLASTN		BLASTX						
	macpi			Similitud	P(N)	Simillitud	P(N)					
C2-31	1	72h9	340-345-361-387-398	NONE	-	NONE						
A7-34	1	72h9	340-346-361-363-387-398	Human DSCR1 (3'UTR) gb:U28833	1.7e-98	NONE	-					
A12-31	4	72h9 ·	340-346-361-363-387-398	Human DSCR1 gb:U28833	6.4e-116	Human DSCR1 sp:P53805	1.1e-55					
C10-35						,						
E6-32												
F4-34												
G3-34	1	72h9	340-346-361-387-398	Human DSCR1 (3'UTR) gb:U28833	7.8e-98	NONE	-					
A1-36	17	72h9	340-346-361-363-387	NONE	-	NONE	-					
A3-34												
A3-33												
B6-31			4									
B10-32												
C7-36												
E8-36												
E10-34												
F4-32												
F5-31												
G5-31												
G6-36												
H7-31												
H7-34												
H12-34												
D8-31	1	72h9	340-346-361-363-387-398	NONE	141	NONE						
F6-35	1	336G11	48-66-67-87-94	Human EST yc08c11.s1 gb:T63265	3.6e-16	NONE	-					
F11-36	1	336G11	48-66-67-81-87-94	NONE	-	NONE	-					
C4-32	2	336G11	48-66-67	Human EST HUM21ES111 gb:L25473	2.8e-22	NONE	-					
B3-31												
A5-31	1	336G11	93	NONE	•	NONE						
C7-32	1	336G11	93	NONE	-	NONE						
D5-35	1	336G11	45-57-60	NONE	•	NONE						
D1-34	1	336G11	45-60-76-93	Human MNB gb:U52373	1.2e-62	Human MNB gb:U52373	1.3e-30					
D8-34	11	336G11-	49-57-50-99	Human MNB gb:U52373	1.2e-71	Human MNB gb:U52373	8.4e-45					
A3-33												
B3-34 B5 31												
C6-33												
D7-35												
D11-31												
E5-34												
G5-34												
G11-33												
H6-32		02071	10 42 44									
	-	23881	10-43-44	Ilumon MND - L-1160070			20.01					
D2-34	.3	330011-	49-57-00-99	numan MINB gb:U52373	1.5e-59	Human MNB gb:U52373	3.0e-31					
E7-31		238B1	10-43-44									
B11-34	2	336G11-	45-49-57-60	NONE		NONE	-					
G8-35		encou=1071651762		na na serie de la constanción de			2 C 2					
		238B1	10-43-44									
A4-34	4	336G11-	49-57-60-78-99	Human STS SHGC-5680 gb:G07613	2.2e-21	NONE	-					
G9-36												
G10-36												
H1-33		238P1	10-22-37-43-44									
A10-32	2	336G11-	49-57-58-78-85-96-99	Human hereditary haemochromatosie	7 0=-07	NONE						
H7-33				in the second providence of the second se								
		238B1	10-37-43-44	region gb:U91328								
C10-32	1	336G11-	95	NONE	-	NONE	-					
		238B1	30-41									

Resum i discussió 237

G2-31 H12-33	2	336G11- 238B1	95 - 30-41	NONE	-	NONE	
B3-33	1	336G11-	95	Human EST HSC3OC101 embl:F11685	3.8e-66	NONE	-
C5-32	1	238B1 336G11-	79	NONE		S. cerevisae Centractin sp:P38696	1.3e-05
A4-33	1	238B1 336G11- 238B1	79	Human EST yb67h06.r1 gb:T59992	1.5e-24	Lupinus luteus Early nodulin 75 protein	8.3e-11
F2-36	1	238B1	1-7-15-24-26-29-31-33-40	NONE	-	Bovine herpesvirus Capsid protein P40	1.2e-04
D2-36	2	238B1	1-7-15-24-26-29-31	NONE	-	NONE	-
E7-34	1	238B1	7-24-31-38	NONE	-	Human ZINC FINGER Prot. gp:001101	1.6e-04
D9-32	1	238B1	12-20	Human DNA sequence from cosmid N117B5 gb:Z75744	2.2e-08	NONE	-
D6-36	1	238B1	13-19	NONE	-	NONE	-
H11-32	1	238B1	6-8-14-27-28	Human EST gb:W27025	4.8e-20	NONE	
E12-34	1	238B1	6-8-14-27-28	NONE		NONE	-
F11-34	1	238B1	6-8-14-27-28	Human EST HSCZTB032 gb:F04620	2.0e-05	NONE	-
H3-36	5	238B1	6-8-14-27-28	NONE		NONE	-
B1-36							
B8-36	-						
F9-34							
G6-33	3	238B1	6-8-14-27-28	NONE	-	NONE	-
F8-35							
H10-36							
E3-32	1	238B1	6-8-13-14-19	Human chromosome BAC clone gb:U95738	4.2e-06	NONE	-
B7-32	12	238B1	6-8-13-14-19	Human cosmid 26hf gb:U09822	5.7e-08	NONE	-
A2-36							
A5-33							
A11-36							
C7-31							
C7-33		-					
D2-31	-						
D7-34							
D9-33							
D12-32				2			
E3-35						at at a firm	
H5-34	5	238B1	6-8-13-14-19	NONE	-	NONE	-
A11-32	1						
E3-33							
H5-34	1						
H11-33							
B7-36	2	238B1	6-8-13-14-19	NONE	-	NONE	-
F3-35							
A7-31	5	238B1	6-8-13-19	NONE	-	NONE	•
A9-34							
B9-31							
G6-32							
E12-35	5	238B1	13-19	Human inward rectifying potassium	2.9e-49	NONE	-
C4-33	P			channel (hiGIRK2) gb:U24660			
E5-35	5						
H6-33							
H12-35 R5-1 1	1	22180	5-8-30-43-44-47	Human cDNA clone J7194F sb:N56107	3.8e-19	NONE	-
F4-31	2	552A3	276-279-303-311-325	NONE	1	NONE	
G2-33							
H5-32 F9-33	2	552A3	273-278-288-309-312	NONE	•	NONE	-
G1-34	3	552A3	282-286-308	NONE	-	NONE	-
B11-31							
D5-31							
G1-36	1	552A3	273-288-312-320	NONE	· · · · · · · · · · · · · · · · · · ·	NONE	
A9-32	1	552A3	271-288-312-320	NONE	-	NONE	-



YAC 221B9



Figura 12. Mapa Transcripcional dels YAC 72H9, 336G11, 238B1, 221B9, 552A3 (2,2 Mb)

YAC 552A3





Figura 13. Estratègia utilitzada per a l'aïllament de nous EST

truncat en MNBH-iso2 i MNBH-iso3. La repetició de 13 histidines i la regió rica en serines i treonines solament estan presents en la variant més llarga: MNBH-iso1.

L'expressió de MNBHa és ubíqua, tant en teixits fetals com en teixits adults, mentre que MNBHb s'expressa en cor i múscul esquelètic, però no en cervell. L'expressió de *Mnbh* en el cervell adult de ratolí és marcadament elevada en el bulb olfactori, el cerebel, l'escorça cerebral, la capa de cèl·lules piramidals de l'hipocamp i en diversos nuclis hipotalàmics.

5.1. CONSTRUCCIÓ DE LA GENOTECA DE COSMIDIS

La construcció d'una genoteca de cosmidis és important per a la seva aplicació en la tècnica de la selecció de cDNA. Hem construït una genoteca de 502 cosmidis a partir de cinc YAC (72H9, 336G11, 238B1, 221B9 i 552A3) específics de tres regions de l'HC21: 1) la regió crítica de la síndrome de Down, una regió distal, i una altra proximal (2,2 Mb). Aquests cosmidis s'han ordenat seqüencialment, i conformen un mapa contigu de cosmidis solapats. Respecte a altres mapes de cosmidis construïts a partir de l'HC21, el mapa de Nizetic et al., (1994) pretén cobrir tot el braç llarg de l'HC21 amb cosmidis. però aquests estan agrupats i no s'ha pogut establir un ordre seqüencial dels mateixos. Els mapes de Patil et al., (1994), el de Lucente et al., (1995) i el d'Osoegawa et al., (1996) es centren en la regió delimitada pels marcadors CBR i ETS2, que inclou la DSCR. Lucente et al., (1995) tampoc ha establert un mapa contigu de cosmidis, i solament Patil et al., (1994) i Osoegawa et al., (1996) han construït mapes de clons solapats, que quan s'intenten comparar mostren nombroses diferencies. D'aquesta manera, tots dos difereixen no solament en l'ordre dels grups de cosmidis, sinó també en la seva assignació al HC21. Solament un grup de cosmidis de Patil et al., va poder ésser integrat en el mapa d'Osoegawa et al. El conjunt de cosmidis de tots dos mapes s'ha obtingut a partir d'hibridacions de membranes de cosmidis, en les que s'ha utilitzat cada YAC de l'HC21 com a sonda. Aquest mètode redunda en la falsa assignació d'un cosmidi a un YAC, perquè el nombre de falsos positius és molt alt quan s'empren sondes d'aquesta naturalesa. S'estima que el 30% dels clons seleccionats amb aquesta tècnica estan localitzats erròniament (Osoegawa et al., 1996). Atessos els problemes inherents d'aquesta tècnica i els obtinguts en l'aplicació de la tècnica de la selecció de cDNA directament a partir del DNA dels YAC (com es comentarà en l'apartat 5.2.), vàrem decidir construir una genoteca de cosmidis directament de cada YAC no quimèric, a

partir de cadascuna de les tres regions d'interès de l'HC21. Aquests cosmidis ens havien de facilitar també, la localització precisa a l'HC21 de cadascun dels EST aïllats, conformant un mapa transcripcional d'alta resolució.

Prèviament a la construcció de la genoteca de cosmidis, s'ha comprovat el possible quimerisme i la integritat de cada YAC per dues estratègies diferents:

1) La tècnica de FISH sobre metafases humanes.

2) L'anàlisi de restricció amb la tècnica de PFGE.

Respecte a la tècnica, hem utilitzat el producte d'una PCR inter-Alu de cada YAC com a sonda fluorescent per a estudis de FISH. La manca de bandes discretes en la PCR amb els dos encebadors consensus de la següència Alu indica que aquest mètode és una bona manera per marcar uniformement el DNA genòmic, i que el resultat d'aquesta amplificació és una mostra prou específica i representativa del contingut d'un determinat YAC. La sensibilitat d'aquesta tècnica queda palesa quan demostrem que el nombre de YAC quimèrics detectats per aquest mètode s'eleva al 60%, xifra lleugerament superior al 46% descrita per a aquesta genoteca. D'aquesta manera, l'extrema sensibilitat d'aquest mètode per a la detecció de quimerisme mostra que el 29% dels YAC descrits com a no quimèrics, realment sí ho són. Aquells YAC que mitjançant FISH donaren dues senyals en cromosomes diferents d'una mateixa metafase humana foren descartats per quimerisme. Aquells que donaren senyal en dos loci de l'HC21, indicant que el YAC tenia una llarga deleció o menys possiblement que, un mateix clon de llevat contenia dos YAC diferents de la mateixa mida, també foren descartats. Solament s'han seleccionat aquells YAC que donaren un únic i fort senval a l'HC21.

En relació a la segona estratègia, les possibles reorganitzacions i delecions internes de cada YAC s'han posat en evidència mitjançant la consistència en els resultats dels mapes de macrorestricció amb la tècnica de PFGE. La detecció de discrepàncies en el patró de bandes electroforètiques entre un determinat YAC i els altres que el contenen o que el solapen s'ha interpretat com una reorganització anòmala del DNA del YAC, i per tant aquests YAC s'han descartat.

Amb aquests criteris de selecció s'han escollit cinc YAC que *a priori* poden ésser considerats com a no quimèrics. D'aquesta manera, tots els EST aïllats a partir d'aquests YAC hauran de pertànyer a les regions d'interès de l'HC21. Els cinc YAC representen tres regions de l'HC21:

1) Els YAC 336G11, 238B1 i 221B9 cobreixen una regió aproximada d'1,3 Mb al voltant del marcador D21S55, que inclou l'anomenada DSCR.

2) El YAC 552A3 de 375 kb de mida, localitzat a prop del marcador D21S53 i representa un fragment més distal que els anteriors.

3) El YAC 72H9 de 450 kb està localitzat a prop del gen AML1. Aquest darrer YAC fou escollit perquè a partir d'ell el nostre Departament havia aïllat i caracteritzat un nou gen, el gen DSCR1 (Fuentes *et al.*, 1995), i podia ésser un control positiu excel·lent en l'aplicació de la tècnica de la selecció de cDNA.

Respecte a la construcció de la genoteca de cosmidis, s'ha fet una transferència de DNA de tota la col·lecció de clons obtinguts del clonatge en el cosmidi supercos1 dels fragments de restrició generats a partir de cada YAC. D'aquesta manera, la membrana que conté el DNA de tots ells, s'ha hibridat amb una sonda humana de DNA repetitiu Alu. Això va permetre seleccionar els clons que contenien següències de l'HC21 i descartar aquells que solament contenien sequències del genoma de llevat. Per a descartar aquells cosmidis quimèrics humà-llevat s'ha fet una digestió amb l'enzim de restricció EcoRI de la minipreparació de DNA de cadascun dels cosmidis, i els productes s'han separat electroforèticament. La transferència de Southern del DNA de cadascuna de les digestions, discriminades electroforèticament en un gel d'agarosa, s'ha hibridat, en experiments paral·lels, amb sondes de DNA específiques d'humà i de llevat, respectivament. Aquest segon clivellatge va resultar en la no identificació de cap cosmidi quimèric humà-llevat. Aquests resultats suggereixen que, possiblement, tampoc no hi ha cap clon quimèric humà-humà, ja que la possibilitat d'obtenir un clon quimèric humàhumà és molt més petita (unes 50 vegades inferior) que la d'obtenir un clon quimèric humà-llevat. D'aquesta manera obviem dos grans possibles problemes, ja que els clons quimèrics (humà-llevat) haguessin pogut disminuir l'eficiència de la tècnica de la selecció de cDNA, i els clons quimèrics humà-humà haguessin dificultat enormement la construcció del mapa contigu de cosmidis solapats

Respecte a les diferents tècniques utilitzades en la construcció del mapa contigu de cosmidis i l'establiment de llur ordre (mencionades en l'apartat 3.4.2.), destaquem la síntesi de ribosondes per la seva gran especificitat, malgrat que és un mètode molt laboriós. Contràriament, creiem que la construcció de mapes de cosmidis segons el seu contingut de STS pot generar forats aparents o solapaments no detectables, però és molt útil per la seva rapidesa.

D'altre banda, l'endreçament de la col·lecció d'aquests cosmidis en un mapa contigu i solapat ofereix informació sobre llur ordre seqüencial i l'existència de possibles forats, a diferència d'un simple conjunt de cosmidis o cosmid-pocket que deriven d'un YAC o d'una regió determinada, on el nombre de forats o gaps i la seva posició són desconeguts. Aquest mapa contigu de cosmidis implica un millorament en la resolució del mapa físic, alhora que defineix el nombre i la posició dels forats existents, on dirigir posteriors esforços i completar el solapament de cosmidis. Aquesta informació és imprescindible per assegurar-nos que tot el material genòmic del YAC està representat en la genoteca de cosmidis. D'aquesta manera, la generació del mapa contigu de cosmidis/PAC ens ha permès localitzar un forat en el YAC 221B9 i definir la seva mida. Possiblement els forats obtinguts en el mapa contigu de cosmidis són causats perquè en una determinada regió de l'HC21 hi ha una mancança, o bé un elevat nombre de dianes per a l'enzim de restricció emprat en la construcció de la genoteca de cosmidis, de tal manera que els fragments de DNA obtinguts en la digestió parcial del YAC són massa grans, o bé massa petits per a ésser posteriorment encapsulats en les partícules del fag lambda, donades les restriccions en la mida del DNA que aquest procés comporta. Aquesta interpretació està avalada per l'observació que el forat obtingut té una mida considerablement petita i per tant, no altera significativament el mapa de transcripció. Mitjançant la comparació entre la mida del YAC 221B9 obtinguda per PFEG i la mida de la suma de les mides dels fragments de DNA de la digestió amb l'endonucleasa EcoRI de cadascun dels cosmidis contigus per a aquest YAC, estimem que la mida del forat d'aquest YAC és inferior a 10 kb. Això significa que aproximadament el 99,6% de la regió genòmica coberta pels cinc YAC està representada en el mapa contigu de cosmidis.

Un mapa contigu de cosmidis com el que hem construït, a més d'ésser el marc de treball imprescindible per a la construcció d'un mapa transcripcional de la regió, ha de permetre identificar més acuradament els punts de trencament o *breakpoints* de les trisomies parcials de l'HC21, i per tant, les correlacions genotip-fenotip d'aquestes persones ens permetrà definir la DSCR amb més precisió.

5.2. SELECCIÓ DIRECTA DE CDNA DE L'HC21

Hem optat pel mètode de selecció de cDNA perquè és l'únic dels presentats en l'apartat 1.4. que parteix de cDNA. D'aquesta manera, aquest mètode ens ha permés identificar gens específics de determinades regions cromosòmiques de l'HC21 i que s'expressen en un determinat teixit d'interés (p.e. el cervell humà fetal), la qual cosa facilita l'aïllament de gens involucrats en el desenvolupament o manteniment de l'SNC, possiblement implicats en l'etiologia molecular de la discapacitat intel·lectual de la DS. Per tal de minimitzar les variacions de l'expressió segons l'edat, el sexe i les individuals, vàrem construir una genoteca no clonada de cDNA a partir de RNA de cervell humà fetal de diferents individus i edats, perquè la genoteca de cDNA reflectís la complexitat real d'aquest teixit. A més a més, el teixit fetal del cervell humà reuneix la complexitat suficient com per tenir una població de RNA prou representativa de la cèl·lula, ja que aquest teixit és el més complex de tots en quant al nombre de gens que s'expressen (Adams *et al.*, 1995). Així mateix, malgrat que en aquest teixit no s'expressen tots els gens del genoma, representa el teixit més interessant per a un dels trets fenotípics més important de la DS, com és la discapacitat intel·lectual. Cal destacar que la quantitat d'espècies de mRNA d'un teixit no neural com fetge o ronyó és normalment de dos a tres vegades inferior respecte al trobat en cervell fetal (Van Ness i Hahn 1980).

Per a l'aplicació del mètode de la selecció de cDNA cal destacar dos elements molt importants:

1) El primer element imprescindible és l'RNA. S'ha comprovat la qualitat de l'RNA, emprat per a la construcció de la genoteca de cDNA, mitjançant una electroforesi en un gel desnaturalitzant d'agarosa al 1%. Per a estimar la qualitat de la genoteca de cDNA construïda, s'ha comprovat la presència de material genòmic mitjançant la reacció de la PCR, utilitzant encebadors intrònics del gen *DSCR1*. Cal esmentar que l'hnRNA (RNA nuclear immadur que no ha processat encara els introns) s'ha eliminat en el procés de purificació dels RNA poli(A)⁺ mitjançant una columna poli-(dT).

2) L'altre element imprescindible per a la selecció de cDNA és el DNA genòmic, que pot estar representat en YAC, PAC, BAC o en cosmidis. En experiència del Dr. M. Lovett, l'ús de cosmidis en la tècnica de la selecció de cDNA rendeix nivells d'enriquiment superiors amb un soroll de fons menor, que quan s'utilitzen YAC en comptes de cosmidis.

La nostra experiència ens indica que les genoteques comercials de cDNA, sense discriminació de marca, solen tenir una bona representació i preponderància de clons genòmics i d'artefactes com a *bonus* afegit, malgrat el seu alt preu. Aquestes desavantatges de les genoteques de cDNA comercials, amplificades moltes vegades (sovint és difícil trobar aquells trànscrits poc representats, o sia $<1 \times 10^{-6}$) es poden superar emprant genoteques primàries, no subclonades, i lligades a uns adaptadors per amplificar-les mitjançant la PCR. Amb una genoteca de cDNA com aquesta, cada molècula amplificada pot ésser potencialment seleccionada. Després de dos cicles de selecció, és possible identificar trànscrits que inicialment estaven representats a nivells molt baixos, sense ésser desplaçats per aquells trànscrits que tenen una alta expressió. D'aquesta manera, aquest mètode tendeix a una certa normalització dels cDNA, amb un factor d'enriquiment més gran per a aquells trànscrits d'expressió baixa respecte a aquells més abundants (Tagle *et al.*, 1993). Atès que la hibridació es duu a terme en solució aquosa, la reacció de reaparellament de les cadenes del DNA seguirà una cinètica de segon ordre, de les clàssiques reaccions de " $C_0 t$ ": les seqüències més abundants es reaparellen més ràpidament que les seqüències menys abundants. D'aquesta manera, en el tub de reacció de la selecció de cDNA, els trànscrits d'expressió molt baixa (alt $C_0 t_{1/2}$) quedaran enriquits respecte als primers (cDNA més abundants o baix $C_0 t_{1/2}$).

La primera vegada que vàrem utilitzar la tècnica de selecció de cDNA, es va emprar directament el DNA genòmic d'un YAC, doblement purificat mitjançant dues PFGE, per seleccionar cDNA a partir d'una genoteca de cDNA comercial de cervell fetal. El resultat no podia haver estat més negatiu: el 100% dels clons (480/480) van resultar ésser seqüències ribosomals. Possiblement la purificació del YAC lliure de contaminats no sigui fàcil, ja que indetectables quantitats de degradació del cromosoma XII, el qual porta >100 còpies de *loci* ribosomals de llevat, pot comigrar en l'electroforesi amb el nostre YAC. Suposem que aquesta petita quantitat de gens ribosomals del llevat ha estat suficient per seleccionar d'una manera efectiva els contaminants de cDNA ribosòmics, presents també en la genoteca comercial de cDNA, malgrat que *a priori* s'ha intentat bloquejar-les.

La segona vegada que vam emprar aquesta tècnica, a partir també d'una genoteca comercial de cDNA de cervell fetal, ho vam fer a partir dels cosmidis que derivaven del YAC 221B9. Els resultats no foren gaire millors que els anteriors. D'un total de 150 clons, el 60% dels clons seleccionats contenien sequències de DNA mitocondrial. La resta de clons (40%) foren DNA repetitiu del tipus *Alu* i artefactes de PCR (sequències poli(A) i encebadors concatenats). Les sequències de DNA mitocondrials provenen dels cDNA mitocondrials derivats dels trànscrits mitocondrials poliadenilats que estan presents, de manera exacerbada, en totes les genoteques comercials de cDNA (Adams *et al.*, 1995), i que són seleccionades per l'alt grau de similitud amb una part del genoma bacterià d'*E.coli*, que està present com a contaminant en algunes preparacions de DNA dels cosmidis. Les sequències de DNA repetitiu del tipus *Alu* se seleccionen atès l'alt contingut de contaminants de sequències genòmiques que també es troben de manera exacerbada en les genoteques de cDNA comercials (Adams *et al.*, 1992).

Els resultats d'un tercer intent, a partir dels cosmidis del YAC 336G11 i d'una genoteca de cDNA d'una altra casa comercial no tingueren millor sort que les anteriors. És per tot això que, finalment, s'ha optat per construir una genoteca pròpia de cDNA, i purificar els cosmidis per ultracentrifugació en gradient de CsCl. L'aplicació de la selecció de cDNA amb aquests darrers elements ha donat els següents resultats: dels 502 cosmidis generats a partir dels cinc YAC mencionats hem obtingut 576 clons, dels quals s'han localitzat 120 clons a l'HC21. El nombre de clons no redundants és de 45, entre els quals es troben:

1) Dos gens descrits: GIRK2 (Ferrer et al., 1995) i DSCR1 (Fuentes et al., 1995),

2) Un nou gen que, per la seva similitud amb el gen minibrain de Drosophila, l'hem anomenat MNBH.

3) Un EST descrit prèviament a l'HC21.

4) Sis EST descrits per projectes de seqüenciació de clons de cDNA a partir de l'extrem poli(A), però no localitzats enlloc.

5) Quatre clons amb similituds baixes amb altres EST.

6) Un STS de l'HC21.

7) Un clon amb una similitud moderada amb el gen de la proteïna centractina de llevat.

La resta dels vint-i-cinc EST descrits no comparteixen similituds amb cap tipus de sequència del banc de dades.

L'elevat nombre de clons obtinguts que contenen DNA ribosomal, després d'aplicar la tècnica de la selecció de cDNA (471/576), suggereix que el bloqueig d'aquest tipus de seqüència durant la prehibridació del cDNA amb DNA ribosomal no fou prou bo. Tampoc podem descartar la possibilitat que a la regió estudiada de l'HC21 hi hagi algun gen ribosomal i que s'ha seleccionat de manera específica. Tanmateix, el bloqueig de les seqüències repetitives amb DNA COT1 sembla haver estar suficient, ja que solament tres clons contenien aquest tipus de seqüència.

Com s'ha mencionat anteriorment, l'altre element important en la tècnica de selecció de cDNA són els cosmidis. Llurs DNA foren purificats mitjançant un gradient de CsCl i marcats posteriorment amb biotina. A fi i efecte de comprovar el marcatge dels cosmidis amb biotina, s'ha incorporat una petita quantitat de dNTP radioactiu en la reacció de la biotinilació dels cosmidis. Mitjançant aquesta petita quantitat isotòpica incorporada s'ha mesurat la proporció entre la radioactivitat associada al cosmidi respecte la lliure, i d'aquesta manera s'ha comprovat que la reacció de marcatge amb

biotina fos l'esperada. Aquesta petita quantitat isotòpica incorporada al DNA dels cosmidis també ens ha permès fer un seguiment del material biotinilat durant els cicles de selecció.

S'ha obtingut una mesura quantitativa del grau d'enriquiment dels cDNA. mitjançant la comparació del nombre de colònies positives del clivellatge d'una genoteca de cDNA de cervell humà fetal i el nombre de clons positius que resulten després de cada cicle de selecció, quan s'hibriden amb la sonda del gen que actua de control positiu. En aquest sentit, quan s'ha utilitzat una sonda del gen DSCR1 (control positiu) s'ha obtingut 1/10⁵, 1/10⁴ i 2/10² clons positius després dels cicles de selecció zero, primer i segon, respectivament, aconseguint un enriquiment final de 2.000 vegades per al gen DSCR1. Per a altres EST d'expressió no tan elevada com el gen DSCR1, hem calculat que el grau d'enriquiment ha estat al voltant d'unes 50.000 vegades. Una altra mesura estimada del grau d'enriquiment del gen control fou la hibridació de la membrana que contenia els productes electroforètics de l'amplificació per PCR de cadascun dels controls relacionats en l'apartat 3.6.5.5. amb una sonda del gen DSCR1. Els resultats d'aquesta hibridació (Fig. 14) mostren com el gen DSCR1 s'enriqueix de manera conspicua durant el primer i segon cicle de selecció, però no pas en els successius cicles. Aquests resultats suggereixen que dos cicles de selecció han estat suficient per tenir una població prou enriquida de cDNA per al seu posterior subclonatge i estudi. Generalment, una tercera ronda de selecció no incrementa el nombre d'espècies de cDNA seleccionats (Morgan et al., 1992; Gecz et al., 1993) i a més, fa que els fragments seleccionats tinguin cada cop una mida més petita, ja que la PCR té la tendència a afavorir-los.

La tècnica de la selecció de cDNA pot també seleccionar seqüències de DNA repetitiu del tipus *Alu*, ja que és possible que en la genoteca de cDNA original hi hagués una petita proporció de trànscrits immadurs. Aquest clons es poden detectar mitjançant l'hibridació de la transferència de *Southern* de tots els clons resultants de la tècnica de la selecció de cDNA amb una sonda de DNA total humà. *A priori*, aquests clons positius de l'hibridació amb la sonda humana s'han de descartar. Tanmateix, la inserció d'una seqüència *Alu* dintre d'un cDNA és un esdeveniment poc conegut però descrit (Mittaz *et al.*, 1997; Guimera *et al.*, sotmès a publicació, veure apartat 4.1.4.). En el nostre cas, com que solament teniem tres clons, els vàrem seqüenciar per saber en quin dels dos casos estàvem. Tots tres clons van resultar ésser seqüències de DNA repetitiu del tipus *Alu*.

Dos paràmetres a tenir en compte per a avaluar la qualitat dels resultats de la selecció de cDNA són la sensibilitat i l'especificitat. Respecte a la sensibilitat, l'observació de que algun EST aïllat i posteriorment utilitzat com a sonda radioactiva d—na senyal en una transferència de *Northern* en teixits diferents a partir del qual s'ha obtingut l'RNA utilitzat en la tècnica de selecció de cDNA, suggereix que les condicions de la selecció de cDNA han estat prou sensibles com per seleccionar i amplificar, durant els dos cicles de selecció aplicats, aquelles espècies de RNA d'expressió molt reduïda. Respecte a l'especificitat, tots els clons que van restar després de descartar aquells que contenien seqüències de DNA ribosomal i repetitiu, s'han localitzat en algun cosmidi dels utilitzats en la selecció de cDNA, obtenint així un 100% d'especificitat dels EST.



Figura 14. Especificitat i sensibilitat del procés de selecció i enriquiment dels cDNA. 1: Genoteca de cDNA original sense amplificar. 2: PCR de la genoteca de cDNA, sense posar l'encebador OLIGO3 (1r control negatiu de la PCR). 3: PCR del cDNA després d'un cicle de selecció. 4: PCR del cDNA després de dos cicles de selecció. 5: PCR del cDNA després de tres cicles de selecció. 6: PCR del cDNA després de quatre cicles de selecció. 7: PCR amb encebadors, però sense cDNA (2n control negatiu de la PCR). 8: PCR del material no específic del 6è rentat després del 2n cicle de selecció. 9: PCR d'un fragment del gen DSCR1 (control positiu de l'hibridació). 10: Marcador de la mida del DNA (# 520-5615SA de Gibco BRL).

La limitació més important de la tècnica de la selecció de cDNA és que l'habilitat per seleccionar un determinat cDNA depèn de l'expressió temporal i espaial d'aquest. Sembla clar que si es vol obtenir un mapa de transcripció complet d'aquesta regió caldrà aplicar aquesta tècnica a altres teixits. També serà necessari emprar aquesta tècnica en combinació amb altres estratègies. La nostra experiència amb les tècniques de *cDNA selection* i *exon trapping* ens mostra que la primera permet identificar clons de seqüència més llarga, la qual cosa facilita força els anàlisis posteriors, mentre que la segona tècnica permet trobar la cadena amb sentit i l'orientació del corresponent cDNA.

5.3. CONSTRUCCIÓ D'UN MAPA TRANSCRIPCIONAL

La genoteca construïda de cosmidis que se solapen entre si i endreçats seqüencialment ha representat un recurs immillorable per a la construcció d'un mapa d'alta resolució de les regions de l'HC21 estudiades. Aquest darrer mapa, a més a més d'ésser el marc de treball següenciació de tot 1'HC21 idoni pel projecte de (http://wwweri.uchsc.edu/chr21/eridna.html), ha estat molt útil per a integrar tota la informació genètica, física i transcripcional generada en un sol mapa funcional. D'aquesta manera, aquest nou mapa d'alta resolució, ha permès assignar acuradament cadascun dels nous cDNA parcials aïllats a un determinat cosmidi. Això ha estat de particular ajuda en l'agrupament d'aquests fragments de cDNA en unitats de transcripció. Al mateix temps, l'assignació de cada unitat transcripcional a una determinada localització en el mapa posa de manifest aquelles regions genòmiques que són marcadament pobres des del punt de vista transcripcional.

L'assignació de cada EST a un determinat YAC mitjançant la tècnica del dot blot (segons s'ha esmentat en l'apartat 3.3.5.1.) ens hagués donat un patró similar d'hibridació per a aquells clons que contenien cDNA ribosòmic que per a aquells que contenien cDNA d'un gen humà amb una alta similitud a nivell nucleotídic amb algun gen de llevat. Per evitar això, s'han utilitzat membranes que contenien una concentració de DNA diferent de llevat i d'humà, però equivalent en el nombre de molècules de DNA. D'aquesta manera, rentats successius d'alta astringència després de l'hibridació individual dels *dots blots* amb cada EST aïllat, ens ha permès discriminar cadascun dels dos casos.

Malgrat la gran quantitat d'EST aïllats en la DSCR per altres laboratoris (veure apartat 1.2.3.6.), solament un EST dels presentats en aquest treball coincideix amb un

dels 559 exons potencials descrits per Chen *et al.*, (1996). Una possible explicació al fet que no s'han trobat similituds entre els 37 nous EST descrits en aquest treball i altres prèviament descrits en aquesta zona de l'HC21 pot ésser perquè tant uns com els altres tenen una mida relativament petita, i malgrat que pertanyin al mateix gen, és molt difícil que les seves seqüències se solapin entre si. Una altra explicació podria ésser que malgrat la gran quantitat d'EST descrits en el HC21, aquests són putatius, ja que no s'ha demostrat que aquests cDNA parcials siguin veritables seqüències transcripcionals. En aquest darrer sentit, moltes de les seqüències descartades per nosaltres com a clons ribosomals, mitocondrials o DNA repetitiu, les hem trobat també a les bases de dades com a EST específics de l'HC21. Contràriament a la tendència del que s'ha publicat, el nostre treball demostra, mitjançant experiments de transferència de *Northern* i d'RT-PCR, que cadascun dels EST identificats són veritables seqüències que s'expressen.

De les 37 noves seqüències d'EST aïllades i caracteritzades, no podem afirmar encara que cadascuna d'elles representi un gen diferent. És possible que diferents clons que no hibriden entre si continguin fragments del mateix cDNA i pertanyin a la mateixa unitat de transcripció, ja que la genoteca de cDNA emprada en la selecció de cDNA fou construïda a partir d'oligo-(dT) i d'hexàmers a l'atzar (dN)₆ en experiments paral·lels. Això és causat, principalment, per la mida petita dels clons de cDNA, de manera que aquest nombre d'EST és possible que sobreestimi el nombre real de gens trobats en aquesta regió. La recerca de més seqüència per a cadascun dels cDNA parcials aïllats es fa necessària per dues raons: 1) ens permetrà saber quins clons pertanyen a la mateixa unitat de transcripció i 2) ens permetrà conèixer la regió codificant, la qual cosa ha d'ajudar a trobar similituds amb altres gens dipossitats en els bancs de dades.

Aquesta informació ajudarà a decidir quin dels clons caracteritzats pot ésser millor candidat per a explicar determinats fenotips observats en la DS. De totes maneres, el fet que cadascun d'aquest clons representi espècies de RNA, amb mobilitats electroforètiques diferents en les anàlisis de transferència de *Northern*, suggereix que aquests cDNA parcials probablement pertanyen a diferents gens que no pas siguin múltiples fragments del mateix gen.

El mapa de transcripció reflecteix una distribució desigual dels cDNA trobats. La diferència en la densitat de cDNA en cada YAC, pot ésser causada en part a les diferències en l'activitat transcripcional de cada regió, però també pot estar causada per la diferent complexitat genòmica inicial amb la que cada YAC estava representat o bé per diferències en l'amplificació de la PCR.

La distribució regular d'alguns cDNA en alguns YAC i el fet que molt pocs dels 120 EST trobats i localitzats en l'HC21 siguin idèntics, indica que el gran nombre de clons (576) que s'han recuperat després de la selecció de cDNA fou necessari i suficient per aïllar aquells trànscrits pocs representats. La comparació entre el nombre de clons amb aquests cDNA d'expressió molt baixa i el nombre de clons que pertanyen a gens coneguts, que sabem que es troben de manera molt representada en la genoteca inicial de cDNA, suggereix que s'ha obtingut un bon grau de normalització dels cDNA durant el dos cicles de selecció aplicats. Aquests resultats corroboren l'observació d'estudis anteriors que indicaven que aquesta tècnica tendeix a una certa normalització en els nivells dels trànscrits (Tagle *et al.*, 1993).

Al finalitzar el treball d'aquesta tesi, els treballs de seqüenciació de determinades regions de l'HC21 s'han fet accessibles a tothom. Més concretament, la regió que correspon als YAC 336G11, 238B1 i 221B9, i 72H9 evidencia que l'ordre seqüencial dels EST presentat en aquest treball és correcte, i també l'ordre dels marcadors i cosmidis, i les distàncies físiques proposades (Carles Pucharcós, comunicació personal).

Finalment, els cosmidis ordenats i localitzats en el mapa físic d'aquesta regió poden ésser utilitzats individualment com a sondes en l'anàlisis de FISH per determinar el nombre de còpies que hi ha de la DSCR en un determinat individu. Sens dubte, la col·lecció generada de cosmidis de la DSCR ha de permetre un millor diagnòstic de la DS, mitjançant anàlisis de FISH en aquells individus amb trisomies parcials que impliquin la DSCR, indetectable en un anàlisi citogenètic convencional.

5.4. CARACTERITZACIÓ D'UN NOU GEN HUMÀ: MNBH

La sequència de DNA del clon D1-34, obtinguda mitjançant la tècnica de la selecció de cDNA, resultar ésser un fragment d'un nou gen que, atesa la seva similitud amb el gen *mnb* de *Drosophila*, l'hem anomenat *Minibrain homolog (MNBH)*.

En Drosophila, el gen mnb codifica una proteïna d'una nova família de quinases: una quinasa serina-treonina, i sembla tenir un paper essencial en la neurogènesi postembrionària, ja que podria regular el nombre i la diferenciació de diferents tipus de cèl·lules neuronals (Tejedor *et al.*, 1995). Les mosques amb mutacions en el gen mnb es caracteritzen per tenir una reducció marcada de la mida dels lòbuls òptics i dels hemisferis centrals del cervell. Aquest fenotip està causat per un espaiament anormal dels neuroblasts i per tant una reducció de la progènie neuronal (Tejedor *et al.*, 1995).

Des del punt de vista morfològic, els cervells de les persones amb la DS són més

petits respecte als individus sense la DS (Kemper, 1989) i tenen una disminució en el nombre de neurones en algunes àrees del cervell, que inclou el cerebel, l'hipocamp i l'escorça cerebral (Becker *et al.*, 1991). L'hipocamp, el cerebel i l'escorça cerebral, a més d'ésser les àrees de l'SNC més freqüentment alterades en la DS, són també les regions on l'expressió del gen *Mnbh* de ratolí és més elevada. Atès que aquestes zones estan majoritàriament involucrades en els processos cognitius, d'aprenentatge, de memòria de llarg termini, en la presa de decisions i en el processament del llenguatge i del pensament abstracte, no és difícil imaginar que aquells gens que s'expressen majoritàriament en aquestes àrees, com *MNBH*, puguin estar involucrats, en un o altre grau, en la discapacitat intel·lectual que s'observa en la DS.

Durant la caracterització del gen humà MNBH, Kentrup et al., (1996) van aïllar i caracteritzar el gen de rata homòleg de minibrain, anomenat Dyrk (o també anomenat Mnbh). Dyrk i MNBH són gens homòlegs. La seqüència d'aminoàcids deduïda de llurs pautes de lectura conté el mateix nombre d'aminoàcids, compartint un 99,7% d'identitat a nivell aminoacídic i una similitud del 96,6% a nivell nucleotídic. La regió catalítica és quasi idèntica excepte en un sol aminoàcid i comparteixen la totalitat dels possibles dominis funcionals d'aquesta regió. Els motius pels quals l'evolució ha conservat aquesta estreta similitud entre la sequència de Dyrk i de MNBH són encara desconeguts, però suggereix un paper essencial d'aquest gen en la cèl·lula. ¿Com es pot entendre sinó, que entre la proteïna de rata i l'humana solament hi hagi dos canvis aminoacídics en un total de 763 aminoàcids? Per què la identitat d'aminoàcids és tan estreta entre les proteïnes de Drosophila i l'humana, malgrat la gran quantitat de polimorfismes (tant silenciosos com no) que s'han trobat en la seqüència de mnb de Drosophila (Tejedor et al., 1995) i en la seqüència del gen humà MNBH (Guimera et al., sotmès a publicació)?

A igual que *mnb* en *Drosophila*, i com s'ha esmentat anteriorment, *Dyrk* codifica per a una serina-treonina quinasa. La fosforilació reversible de les proteïnes representa el mecanisme regulador més important en la senyalització cel·lular (Edelman *et al.*, 1987; Hunter 1991), a més de representar una de les famílies més llargues de proteïnes (s'han caracteritzat més de 100 quinases de mamífer). Segons criteris funcionals i estructurals, coneixem principalment dos grups de quinases: les tirosina quinases i les serina/treonina quinases, segons quin sigui el residu que fosforilen. La quinasa Dyrk sembla ésser una de les poques proteïnes amb especificitat dual, doncs té la capacitat d'autofosforilar-se en residus tirosina i en residus serina/treonina, i també fosforilar residus tirosina de

254 Resum i discussió -

substrats exògens com histones. La seva activitat quinasa sembla estar regulada per la fosforilació d'un o tots dos residus tirosina que hi ha entre els subdominis VII i VIII de la regió catalítica, els quals conformen l'anomenat lloc d'activació o *activating loop* (Kentrup *et al.*, 1996). Podriem especular amb la possiblitat que la quinasa Dyrk, i molt possiblement MNBH, puguin ésser components de la cascada de senyalització cel·lular, mitjançant la fosforilació específica de factors de transcripció en el nucli, ja que totes les altres proteïnes regulades per fosforilació en tirosina entre els subdominis VII i VIII, p.e. MAPK/ERK, JNK i GSK3ß, són components de la senyalització cel·lular implicades en la transmissió de senyals extracel·lulars al nucli i que fosforilen, directament o indirectament, factors de transcripció (Kentrup *et al.*, 1996).

Aquestes tres quinases (*mnb*, *Mnbh* i *MNBH*) comparteixen una gran similitud a nivell nucleotídic amb altres proteïnes quinases dependents de ciclina (cdks), les quals participen en la regulació de la proliferació cel·lular en diverses espècies, la qual cosa suggereix un mecanisme de regulació universal (Sherr, 1993; Nigg, 1995).

El gen MNBH consta de 17 exons i abarca 150 kb de DNA. Mitjançant estudis de FISH sobre metafases de cromosomes humans hem obtingut senyal en un sol *locus*, suggerint que MNBH és un gen de còpia única. L'anàlisi de transferència de Southern de DNA genòmic, digerit amb un enzim de tall freqüent, mostra una única banda de DNA quan s'hibrida amb una sonda específica del gen de MNBH en condicions de baixa astringència, la qual cosa corrobora els resultats anteriors de FISH. Aquests mateixos resultats indiquen que la banda de DNA obtinguda en la digestió del DNA genòmic total correspon a l'HC21, al observar-se la mateixa banda en la digestió del DNA de la línia cel·lular híbrida WAV-17, que conté l'HC21 com a únic component humà. MNBH és doncs un gen de còpia única i es localitza en la DSCR de l'HC21. Mitjançant anàlisis moleculars, MNBH està localitzat entre els marcadors D21S335 i D21S337, i és a 186 nt del marcador de referència D21S55.

Segons quin sigui l'inici de transcripció per l'ús alternatiu del primer exó, hem anomenat trànscrits MNBHa i MNBHb a les dues variants de *MNBH*. L'inici de transcripció del trànscrit MNBHb és concret, mentre que els resultats de tres experiments diferents independents, cadascun d'ells amb teixits diferents i amb encebadors diferents, mostren que l'inici de transcripció del trànscrit MNBHa és múltiple, i sempre dins d'una finestra de 32 nt, en tots els experiments duts a terme. Aquesta regió d'inici pot estar controlada per la presència d'un element anomenat MED-1 (seqüència de DNA que promociona un inici de transcripció irregular) (Fig. 15) que hem trobat en l'intró 2 de MNBHa. Aquests elements reguladors es troben sempre a 20-45 pb corrent abaix o *downstream* de l'extrem 3' del final del primer exó, i a un màxim de 110 pb corrent abaix del 5' del començament de la finestra d'iniciació de la transcripció, i estan associats a múltiples llocs d'inici de la transcripció (Ince i Scotto, 1995).

D'aquesta manera, cadascun d'aquests trànscrits estan probablement sota el control d'un promotor diferent. Aquesta hipòtesi està corroborada pel diferent patró d'expressió que mostra cadascun dels trànscrits en anàlisis de transferència de *Northern*.



Figura 15. Estructura genòmica de les regions 5' de MNBHa i MNBHb

El primer exó del trànscrit MNBHa (exó 2) es troba dins d'una illa CpG. Hem comprovat també, mitjançant experiments de transferència de *Southern*, que en la seqüència de DNA genòmic d'aquesta illa CpG hi ha tantes dianes per a enzims de restricció sensibles a la metilació i de tall infreqüent com dianes hi ha en el DNA d'aquells YAC o cosmidis que contenen aquesta seqüència. Ja que el DNA subclonat en el YAC no pateix metilació en citosina (Butler *et al.*, 1992) ens fa pensar que aquesta illa CpG roman també de manera no metilada *in vivo*, com caldria esperar d'una autèntica illa CpG. La presència d'aquesta illa CpG o HTF en la regió promotora és consistent amb les característiques d'un gen d'expressió ubiqua. S'especula que aquestes illes CpG mantenen el DNA d'aquesta regió en conformació Z, de manera que per desenrotllar la doble cadena de DNA i possibilitar la seva transcripció no és necessari el concurs d'altres enzims que catalitzen el tall, desenrotllamenet i unió del DNA, amb el conseqüent estalvi energètic per a la cèl·lula. D'altra banda, l'absència d'aquesta illa en la regió promotora

del trànscrit MNBHa és també consistent amb les característiques dels gens amb expressió específica de teixit.

L'exó 3 de MNBH és en la seva totalitat part d'una sequència Alu Jo (Fig. 16).

BLASTN Search Results with RepeatMasker filtering, Entrez and SRS links BLAST search performed using the National Center for Biotechnology Information's BLAST WWW Ésserver. Repeat reporting and filtering performed by RepeatMasker from U. Washington. RepeatMasker repeats found in sequence: SW % % % name query position in query (bp left) name matching type of repeat 631 15.6 0.0 2.8 Exon 2+3+4 65 173 (86) C AluJo/FRAM SINE/Alu pos. in repeat (bp left) (39) 263 158 Masked Sequence: >Exon 2+3+4 of MNBHa GGGAGCGCGCGCGGGGAGCCCGAGGCTGAGACTCACCGGAGGAAGCGGCG NNNNNNNNNNNNNNNNNTGTTATAGTTTTGCCGCTGGACTCTTC CCTCCCTTCCCCCACCCCATCAGGATGATATGAGACTTGAAAAGAAGACGA TGCATACAG Summary: file name: /genome/www/cgi-bin/tmp/ sequences: 259 bp 59.46 % total length: GC level: bases masked 109 bp (42.08 %) number of length percentage occupied of sequence elements* 109 bp SINES: 1 42.08 % 109 bp 42.08 % 1 ALUS ō 0 bp MIRS 0.00 % LINEs: 0 0 bp LINE1 0 0 bp 0.00 0 0 bp 0.00 % LINE2 0 0 bp LTR elements: 0.00 % 0 bp 0 bp 0.00 % MaLRs 0 Retrov. 0 0 bp MER4_group 0 0.00 % 0 0 bp 0.00 % DNA elements: MER1 type MER2 type 0 bp 0.00 % 0 0 bp 0 bp 0.00 % 0 õ Mariners 0 0 bp 0.00 % Unclassified: . Total interspersed repeats: 109 bp 42.08 % Small RNA: 0 0 bp 0.00 % Simple repeats: 0 0 bp 0.00 % 0.00 % 0 bp 0 Low complexity: _____ The sequence(s) were assumed to be of primate origin. RepeatMasker version 09/19/97 ProcessRepeats version 12/15/97 default Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402. Query= Exons 2+3+4 of MNBHa

(259 letters)

Database: Non-redundant GenBank+EMBL+DDBJ+PDB sequences

Figura 16. Resultat de l'estudi dels exons 2, 3 i 4 del trànscrit MNBHa amb el programa BLASTN amb emmascarament de les sequències repetitives. La lletra N representa quansevol nucleòtid que ha estat emmascarat.

L'inserció d'una sequència repetitiva del tipus Alu en el cDNA d'un gen no és un esdeveniment habitual, sinó més aviat bastant excepcional, però no és el primer descrit (Mittaz et al., 1997).

La funció biològica de la presència d'un element Alu en el 5'-UTR d'un cDNA no és prou coneguda, però es creu que l'allargament de la seqüència en aquest extrem pot ajudar a tenir una transcripció més eficient, sobretot en aquells gens, l'expressió dels quals pot estar sota el control d'una illa CpG, superant els efectes d'inhibició de la transcripció que suposa les estructures secundàries de les seqüències *leader* riques en G i C (Kozac, 1987). De fet, la inserció d'aquest exó en el trànscrit MNBHa, allarga la mida del 5'-UTR de 148 a 249 pb, amb la qual cosa s'obté una mida insòlita respecte a la resta de gens (5%). Si tenim en compte que el 5'-UTR i el 3'-UTR de la seqüència coneguda de *Dyrk* comparteix el mateix grau de similitud amb *MNBH* que l'observada en la regió codificant, seria interessant saber com *Dyrk* soluciona aquest problema.

La part 3'-UTR del cDNA del gen MNBH presenta diversos motius ATTTA. Aquests motius també estan presents en el 3'-UTR dels seus homòlegs de rata i Drosophila. Es tracta d'elements de desestabilització de l'RNA missatger (Wilson i Treisman 1988; Asson-Batres *et al.*, 1994) i es troben en la regió 3'-UTR de gens que s'expressen dinàmicament i que estan implicats en un ràpid recanvi de l'mRNA. (Frankel *et al.*, 1991).

La comparació de l'estructura genòmica de *mnb* i *MNBH* mostra que la majoria dels límits entre exó-intró estan conservats en el domini catalític, amb l'excepció que el gen humà incorpora un i dos introns en els exons 4 i 8 de *mnb*, respectivament. Aquesta observació, juntament amb la gran similitud de seqüència de DNA entre els dos gens, fa pensar que provenen d'un mateix gen ancestral.

MNBH s'expressa en tots els teixits fetals assajats. En teixits adults s'expressa de manera ubiqua, amb una certa preponderància en teixit de l'SNC, però també s'expressa en teixit germinal com testicles i ovaris. *MNBH* s'expressa de manera molt feble en ronyò i fetge.

L'existència de splicings alternatius en els exons 13a, 13b i 13c del gen MNBH, determina quatre pautes possibles de lectura oberta. Cadascuna de les quatre proteïnes inferides, MNBH-iso1, MNBH-iso2, MNBH-iso3 i MNBH-iso4 tenen en la seva seqüència aminoacídica extrems C-terminals diferents. MNBH-iso1 és el més llarg i abundant i codifica una proteïna de 754 aminoàcids (Fig. 17), mentre que els altres tres cDNA s'han de considerar com trànscrits de presència molt baixa. La importància de cadascuna de les proteïnes inferides i la seva funció biològica encara resta per determinar.

Cal destacar que l'exó 6 del cDNA de MNBH-iso1 pot tenir una inserció de 27 pb que codifica per a 9 aminoàcids (VMPDIVMLQ), generat per *splicing* alternatiu de l'exó 6. Aquesta darrera variant del cDNA de MNBH-iso1 és més llarga (763 aminoàcids) però és la forma minoritària, representant l'1% de la població de cDNA de MNBH-iso1. Aquest *splicing* alternatiu de l'exó 6 també s'ha descrit per al gen *Dyrk* (Kentrup *et al.,* 1996), la qual cosa suggereix l'existència de sistemes de regulació semblants entre el gen humà i el de rata.



Figura 17. Pautes de lectura possibles de la seqüència del cDNA de *MNBH-iso1*. Solament la pauta 1 sembla una pauta de lectura oberta veritable. Cada ratlla vertical representa un triplet de parada o *stop codon* en la seqüència de l'RNA.

Anàlisis de transferència de Northern indiquen que el gen MNBH està sobreexpressat 1,5 vegades en l'SNC de fetus amb la DS, repecte als fetus sense la DS (Guimera et al., sotmés a publicació, veure aparat 4.1.4.). Aquests estudis mostren també com el gen homòleg de ratolí, Mnbh, també està sobreexpressat 2,1 vegades en el ratolí trisòmic parcial Ts65Dn. Cal esmentar que Mnbh s'ha localitzat entre els marcadors Cbr i Pcp-4 del genoma de ratolí, dins de la regió sintènica entre el cromosoma 16 de ratolí i l'HC21.

Estudis funcionals *in vivo* de dos Mb de la banda cromosòmica 21q22 de l'HC21 impliquen el gen *MNBH* en les alteracions d'aprenentatge i de memòria que mostren els ratolins transgènics amb un fragment telomèric de 180 kb del YAC 152F7 (Smith *et al.*, 1997). *MNBH* és l'únic gen caracteritzat en aquest fragment.

El fenotip que pot causar la sobreexpressió del gen *MNBH* en la DS encara no està clar, però atès el fenotip de *Drosophila* amb mutacions en el gen *mnb*, podem pensar que podria tenir un paper molt important també en la regulació de la neurogènesi en humans. La seva expressió aberrant podria afectar el desenvolupament de l'SNC, la qual cosa podria explicar la gran reducció (20-50%) de la densitat neuronal en l'escorça cerebral, cèl·lules granulars de l'hipocamp, cerebel i hipotàlem dels cervells dels individus amb la DS, i també una reducció extraordinària del volum del cerebel (hipoplasia). És precisament en aquestes parts de l'SNC on aquest gen s'expressa de manera més conspícua. La pèrdua difusa de les neurones granulars que s'observa en nounats amb la DS, fa pensar que els gens que causen les alteracions neurològiques s'expressen sobretot en edats primerenques del desenvolupament de l'SNC, com també és el cas del gen *MNBH*.

La proteïna inferida del trànscrit MNBH-iso1 té dos senyals de localització nuclear conservats en Drosophila, rata i humans:

1) Un primer motiu simple KKKRR, conservat en les tres quinases mnb, Mnbh i MNBH.

2) Un senyal bipartit.

Aquest darrer senyal no solament s'ha conservat en els tres gens mnb, Mnbh i MNBH, sinó que s'ha delecionat un triplet de la seqüència espaiadora típica de Drosophila [RR(N)₁₁HKKER], transformant-se en [RR(N)₁₀HKKER], que és una seqüència de localització nuclear pròpia dels vertebrats. Tot plegat suggereix que, aquests motius putatius de localització nuclear són realment funcionals. Darrers estudis de la localització de Mnbh corraboren aquesta hipòtesi (Song *et al.*, 1997; Francisco Tejedor, comunicació personal). Aquesta darrera senyal de localització nuclear és típica de factors de transcripció i de proteïnes involucrades en la transcripció o en la mitosis (Dingwall i Laskey, 1991).

Altres dominis de la proteïna MNBH són:

1) Un domini catalític amb 8 subdominis típics de les proteïnes quinases.
260 Resum i discussió -

2) Una cremallera de leucines, on els quatre residus leucina estan separats entre si per 7 aminoàcids. Moltes proteïnes reguladores, com c-Fos i c-Jun, tenen aquest motiu, el qual és crític per a la interacció entre proteïnes (Landschulz *et al.*, 1988) i per la seva habilitat d'unir-se al DNA (Ellenberger *et al.*, 1992; Okagami *et al.*, 1995).

3) Una repetició truncada d'histidines $(His)_4X_4(His)_{13}$ no polimòrfica, segons hem verificat en 40 individus no relacionats. Aquestes cues d'histidines es troben principalment en factors de transcripció.

4) Una regió rica en serines/treonines. Aquest motiu té moltes seqüències S/T-P i S/T-S/T, típic de proteïnes que interaccionen amb el solc menor de la doble hèlix del DNA (Suzuki, 1989).

Un altre motiu de la proteïna de MNBH es la regió PEST (motiu ric en prolina, àcid glutàmic, serina i treonina). Aquesta regió, que també està present en els seus homòlegs de rata i *Drosophila*, és també un motiu d'interacció proteïna-proteïna, i típica d'aquelles proteïnes que pateixen una degradació molt ràpida (Rogers *et al.*, 1986). Aquest motiu PEST juntament amb els motius desestabilitzadors de l'RNA, que també estan conservats en les tres espècies, suggereixen que la regulació d'aquest gen està molt controlada i que el recanvi o *turnover* és molt ràpid, al igual que s'observa en aquells gens implicats en el cicle cel·lular.

Tots aquests motius suggereixen que la proteïna MNBH pot formar part d'un complex proteic en el nucli i pot estar implicada en la senyalització i/o en la transcripció cel·lular. Aquest complex proteic podria estar format per unitats de la proteïna MNBH i per altres proteïnes amb les que MNBH podria interaccionar. Si aquesta hipòtesi fos certa, podriem contemplar la possibilitat que el fenotip motivat per una subexpressió del locus minibrain pugui ésser molt similar al fenotip de la seva sobreexpressió. D'aquesta manera, podria ésser que el més important pel fenotip minibrain no sigui la seva sobreexpressió o subexpressió, sinó un canvi en l'expressió d'aquest. Canvis en l'expressió de MNBH podrien modificar l'estequiometria entre les diferents proteïnes que formen el possible complex proteic, causant el mateix fenotip. En essència, aquesta especulació admet que causes diferents (sobreexpressió i subexpressió d'un gen) podrien provocar un efecte similar. Cal tenir present que el fenotip observat en Drosophila està causat per una reducció de l'expressió de mnb, puix que hi ha una sola còpia gènica, mentre que en la DS el gen MNBH es troba sobreexpressat, puix que hi ha tres còpies gèniques, i en tots dos casos, s'observa una reducció de la mida i una hipocel·lularitat de certes estructures de l'SNC, segons s'ha esmentat anteriorment. Tanmateix, quan la sobreexpressió del gen *MNBH* es troba en conjunció amb l'expressió alterada d'altres gens com en el ratolí amb trisomia total Ts65Dn, o de forma individual com en el ratolí transgènic de la part telomèrica pel YAC 152F7, aquests no mostren diferències morfològiques macroscòpiques de l'SNC, contràriament al fenotip observat en el mutant *minibrain* de *Drosophila*, en el ratolí trisòmic total Ts16 i en la DS.

Una possible explicació a aquesta observació podria ésser que l'impacte fenotípic d'un gen pot variar entre les diferents espècies. També hem de tenir en compte que per a que un gen pugui motivar una patologia determinada, cal que aquest gen s'expressi en el moment i en el teixit escaient. D'aquesta manera, l'expressió temporal i espaial de MNBH en els diferents models pot no ésser exactament la mateixa. Una altra explicació podria ésser que la trisomia del ratolí transgènic per la part telomèrica del YAC 152F7 sigui tan petita, que malgrat sobreexpressar el gen MNBH, no sobreexpressa altres gens de l'HC21 que podrien interaccionar amb MNBH, conformant així un impacte fenotípic similar, però no idèntic. Aquesta hipòtesi pot estar recolçada per l'observació que, contràriament al que s'observa en els cervells de les persones amb la DS, cap dels models de ratolins amb trisomies parcials ni del transgènic esmentat anteriorment mostren alteracions cerebrals apreciables a escala macroscòpica, però que sí s'observen en el ratolí trisòmic total Ts16, on la interacció gènica de MNBH amb altres gens podria ésser més elevada. D'aquesta manera, tots aquests models ens poden suggerir que l'impacte fenotípic de l'expressió aberrant d'un sol gen com MNBH (com en el mutant mnb de Drosophila o en el ratolí transgènic per la regió telomèrica del YAC 152F7) o d'uns pocs gens, que inclou Mnbh (ratolí trisòmic Ts65Dn) pot ésser quelcom diferent de l'impacte fenotípic de la trisomia d'una regió cromosòmica sencera (com en el ratolí trisòmic Ts16 o en la DS), atesa per la interacció entre MNBH i la resta de gens presents en cadascun d'aquests models. Malgrat aquestes diferències morfològiques, tots aquests models fenotípics esmentats (i a on l'expressió de mnb/Mnbh/MNBH és aberrant) mostren alteracions molt similars en el comportament, que suggereixen que aquesta alteració conductual no sempre està associada a una alteració morfològica macroscòpica, i que les diferències entre aquestes darreres poden estar motivades per les diferents interaccions entre l'expressió aberrant de mnb/Mnbh/MNBH i la resta de background genètic propi de cada model. Cal fer esment, que els darrers estudis neuromorfològics dels ratolins Ts65Dn amb mètodes més precisos com l'estereologia, s'aprecien algunes diferències respecte als controls:

1) Un menor volum de la regió CA2 de l'hipocamp (Insausti et al., 1998).

2) Una disminució en el volum de l'escorça cerebral (Haydar et al., 1996).

3) Una disminució en la densitat cel·lular que afecta a diverses regions de l'SNC, com en el gir dentat (Insausti *et al.*, 1998).

Si puguèssim establir paral·lelismes entre la marcada disminució (20-50%) del nombre de neurones de l'SNC dels individus amb la DS i la gran disminció observada (40-70%) en el nombre de neurones de l'SNC del mutant minibrain de Drosophila. podriem pensar que la marcada disminució del nombre i complexitat de les arboritzacions i de les espines dendrítiques dels nounats amb la DS (fins i tot arriben a ésser atròfiques als quatre mesos de vida), es causada per a una reducció de sinapsis provocada per l'espaiament anormal de les cèl·lules nervioses, tal i com succeeix en la mosca minibrain. Aquesta reducció d'arboritzacions i dendrites, sinapsis i cèl·lules neuronals, sobretot en aquelles àrees de l'SNC humà que coincideixen amb aquelles àrees de l'SNC de ratolí on Mnbh s'expressa majoritàriament, juntament amb una comunicació interneural inadequada, podrien provocar algunes de les alteracions de memòria, atenció i coordinació motora observades en la DS. En aquest sentit, també podria explicar, en part, que el CI dels nounats amb aquesta símdrome decaigui dràsticament fins els tres anys, que el pes del cervell comenci a descendre a partir dels tres o sis mesos de vida i que la mida del cerebel sigui extraordinariament petita. És precisament en el cerebel a on l'expressió de MNBH és més intensa, tan en teixit fetal com en adult.

En el ratolí trisòmic parcial Ts65Dn també s'observa un dèficit d'atenció i de coordinació motora similar al de la DS, i també alteracions en els processos de memòria de treball i de curt termini, i d'aprenentatge (test del laberint aquàtic de Morris). Aquest fenotip és també el fenotip del ratolí transgènic amb la part telomèrica del YAC 152F7, on l'ùnic gen caracteritzat és el gen *MNBH*. Val a dir que la regió sintènica del fragment de DNA d'aquest YAC de l'HC21 es troba dindre la regió cromosòmica per la qual el ratolí Ts65Dn és trisòmic. És important destacar que el gen *MNBH* es troba dins la DSCR de l'HC21, una regió amb una mida estimada de 300-4.000 kb i que per sí sola, en condició de trisomia parcial, és capaç de motivar 14 dels 27 trets fenotípics de la DS, inclosa una greu discapacitat intel·lectual.

La comparació de l'estructura entre la proteïna Minibrain humana i la de Drosophila mostra que són molts els dominis funcionals conservats, com els dos senyals de localització nuclear, o cadascun dels subdominis catalítics de la quinasa, o la cremallera de leucines. Malgrat que no coneixem bé el paper del domini GAS, ni el de la repetició d'alanines ni el del domini ric en serines, presents en el gen *mnb* de *Drosophila*, la comparació de l'estructura primària de les tres proteïnes quinases suggereix una possible similitud de funció entre aquests dominis mencionats i els dominis PEST, repetició d'histidines i el domini ric en serines/treonines respectivament, conservats en rata i en humà. És interessant destacar que la seqüència GAS també s'ha trobat en proteïnes de la família de quinases sgg (Siegfried *et al.*, 1990; Bourouis *et al.*, 1990), que són necessàries pel creixement normal de la majoria dels teixits larvaris i per a la neurogenesi (Bourouis *et al.*, 1989).

Aquestes observacions, juntament amb el fenotip del mutant *mnb* de *Drosophila* i el seu possible paper en la neurogènesi postembrionària d'aquesta, la correlació observada entre el patró d'expressió de *Mnbh* en l'SNC del cervell de ratolí i aquelles àrees afectades en el cervell dels individus amb la DS, el fenotip del transgènic amb la part telomèrica del YAC 152F7 i la localització del gen *MNBH* en la DSCR del cromosoma 21, sustenten la hipòtesi de que *MNBH* és un excel·lent candidat per a explicar algunes de les alteracions observades en la DS, i que la seva sobreexpressió pot tenir conseqüències importants en la discapacitat intel·lectual de la DS.

També és dóna el cas que, MNBH és un ferm candidat com a possible gen responsable de la microcefàlia i el retard del creixement intrauterí del fetus, puix que MNBH es troba en la regió cromosòmica mínimament delecionada (D21S11-D21S55) en aquells individus amb aquestes alteracions associades a una monosomia parcial de l'HC21 (Matsumoto et al., 1997). Potser tampoc sigui una coincidència que els mutants minibrain de Drosophila pateixin un endarreriment en el creixement, i requereixen un 10% més de temps pel seu desenvolupament (Heidenreich, 1982). En aquests dos casos, el gen es troba en condició d'una sola còpia gènica.

Per tal de relacionar millor la funció de MNBH en la neurogènesi, caldrà esperar a obtenir els resultats de les anàlisis dels ratolins transgènics que sobreexpressen el gen MNBH i dels ratolins amb aquest gen anul·lat o knock out del gen MNBH. D'aquesta manera, podrem saber si MNBH pot ésser la connexió molecular entre canvis en la dosi gènica i algunes de les alteracions del desenvolupament de l'SNC observats en la DS, i la seva possible relació amb la discapacitat intel·lectual.

L'estudi de la naturalesa genètica i molecular de gens com MNBH ha de permetre conèixer millor els mecanismes que controlen la proliferació i diferenciació cel·lular durant el desenvolupament normal del sistema nerviós humà i la seva implicació en la patogènesi en la DS. Finalment, la caracterització de cadascun dels gens del cromosoma 21 ens permetrà estudiar la sensibilitat de la seva dosi d'expressió, ja sigui de manera individual o bé de manera conjunta amb d'altres gens, a més de la interacció entre cadascuna de les proteînes per a les que codifiquen i la seva interacció amb l'expressió dels productes gènics de la resta del genoma. El coneixement sobre la DS avança de mica en mica, aportant importants informacions sobre aspectes de la fisiologia del nostre organisme i la complexitat de les alteracions de l'HC21.

6. CONCLUSIONS

•

•1

1. S'ha construït una genoteca de 502 cosmidis a partir de cinc YAC (72H9, 336G11, 238B1, 221B9 i 552A3) específics de tres regions del cromosoma 21 humà: 1) una regió que inclou la regió crítica de la síndrome de Down, 2) una regió més distal, i 3) una altra proximal, representant una longitud cromosòmica total de 2,2 Mb. Aquests cosmidis s'han ordenat seqüencialment en un nou mapa genòmic contigu.

2. S'ha aplicat la tècnica de selecció de cDNA a les tres regions esmentades del cromosoma 21 humà, utilitzant una genoteca de cDNA de cervell fetal humà. S'han identificat, caracteritzat i analitzat 120 EST, dels quals 45 són no redundants i 37 són EST inèdits. Tots ells s'han integrat en un mapa transcripcional.

3. S'ha expandit i caracteritzat tota la sequència gènica d'un dels 37 nous EST aïllats. Per la seva similitut amb el gen *minibrain* de *Drosophila* l'hem anomenat *Minibrain* (*MNBH*). El gen humà *MNBH* és un gen de còpia única, abarca 150 kb de DNA genòmic i consta de disset exons, es localitza dins la regió crítica de la síndrome de Down del cromosoma 21 humà, en la banda 21q22.2, entre els marcadors de DNA *D21S335* i *D21S337*, a 180 kb del marcador de referència *D21S55*.

4. Segons quin sigui l'inici de transcripció del gen *MNBH*, s'ha aïllat i caracteritzat dues variants transcripcionals: 1) MNBHa, té l'inici de la transcripció a l'exó 2, la seva expressió és ubíqua, tant en teixits fetals com en adults, i la seva expressió pot estar controlada per una illa CpG, i 2) MNBHb, té l'inici de transcripció a l'exó 1, s'expressa en cor i múscul esquelètic, però no en cervell, i la seva expressió pot estar sota la influència d'una caixa TATA-like.

5. S'han caracteritzat quatre *splicings* alternatius de la regió codificant de *MNBH* que donen lloc a quatre pautes de lectura oberta: MNBH-iso1, MNBH-iso2, MNBH-iso3 i MNBH-iso4. La variant transcripcional MNBH-iso1 és la forma majoritària.

6. L'anàlisi bioinformàtic de la proteïna MNBH mostra dos possibles senyals de localització nuclear, un domini serina-treonina quinasa, una regió PEST, una cremallera de leucines, una repetició de tretze histidines, i una regió rica en serines-treonines. La comparació d'aquests possibles dominis funcionals de la proteïna MNBH amb les seves homòlogues de *Drosophila* i de rata, mostra un elevat grau de similitud.

7. El gen murí *Mnbh* s'expressa a l'hipotàlem, hipocamp, escorça cerebral, bulb ofactori i cerebel del cervell adult de ratolí.

8. L'expressió del gen MNBH està 1,5 vegades sobreexpressat en la síndrome de Down i 2,1 vegades en el ratolí Ts65Dn.

£.

Referències 271

А

Aaltonen J, Bjorses P, Sandkuijl L, Perheentupa J and Peltonen L, (1994). An autosomal locus causing autoimmune disease: autoimmune polyglandular disease type I assigned to chromosome 21. Nat Genet., 8: 83-87.

Abuelo D, Barsel-Bowers G, Busch W, Pueschel S and Pezzullo J, (1986). Risk for trisomy 21 in offspring of individuals who have relatives with trisomy 21. Am J Med Genet., 25: 365-367.

Adams MD, Dubnick M, Kerlavage AR, Moreno R, Kelley JM, Utterback TR, Nagle JW, Fields C and Venter JC, (1992). Sequence identification of 2,375 human brain genes. *Nature*, **355**: 632-634.

Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O, (1995). Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature*, **377**: 3-174.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ, (1990). Basic local alignment search tool. J Mol Biol., 215: 403-410.

Antonarakis SE, Adelsberger PA, Petersen MB, Binkert F, Schinzel AA, (1990). Analysis of DNA polymorphisms suggests that most *de novo* dup(21q) chromosomes in patients with Down syndrome are isochromosomes and not translocations. *Am J Hum Genet.*, **47**: 968-972.

Antonarakis SE, Avramopoulos D, Blouin J-L, Talbot CC and Schinzel AA, (1993). Mitotic errors in somatic cells cause trisomy 21 in about 4.5% of cases and are not associated with advanced maternal age. *Nature Genet.*, **3**: 146-150.

Antonarakis SE, (1996). Chapter for Textbook of Molecular Medicine. Down syndrome. New York. Jameson JL and Jabs EW (eds). Blackwell Science Inc.

Arnaout MA, Dana N, Gupta SK, Tenen DG and Fathallah DM, (1990). Point mutations impairing cell surface expression of the common beta subunit (CD18) in a patient with leukocyte adhesion molecule (Leu-CAM) deficiency. J Clin Invest., 85: 977-981.

Arnheim N, Krystal M, Schmickel R, Wilson G, Ryder O and Zimmer E, (1980). Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc Natl Acad Sci U S A, 12: 7323-7327.

Asson-Batres MA, Spurgeon SL, Diaz J, DeLoughery TG and Bagby, GC Jr, (1994). Evolutionary conservation of the AU-rich 3' untranslated region of messenger RNA. Proc Natl Acad Sci U S A, 91: 1318-1322.

Auch D and Reth M, (1990). Exon trap cloning: using PCR to rapidly detect and clone exons from genomic DNA fragments. Nucleic Acids Res., 18: 6743-6744.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K, (1994). Current Protocols in Molecular Biology. 6.10 (Suppl. 20).

Avraham KB, Sugarman H, Rotshenker S and Groner Y, (1991). Down's syndrome: morphological remodelling and increased complexity in the neuromuscular junction of transgenic CuZn-superoxide dismutase mice. J Neurocytol., 20: 208-215.

Avramopoulos D, Cox T, Kraus JP, Chakravarti A and Antonarakis SE, (1993). Linkage mapping of the cystathionine beta-synthase (*CBS*) gene on human chromosome 21 using a DNA polymorphism in the 3' untranslated region. *Hum Genet.*, 90: 566-568.

B

Banchs I, Bosch A, Guimera J, Lazaro C, Puig A, Estivill X, (1994). New alleles at microsatellite loci in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell lines. *Hum Mutat.*, **3**: 365-372.

Becker L, Mito T, Takashima S and Onodera K, (1991). Growth and Development of the Brain in Down Syndrome. In *The Morphogenesis of Down Syndrome* The Wiley-Liss NY. *Prog Clin Biol Res.*, 373, 133-153.

Beckman JS and Weber JL, (1992). Survey of human and rat microsatellites. Genomics, 12: 627-631.

Berr C, Borghi E, Rethore MO, Lejeune J and Alperovitch A, (1990). Risk of Down syndrome in relatives of trisomy 21 children. A case-control study. Ann Genet., 33: 137-140.

Bimboim HC and Doly J, (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl Acids Res., 7: 1513-1523.

Bird AF, (1986). CpG-rich islands and the function of DNA methylation. Nature, 321: 209-213.

Bogart MH, (1992). Current and future modes of prenatal diagnosis, In Lott IT and McCoy EE (eds): Down Syndrome: Advances in Medical Care. New York, Wiley-Liss, p 13.

Bosch A, Nunes V, Patterson D and Estivill X, (1993a). Isolation and characterization of 14 CA-repeat microsatellites from human chromosome 21. Genomics 18: 151-155.

Bosch A, Wiemann S, Guimera J, Ansorge W, Patterson D and Estivill X, (1993b). Two dinucleotide repeat polymorphisms at 21q22.3 (D21S416 and D21S1235). Hum Mol Genet., 2: 1744.

Bosch A, Guimera J, Pereira de Souza A, Estivill X, (1994a). The EUROGEM map of human chromosome 21. Eur J Hum Genet., 2: 244-245.

Bosch A, Wiemann S, Ansorge W, Patterson D and Estivill X, (1994b). Three CA/GT repeat polymorphisms from loci D21S414 and D21S1234 on human chromosome 21. Hum Genet., 93: 359-360.

Bosch A, (1994c). Identificació i caracterització de nous marcadors de DNA hipervariables i construcció d'un mapa integrat del cromosoma 21 humà. Tesi doctoral, Universitat Autònoma de Barcelona.

Referències 273

Bosch A, Guimera J, Patterson D and Estivill X, (1995a). Characterisation of three microsatellite polymorphisms (D21S1262, D21S1419 and D21S1421) from band 21q22.1. Hum Genet., 95: 596-598.

Bosch A, Guimera J, Wiemann S, Ansorge W, Patterson D and Estivill X, (1995b). Identification of two highly polymorphic CA-repeats (*D21S1224* and *D21S1261*) on human chromosome 21q22.3. *Hum Genet.*, 95: 367-369.

Bosch A, Wiemann S, Guimera J, Ansorge W, Patterson D and Estivili X, (1995c). Five new microsatellite polymorphisms at the q21 region of human chromosome 21. Hum Genet., 95:119-122.

Bosch A, Guimera J, Graw S, Gardiner K, Chumakov I, Patterson D and Estivill X, (1996). Integration of 30 CA-repeat markers into the cytogenetic, genetic and YAC maps of human chromosome 21. Eur J Hum Genet., 4: 135-142.

Bourouis M, Heitzler P, el Messal M and Simpson P, (1989). Mutant Drosophila embryos in which all cells adopt a neural fate. Nature, 341: 442-444.

Bourouis M, Moore P, Ruel L, Grau Y, Heitzler P and Simpson P, (1990). An early embryonic product of the gene shaggy encodes a serine/threonine protein kinase related to the CDC28/cdc2+ subfamily. *EMBO J*, 9: 2877-2884.

Buckler AJ, Chang DD, Graw SL, Brook JD, Haber DA, Sharp PA and Housman DE, (1991). Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc Natl Acad Sci USA.*, 88: 4005-4009.

Burmeister M, Kim S, Price ER, de Lange T, Tantravahi U, Myers RM and Cox DR, (1991). A map of the distal region of the long arm of human chromosome 21 constructed by radiation hybrid mapping and pulsed-field gel electrophoresis. *Genomics*, 9: 19-30.

Busciglio J and Yankner BA, (1995). Apoptosis and increased generation of reactive oxygen species in Down's syndrome neurons in vitro. *Nature*, **378**: 776-779.

Bustin M, Alfonso PJ, Pash JM, Ward JM, Gearhart JD and Reeves RH, (1995). Characterization of transgenic mice with an increased content of chromosomal protein HMG-14 in their chromatin. DNA Cell Biol., 14: 997-1005.

Butler R, Ogilvie DJ, Elvin P, Riley JH, Finniear RS, Slynn G, Morten JE, Markham AF, Anand R, (1992). Walking, cloning, and mapping with yeast artificial chromosomes: a contig encompassing D21S13 and D21S16. Genomics, 12: 42-51.

С

Cabin DE, Hawdins A, Griffin C and Reeves RH, (1995). YAC transgenic mice in the study of the genetic basis of Down syndrome. En: Epstein; CJ, Hassold, T, Lott, IT, Nadel, L, Patterson, D, eds. *Etiology and Pathogenesis of Down Syndrome*. Wiley-Lyss, New York: 213-226.

274 Referències

Calonge MJ, Gasparini F, Chillaron J, Chillon M, Gallucci M, Rousaud F, Zelante L, Testar X, Dallapiccola B, Di Silverio F, Falacin M, (1994). Cystinuria caused by mutations in *rBAT*, a gene involved in the transport of cystine. Nat Genet., 6: 420-425.

Carr J, (1994). Long-term-outcome for people with Down's syndrome. J Child Psychol Psychiatry, 35: 425-439.

Chaïb H, Kaplan J, Gerber S, Vincent C, Ayadi H, Slim R, Munnich A, Weissenbach J and Petit C, (1997). A newly identified locus for Usher syndrome type I, USH1E, maps to chromosome 21q21. Hum Mol Genet., 6: 27-31.

Chartier-Harlin MC, Crawford F, Houlden H, Warren A, Hughes D, Fidani L, Goate A, Rossor M, Roques P, Hardy J, (1991). Early-onset Alzheimer's disease caused by mutations at codon 717 of the betaamyloid precursor protein gene. *Nature*, 353: 844-846.

Chen H, Chrast R, Rossier C, Morris MA, Lalioti MD and Antonarakis SE, (1996). Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping. *Genome Res.*, 6: 747-760.

Cheng JF, Boyartchuk V and Zhu Y, (1994). Isolation and mapping of human chromosome 21 cDNA: progress in constructing a chromosome 21 expression map. *Genomics*, 23: 75-84.

Chomczynski F and Sacchi N, (1987). Single-step method of RNA isolation by acid guanidine thiocyanate-phenol-chloroform extraction. Anal Biochem., 162: 156-159.

Choo KH, Vissel B, Brown R, Filby RG and Earle E, (1988). Homologous alpha satellite sequences on human acrocentric chromosomes with selectivity for chromosomes 13, 14 and 21: implications for recombination between nonhomologues and Robertsonian translocations. *Nucleic Acids Res.*, 16: 1273-1284.

Chumakov IM, Le Gall I, Billault A, Ougen P, Soularue P, Giullou S, Rigault P, Bui H, De Tand M-F, Barillot E, Abderrahim H, Cherif D, Berger R, Le Paslier D and Choen D, (1992a). Isolation of chromosome 21-specific yeast artificial chromosomes from a total human genome library. Nat Genet., 1: 222-225.

Chumakov I, Rigault P, Guillou S, Ougen P, Billaut A, Guasconi G, Gervy P, LeGall I, Soularue P, Grinas L, Bougueleret L, Bellanné-Chantelot C, Lacroix B, Barillot E, Gesnouin P, Pook S, Vaysseix G, Frelat G, Schmitz A, Sambucy JL, Bosch A, Estivill X, Weissenbach J, Vigna, LA, Riethman H, Cox D, Patterson D, Gardiner K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Le Paslier D, Heilig R, Antonarakis S and Cohen D, (1992b). Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature*, **359**: 380-387.

Chung CT, Niemela SL and Miller RH, (1989). One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci U S A, 86: 2172-2175.

Church DM, Stotler CJ, Rutter JL, Murrell JR, Trofatter JA and Buckler AJ, (1994). Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nat Genet.*, 6: 98-105.

Clark JM, (1988). Novel non-templated nucleotide addition reactions catalyzed byprocaryotic and eucaryotic DNA polymerases. Nucleic Acids Res., 16: 9677-9686.

Clarke PG, (1990). Developmental cell death: morphological diversity and multiple mechanisms. Anat Embryol., 181: 195-213.

Costa GL, Grafsky A and Weiner MP, (1994). Cloning and analysis of PCR-generated DNA fragments. PCR Methods Appl 3: 338-345.

Cox DR, Burmeister M, Price ER, Kim S and Myers RM, (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science*, **250**: 245-250.

Cuckle HS, Wald NJ and Lindenbaum RH, (1984). Maternal serum alpha-fetoprotein measurement: a screening test for Down syndrome. Lancet, 1: 926-929.

Cuckle HS, Wald NJ, Lindenbaum RH, Jonasson J, (1985). Amniotic fluid AFP levels and Down syndrome. Lancet, 1: 290-291.

 \mathcal{D}

Dahmane N, Ghezala GA, Gosset P, Chamoun Z, Dufresne-Zacharia MC, Lopes C, Rabatel N, Gassanova-Maugenre S, Chettouh Z, Abramowski V, Fayet E, Yaspo ML, Korn B, Blouin JL, Lehrach H, Poutska A, Antonarakis SE, Sinet PM, Creau N and Delabar JM, (1998). Transcriptional map of the 2.5-Mb CBR-ERG region of chromosome 21 involved in Down syndrome. Genomics, 48: 12-23.

Dalton AJ, Seltzer GB, Adlin MS and Wisniewski HM, (1993). Association between Alzheimer disease and Down syndrome-clinical obsevations. In Alzheimer Disease, Down Syndrome, and their Relationship, Oxford University Press, Oxford.

Daniel A, (1979). Normal phenotype and partial trisomy for the G positive region of chromosome 21. J Med Genet., 16: 227-229.

Deininger PL, Jolly DJ, Rubin CM, Friedmann T and Schmid CW, (1981). Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J Mol Biol.*, 151: 17-33.

Delabar JM, Theophile D, Rahmani Z, Chettouh Z, Blouin JL, Prieur M, Noel B and Sinet PM, (1993). Molecular mapping of twenty-four features of Down syndrome on chromosome 21. Eur J Hum Genet., 1: 114-124.

Dierssen M, Vallina IF, Baamonde C, García-Calatayud S, Lumbreras MA, and Florez J, (1997). Alterations of central noradrenergic transmision in Ts65Dn mouse, a model for Down syndrome. Brain Res., 749: 238-244.

Dierssen M, Vallina IF, Baamonde C, Lumbreras MA, Martinez-Cue C, Calatayud SG and Florez J, (1996). Impaired cyclic AMP production in the hippocampus of a Down syndrome murine model. Brain Res Dev Brain Res., 95: 122-124.

Dingwall C and Laskey RA, (1991). Nuclear targeting sequences -- a consensus?. Trends Biochem Sci, 16: 478-481.

Doll, (1992). Chromosome 21 Workshop N.3 Baltimore Abstr. p.7.

Down JLH, (1866). Observation on an ethnic classification of idiots. Clinical lectures and reports by the Medical and Surgical Staff of the London Hospital, 3: 259-262.

Dufresne-Zacharia MC, Dahmane N, Theophile D, Orti R, Chettouh Z, Sinet PM, Delabar JM, (1994). 3.6-Mb genomic and YAC physical map of the Down syndrome chromosome region on chromosome 21. Genomics, 19: 462-469.

Dutriaux A, Rossier J, Van Hul W, Nizetic D, Theophille D, Delabar JM, Van Broeckhoven C and Potier MC, (1994). Cloning and characterization of a 135- to 500-kb region of homology on the long arm of human chromosome 21. *Genomics*, 22: 472-477.

E

Edelman AM, Blumenthal DK and Krebs EG, (1987). Protein serine/threonine kinases. Annu Rev Biochem, 56: 567-613.

Edelman AM, Blumenthal DK, Krebs EG, (1987). Protein serine/threonine kinases. Annu Rev Biochem., 56: 567-613.

Eickbush TH and Moudrianakis EN, (1978). The compaction of DNA helices into either continuous supercoils or folded-fiber rods and toroids. Cell, 13: 295-306.

Eki T, Abe M, Furuya K, Ahmad I, Fujishima N, Kishida H, Shiratori A, Onozaki T, Yokoyama K, Le Paslier D, Cohen D, Hanaoka F and Murakami Y, (1996). A long-range physical map of human chromosome 21q22.1 band from the YAC continuum. *Mamm Genome*, **7**: 303-311.

Ellenberger TE, Brandl CJ, Struhl K, Harrison SC, (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell*, **71**: 1223-1237.

Elson A, Levanon D, Weiss Y and Groner Y, (1994). Overexpression of liver-type phosphofructokinase (PFKL) in transgenic-*PFKL* mice: implication for gene dosage in trisomy 21. *Biochem J.*, 299: 409-415.

Epstein CJ, Cox DR and Epstein LB, (1985). Mouse trisomy 16: an animal model of human trisomy 21 (Downsyndrome). Ann N Y Acad Sci., 450: 157-168.

Epstein CJ, (1986). Consequences of Chromosome Imbalance: Principles, Mechanisms, and Models. Cambridge Univ. Press, New York.

Epstein CJ, Korenberg JR, Anneren G, Antonarakis SE, Ayme S, Courchesne E, Epstein LB, Fowler A, Groner Y, Huret JL, (1991). Protocols to establish genotype-phenotype correlations in Down syndrome. *Am J Hum Genet.*, **49**: 207-235. Epstein CJ, (1995). Epilogue: toward the twenty-first century with Down syndrome, a personal view of how far we have come and of how far we can reasonably expect to go. *Prog Clin Biol Res.*, **393**: 241-246.

Epstein CJ, (1995b). The metabolic and molecular bases of interited disease, Scriver CR, Beaudet AL, Sly WS and Valle D (eds): Chapter 18, Down Syndrome. New York, Mc Graw-Hill, Vol. I. 7th edi.

Escorihuela RM, Fernandez-Teruel A, Vallina IF, Baamonde C, Lumbreras MA, Dierssen M, Tobena A and Flórez J, (1995). A behavioral assessment of Ts65Dn mice: a putative Down syndrome model. *Neurosci Lett.*, **199**: 143-146.

Esquirol JED, (1838). Des maladies mentales considérées sous les rapports médical, higiènique et médico-legal, Paris.

Estivill X and Williamson R, (1987). A rapid method to identify cosmids containing rare restriction sites. Nucleic Acids Res., 15: 1415-1425.

Eunpu DL, McDonald DM and Zackai EH, (1986). Trisomy 21: rate in second-degree relatives. Am J Med Genet., 25: 361-363.

Evans DI and Steward JK, (1972). Down's syndrome and leukaemia. Lancet, 2: 1322.

 \mathcal{F}

Feinberg AP and Vogelstein B, (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal biochem., 132: 6-13.

Ferrer J, Nichols CG, Makhina EN, Salkoff L, Bernstein J, Gerhard D, Wasson J, Ramanadham S and Permutt A, (1995). Pancreatic islet cells express a family of inwardly rectifying K+ channel subunits which interact to form G-protein-activated channels. J Biol Chem., 270: 26086-26091.

Flórez J, del Arco C, González A, Pascual J, Pazos A, (1990). Autoradiographic studies of neurotransmitter receptors in the brain of newborn infants with Down syndrome. Am J Med Genet., Suppl 7: 301-305.

Frankel AD, Mattaj IW and Rio DC, (1991). RNA-protein interactions. Cel 67: 1041-1046.

Freeman S, Grantham M, Hassold T, Herbert M, Hersey J, Muccio J, Petay D, Takaesu N and Phillips C, (1991). Cytogenetic and molecular studies of human spontaneous abortions. Am. J. Hum. Genet., Suppl 49, 916A.

Fryers T, (1986). Survival in Down'syndrome. J. Ment Defic Res., 30: 101.

Fuentes JJ, Banchs I, Volpini V and Estivill X, (1993). Genetic variation of microsatellite markers D1S117, D6589, D11S35, APOC2, and D21S168 in the Spanish population. Int J Legal Med., 105: 271-277.

Fuentes JJ, Pritchard MA and Estivill X, (1995). Genomic organization, alternative splicing, and expression patterns of the DSCR1 (Down syndrome candidate region 1) gene. Genomics, 44: 358-361.

Fuentes JJ, Pucharcos C, Fritchard M and Estivill X, (1997). Alu-splice FCR: a simple method to isolate exon-containing fragments from cloned human genomic DNA. Hum Genet., 101: 346-350.

G

Games D, Adams D, Alessandrini R, Barbour R, Berthelette P, Blackwell C, Carr T, Clemens J, Donaldson T, Gillespie F, (1995). Alzheimer-type neuropathology in transgenic mice overexpressing V717F betaamyloid precursor protein. *Nature*, **373**: 523-527.

Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S, Patterson D, (1990a). Analysis of human chromosome 21: correlation of physical and cytogenetic maps; gene and CpG island distributions. *EMBO J.*, **9**: 25-34.

Gardiner K, Watkins, P, Münke M, Drabkin H, Hones C and Patterson D, (1988). Fartial physical map of human chromosome 21. Somat. Cell Mol. Genet., 14: 623-638.

Gecz J, Villard L, Lossi AM, Millasseau P, Djabali M, Fontes M, (1993). Physical and transcriptional mapping of DXS56-PGK1 1 Mb region: identification of three new transcripts. *Hum Mol Genet.*, 2: 1389-1396.

Gerlai R and Roder J, (1995). Abnormal exploratory behavior in transgenic mice carrying multiple copies of the human gene for \$100 beta. J Psychiatry Neurosci., 20: 105-112.

Gingrich JC, Lowry SR, Kuo W-L, Gray J, Smith CL and Cantor CR, (1993). Cloning and characterization of Eagl YACs from human chromosome 21. Genomics, 15: 228-230.

Glenner GG and Wong CW, (1984). Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun.*, 120: 885-890.

Glücksmann A, (1951). Cell deaths in normal vertebrate ontogeny. Biol. Rev., 26: 59-86.

Goate A, chartier-Harlin M-C, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes AI, Nick J, Louise MR, Newton P, Rooke K, Roques P, Talbot C, Pericak-Vance M, Roses A, Williamson R, Rossor M, Owen M and Hardy J, (1991). Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*, **349**: 704-706.

Goss SJ and Harris H, (1975). New method for mapping genes in human chromosomes. Nature, 255: 680-684.

Gosset P, Crete N, Ait Ghezala G, Theophile D, Van Broeckhoven C, Vayssettes C, Sinet PM and Creau N, (1995). A high-resolution map of 1.6 Mb in the Down syndrome region: a new map between D21S55 and ETS2. Mamm Genome, 6: 127-130.

Groner Y, Elroy-Stein O, Avraham KB, Yarom R, Schickler M, Knobler H and Rotman G, (1990). Down syndrome clinical symptoms are manifested in transfected cells and transgenic mice overexpressing the human Cu/Zn-superoxide dismutase gene. J Physiol., 84: 53-77.

Groner Y, (1995). Transgenic models for chromosome 21 gene dosage effects. Prog Clin Biol Res., 393: 193-212.

Guimera J, Casas C, Pucharcos C, Solans A, Domenech A, Planas AM, Ashley J, Lovett M, Estivill X and Pritchard MA, (1996). A human homologue of Drosophila minibrain (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. *Hum Mol Genet.*, 5: 1305-1310.

Guimera J, Pritchard M, Nadal M and Estivill X, (1997b). Minibrain (MNBH) is a single copy gene mapping to human chromosome 21q22.2. Cytogenet Cell Genet., 77: 182-184.

Guimera J, Pucharcos C, Domenech A, Casas C, Solans A, Gallardo T, Ashley J, Lovett M, Estivill X and Pritchard M, (1997). Cosmid contig and transcriptional map of three regions of human chromosome 21q22: identification of 37 novel transcripts by direct selection. *Genomics*, **45**: 59-67.

 \mathcal{H}

Haddow JE, Palomaki GE, Knight GJ, Canick JA, Wald NJ and Cuckle HS, (1990). Maternal serum unconjugated estriol levels are lower in the presence of fetal Down syndrome. Am J Obstet Gynecol., 163: 1372-1374.

Hamers AJ, Meyer H, Jongbloed RJ, van der Hulst RR and Geraedts JP, (1990). Rate of recombination of chromosomes 21 in parents of children with Down syndrome. *Clin Genet.*, **37**: 463-469.

Hamerton J, (1971). Human Cytonegetics. New York, Academic, vol I.

Hanks SK, Quinn AM and Hunter T, (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**: 42-52.

Hartenstein V and Campos Ortega JA, (1984). Early neurogenesis in wild type Drosophila melanogaster. Roux's Arch Dev Biol., 193: 308-325.

Hassold T, Chiu D, (1985). Maternal age-specific rates of numerical chromosome abnormalities with special reference to trisomy. *Hum Genet.*, 70: 11-17.

Hassold TJ and Jacobs PA, (1984). Trisomy in man. Annu Rev Genet., 18: 69-97.

Hawley RS, Frazier J and Rassoly R, (1994). Separation anxiety: the biology of non-disjunction in flies and people. Hum. Mol. Genet., 3: 1521-1528.

Haydar TF, Blue ME, Molliver ME, Krueger BK and Yarowsky PJ J, (1996). Consequences of trisomy 16 for mouse brain development: corticogenesis in a model of Down syndrome. *Neurosci.*, 16: 6175-6182.

Heidenreich D, (1982). Die Genetik der Mutante minibrain (mnb) von Drosophila melanogaster. Diplomthesis, University of Würzburg.

Hendriks L, van Duijn CM, Cras F, Cruts M, Van Hul W, van Harskamp F, Warren A, McInnis MG, Antonarakis SE and Martin JJ, (1992). Presenile dementia and cerebral haemorrhage linked to a mutation at codon 692 of the beta-amyloid precursor protein gene. *Nat Genet.*, 1: 218-221.

Hook EB, (1978). Spontaneous deaths of fetuses with chromosomal abnormalities diagnosed prenatally. N Engl J Med., 299: 1036-1038.

Hook EB, (1982). Epidemiology of Down syndrome, in Pueschel SM, Rynders JE (eds): Down Syndrome. Advances in Biomedicine and the Behavioral Sciences. Cambridge, MA, Ware, p11.

Hook EB, Mutton DE, Ide R, Alberman E, Bobrow M, (1995). The natural history of Down syndrome conceptuses diagnosed prenatally that are notelectively terminated. *Am J Hum Genet.*, **57**: 875-881.

Hubert RS, Mitchell S, Chen XN, Ekmekji K, Gadomski C, Sun Z, Noya D, Kim UJ, Chen C, Shizuya H, Simon M, de Jong FJ and Korenberg JR, (1997). BAC and PAC contigs covering 3.5 Mb of the Down syndrome congenital heart disease region between *D21S55* and *MX1* on chromosome 21. *Genomics*, **41**: 218-226.

Hunter T, (1991). Protein kinase classification. Methods Enzymol, 200: 3-37.

Hwu HR, Roberts JW, Davidson EH and Britten RJ, (1986). Insertion and/or deletion of many repeated DNA sequences in human and higher ape evolution. *Proc Natl Acad Sci USA*, 83: 3875-3879.

I

Ichikawa H, Hosoda F, Arai Y, Shimizu K, Ohira M and Ohki M, (1993). A Notl restriction map of the entire long arm of human chromosome 21. Nat Genet., 4: 361-366.

Insausti AM, Megías M, Crespo D, Cruz-OriveKM, Dierssen M, Vallina F, Insausti R and Flórez J, (1998). Hippocampal volume and neuronal number in Ts65Dn mice: a murine model of Down syndrome. Neuroscience Letters, 253: 1-4.

J

Jacobs, PA, Baikie, AG, Court-Bourn, WM and Strong, JA, (1959). The somatic chromosomes in mongolism. Lancet i, 710-711.

K

Kemper TL, (1989). Neuropathology of Down Syndrome, In The psychobiology of Down Syndrome (ed Nadel, L.). The MIT Press pp.269-289.

Kentrup H, Becker W, Heukelbach J, Wilmes A, Schurmann A, Huppertz C, Kainulainen H and Joost HG, (1996). Dyrk, a dual specificity protein kinase with unique structural features whose activity is dependent on tyrosine residues between subdomains VII and VIII. J Biol Chem., 271: 3488-3495.

Kesslak JP, Nagata SF, Lott I and Nalcioglu O, (1994). Magnetic resonance imaging analysis of agerelated changes in thebrains of individuals with Down's syndrome. *Neurology*, **44**: 1039-1045.

Korenberg JR, (1991). Down syndrome phenotypic mapping, in Esptein CJ (ed): The Morphogenesis of Down Syndrome. New York, Wiley-Liss, p 43.

Korenberg JR, Bradley C and Disteche CM, (1992). Down syndrome: molecular mapping of the congenital heart disease and duodenal stenosis. Am J Hum Genet., 50: 294-302.

Korenberg JR, Chen XN, Schipper R, Sun Z, Gonsky R, Gerwehr S, Carpenter N, Daumer C, Dignan P, Disteche C, (1994). Down syndrome phenotypes: the consequences of chromosomal imbalance. Proc Natl Acad Sci USA., 91: 4997-5001.

Korenberg JR, (1995). Mental modelling. Nat Genet., 11: 109-111.

Kozak M, (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15: 8125-8148.

Kwiatkowski DJ, Henske EP, Weimer K, Ozelius L, Gusella JF and Haines J, (1992). Construction of a GT polymorphism map of human 9q. Genomics, 12: 229-240.

L

Lafrenière RG, Rochefort DL, Chrétien N, Rommens JM, Cochius JI, Kälviäinen R, Nousiainen U, Patry G, Farrel K, Söderfeldt B, Federico A, Hale BR, Hernandez Cossio O, Sorensen T, Pouliot MA, Kmiec T, Uldall P, Janszky J, Pranzatelli MR, Andermann F, Andermann E and Rouleau GA, (1997). Unsteble insertion in the 5' flanking region of the cystatin B gene is the most common mutation in progressive myoclonus epilepsy type 1, EPM1. *Nature Genet.*, **15**: 298-302.

Lamb BT, Sisodia SS, Lawler AM, Slunt HH, Kitt CA, Karns WG, Pearson PL, Price DL and Gearhart JD, (1993). Introduction and expression of the 400 kilobase amyloid precursor protein gene in transgenic mice. *Nat Genet.*, 5: 22-30.

Landschulz WH, Johnson PF and McKnight SL, (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*, **240**: 1759-1764.

Lapenta V, Sossi V, Gosset P, Vayssettes C, Vitali T, Rabatel N, Tassone F, Blouin JL, Scott HS, Antonarakis SE, Creau N and Brahe C, (1998). Construction of a 2.5-Mb integrated physical and gene map of distal 21q22.3. Genomics, 49: 1-13.

Lehesjoki AE, Koskiniemi M, Sistonen P, Miao J, Hastbacka J, Norio R and de la Chapelle A, (1991). Localization of a gene for progressive myoclonus epilepsy to chromosome 21q22. Proc Natl Acad Sci USA., 88: 3696-3699. Lejeune J, (1959). Le mongolisme, maladie chromosomique. La Nature, 3296: 521-523.

Levanon D, Danciger E, Dafni N andGroner Y, (1986). Genomic clones of the human liver-type phosphofructokinase. Biochem Biophys Res Commun., 141: 374-380.

Lindsay S and Bird AP, (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature*, **327**: 336-338.

Litt M and Luty JA, (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet., 44: 397-401.

Llull M, (1995). Del cromosoma al gen, de l'Institut d'Edicions de la Diputació de Barcelona. Consell genètic i prevenció de les anomalies cromosòmiques. Cap. 7, p 147-169.

Lovett M, Kere J and Hinton L.M, (1991). Direct selection: A method for the selection of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci. U.S.A.*, 88: 9628-9632.

Lovett M, (1994). Fishing for complements: finding genes by direct selection. Trends Genet., 10: 352-357.

Lucente D, Chen HM, Shea D, Samec SN, Rutter M, Chrast R, Rossier C, Buckler A, Antonarakis SE and McCormick MK, (1995). Localization of 102 exons to a 2.5 Mb region involved in Down syndrome. *Hum. Mol. Genet.*, 8 : 1305-1311.

М

Macri JN, Buchanan PD and Gold MP, (1986). Low alfa-fetoprotein and trisomy. Lancet, 2: 405.

Mann DM, (1988). Alzheimer's disease and Down's syndrome. Histopathology, 13: 125-137.

Marchuk D, Drumm M, Saulino A and Collins FS (1991). Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.*, 19: 1154.

Marin-Padilla M, (1976). Pyramidal cell abnormalities in the motor cortex of a child with Down's syndrome. A Golgi study. *J Comp Neurol.*, 167: 63-81.

Masumoto H, Masukata H, Muro Y, Nozaki N and Okazaki T, (1989). A human centromeric antigen (CENP-B) interacts with a short specific sequence in alphoid DNA of human centromeric satellite. *J. Cell. Biol.*, **109**: 1963-1973.

Matise TC, Perlin M and Chakravarti A, (1994). Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. Nat Genet., 6: 384-390.

Matsumoto N, Ohashi H, Tsukahara M, Kim KC, Soeda E and Niikawa N. (1997). Possible narrowed assignment of the loci of monosomy 21-associated microcephaly and intrauterine growth retardation to a 1.2-Mb segment at 21q22.2. Am J Hum Genet 60: 997-999.

Mazzoni DS, Ackley RS and Nash DJ, (1994). Abnormal pinna type and hearing loss correlations in Down's syndrome. J Intellect Disabil Res., 38: 549-560.

McInnis MG, Chakravarti A, Blaschak J, Petersen MB, Sharma V, Avramopoulos D, Blouin JL, Konig U, Brahe C, Matise TC, (1993). A linkage map of human chromosome 21: 43 PCR markers at average intervals of 2.5 cM. *Genomics*, **16**: 562-571.

Minc-Golomb D, Knobler H and Groner Y, (1991). Gene dosage of CuZnSOD and Down's syndrome: diminished prostaglandin synthesis in human trisomy 21, transfected cells and transgenic mice. *EMBO J.*, **10**: 2119-2124.

Mittaz L, Scott HS, Rossier C, Seeburg PH, Higuchi M and Antonarakis SE, (1997). Cloning of a human RNA editing deaminase (ADARB1) of glutamate receptors that maps to chromosome 21q22.3. *Genomics* **41**: 210-217.

Miyabara S, Gropp A, Winking H, (1982). Trisomy 16 in the mouse fetus associated with generalized edema and cardiovascular and urinary tract anomalies. *Teratology*, **25**: 369-380.

Miyoshi H, Shimizu K, Kozu T, Maseki N, Kaneko Y and Ohki M, (1991). t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, *AML1. Proc Natl Acad Sci USA.*, 88: 10431-10434.

Moran PM, Higgins LS, Cordell B and Moser PC, (1995). Age-related learning deficits in transgenic mice expressing the 751-amino acid isoform of human beta-amyloid precursor protein. *Proc Natl Acad Sci USA.*, 92: 5341-5345.

Morgan SB, (1979). Development and distribution of intellectual and adaptive skills in Down syndrome children: implications for early intervention. *Ment Retard.*, 17: 247-249.

Morgan JG, Dolganov GM, Robbins SE, Hinton LM and Lovett M, (1992). The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. *Nucleic Acids Res.*, 20, 5173-5179.

Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Gimenez J, Reis A, Varon-Mateeva R, Macek M Jr, Kalaydjieva L, (1994). The origin of the major cystic fibrosis mutation (delta F508) in European populations. Nat Genet, 7: 169-175.

Morton NE, (1991). Parameters of the human genome. Proc Natl Acad Sci USA., 88: 7474-7476.

Muenke M, Bone LJ, Mitchell HF, Hart I, Walton K, Hall-Johnson K, Ippel EF, Dietz-Band J, Kvaloy K, Fan CM, (1995). Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of *SIM2* as a candidate gene for holoprosencephaly, and mapping of *SIM2* to a region of chromosome 21 important for Down syndrome. *Am J Hum Genet.*, 57: 1074-1079.

Murrell J, Farlow M, Ghetti B and Benson MD, (1991). A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science*, 254: 97-99.

N

Nadel L and Epstein CJ (eds), (1992). Down syndrome and Alzheimer Disease. New York, Wiley-Liss.

Nicole A and Aurias A, (1990). Down syndrome critical region around D21S55 on proximal 21q22.3. Am J Med Genet Suppl., 7: 98-103.

Niebuhr E, (1974). Down's syndrome. The possibility of a pathogenetic segment on chromosome no. 21. Human genetik, 21: 99-101.

Nigg EA, (1995). Cyclin-dependent protein kinases: key regulators of the eukaryoti cell cycle. *Bioessays*, 17:471-480.

Nizetic D, Gellen L, Hamvas R, Mott R, Grigoriev A, Vatcheva R, Zehetner G, Yaspo ML, Dutriaux A, Lopes C, Delabar JM, Van Broeckhoven C, Potier MC and Lehrach H, (1994). An integrated YAC-overlap and 'cosmid-pocket' map of the human chromosome 21. Hum. Molec. Genet., 3: 759-770.

0

Ohira M, Ichikawa H, Suzuki E, Iwaki M, Suzuki K, Saito-Ohara F, Ikeuchi T, Chumakov I, Tanahashi H, Tashiro K and Sakaki Y, (1996). A 1.6-Mb P1-based physical map of the Down syndrome region on chromosome 21. Genomics, 33: 65-74.

Ohira M, Seki N, Nagase T, Suzuki E, Nomura N, Ohara O, Hattori M, Sakaki Y, Eki T, Murakami Y, Saito T, Ichikawa H and Ohki M, (1997). Gene identification in 1.6-Mb region of the Down syndrome region on chromosome 21. Genome Res., 7: 47-58.

Okagami M, Ueno M, Makino K, Shimomura M, Saito I, Morii T and Sugiura Y, (1995). Sequencespecific DNA binding by covalently constrained peptide dimers of the basic leucine zipper protein GCN4. *Bioorg Med Chem.*, 3: 777-784.

Olsen AS, Combs J, Garcia E, Elliott J, Amemiya C, Jong P and Threadgill G, (1993). Automated production of high density cosmid and YAC colony filters using a robotic workstation. *Product Application Focus*, 14: 116-123.

Osoegawa K, Susukida R, Okano S, Kudoh J, Minoshima S, Shimizu N, de Jong FJ, Groet J, Ives J, Lehrach H, Nizetic D and Soeda E, (1996). An integrated map with cosmid/PAC contigs of a 4-Mb Down syndrome critical region. *Genomics*, **32**: 375-387.

Overhauser J and Radic MZ, (1987). Encapsulation of cells in agarose beads for use with pulse-field gel electrophoresis. Focus, 9: 8-9.

P

Parsons R, Li GM, Longley MJ, Fang WH, Papadopoulos N, Jen J, de la Chapelle A, Kinzler KW, Vogelstein B, Modrich P, (1993). Hypermutability and mismatch repair deficiency in RER⁺ tumor cells. *Cell.*, **75**: 1227-1236.

Patil N, Peterson A, Rothman A, de Jong PJ, Myers RM and Cox DR, (1994). A high resolution physical map of 2.5 Mbp of the Down syndrome region on chromosome 21. Hum Mol Genet., 3: 1811-1817.

Patterson D, (1991). Report of the second international workshop on human chromosome 21 mapping. Cytogenet. Cell Genet, 57: 167-174.

Patterson D, Hart I, Lai LW, Brahe C, Moscetti A, Tassone F, Raimondi E and Jones C, (1993). Molecular and cytogenetic characterization of a Chinese hamster/human hybrid cell line containing a der (21)t(Ypter-->cenY::cen21-->21qter) chromosome. Genomics, 15: 177-179.

Pennacchio LA, Lehesjoki AE, Stone NE, Willour VL, Virtaneva K, Miao J, D'Amato E, Ramirez L, Faham M, Koskimiemi M, Warrington JA, Norio R, de la Chapelle A, Cox DR and Myers RM, (1996). Mutations in the gene encoding Cystatin B in progressive myoclonus epilepsy (EPM1). Science, 271: 1731-1734.

Petersen MB, Adelsberger PA, Schinzel AA, Binkert F, Hinkel GK, Antonarakis SE, (1991). Down syndrome due to de novo Robertsonian translocation t(14q;21q): DNA polymorphism analysis suggests that the origin of the extra 21q is maternal. Am J Hum Genet., 49: 529-536.

Petersen MB, Antonarakis SE, Hassold TJ, Freeman SB, Sherman SL, Avramopoulos D, Mikkelsen M, (1993). Paternal nondisjunction in trisomy 21: excess of male patients. *Hum Mol Genet.*, **2**: 1691-1695.

Peterson A, Fatil N, Robbins C, Wang L, Cox DR and Myers RM, (1994). A Transcript map of the Down syndrome critical region on chromosome 21. Hum. Molec. Genet, 10: 1735-1742.

Prasher VP, Farrer MJ, Kessling AM, Fisher EM, West RJ, Barber PC, Butler AC, (1998). Molecular mapping of Alzheimer-type dementia in Down's syndrome. Ann Neurol., 43: 380-383.

Q

Quon D, Wang Y, Catalano R, Scardina JM, Murakami K and Cordell B, (1991). Formation of betaamyloid protein deposits in brains of transgenic mice. Nature, 352: 239-241.

R

Radloff R, Bauer W and Vinograd J A, (1967). Dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA in HeLa cells. *Proc Natl Acad Sci U S A*, **57**: 1514-1521.

Rahmani Z, Blouin JL, Creau-Goldberg N, Watkins PC, Mattei JF, Poissonnier M, Prieur M, Chettouh Z, Nicole A, Aurias A, (1989). Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome. *Proc Natl Acad Sci USA.*, **86**: 5958-5962.

Raz N, Torres IJ, Briggs SD, Spencer WD, Thornton AE, Loken WJ, Gunning FM, McQuain JD, Driesen NR and Acker JD, (1995). Selective neuroanatomic abnormalities in Down's syndrome and their cognitive correlates: evidence from MRI morphometry. *Neurology*, 45: 356-366.

Reeves RH, Yao J, Crowley MR, Buck S, Zhang X, Yarowsky P, Gearhart JD and Hilt DC, (1994). Astrocytosis and axonal proliferation in the hippocampus of S100b transgenic mice. *Proc Natl Acad Sci USA*, **91**: 5359-5363.

Reeves RH, Irving NG, Moran TH, Wohn A, Kitt C, Sisodia SS, Schmidt C, Bronson RT and Davisson MT, (1995). A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nat Genet.*, 11: 177-184.

Rogers S, Wells R and Rechsteiner M, (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science, 234: 364-368.

Rosen DR, Siddique T, Fatterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX, (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*, **362**: 59-62.

Ross MH, Galaburda AM and Kemper TL, (1984). Down's syndrome: is there a decreased population of neurons? Neurology, 34: 909-916.

Rumble B, Retallack R, Hilbich C, Simms G, Multhaup G, Martins R, Hockey A, Montgomery P, Beyreuther K and Masters CL, (1989). Amyloid A4 protein and its precursor in Down's syndrome and Alzheimer's disease. N Engl J Med., 320: 1446-1452.

S

Sanger F, Nicklen S and Coulson AR, (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA, 74: 5463-5467.

Saunders JW (1996). Death in embryonic systems. Science 154: 604-612.

Schapiro MB, Luxenberg JS, Kaye JA, Haxby JV, Friedland RP and Rapoport SI, (1989). Serial quantitative CT analysis of brain morphometrics in adult Down's syndrome at different ages. *Neurology*, **39**: 1349-1353.

Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, Nemens E, White JA, Bonnycastle L, Weber JL, Alonso ME, (1992). Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. *Science*, **258**: 668-671.

Sertie AL, Quimby M, Moreira ES, Murray J, Zatz M, Antonarakis SE and Passos-Bueno MR, (1996). A gene which causes severe ocular alterations and occipital encephalocele (Knobloch syndrome) is mapped to 21q22.3. *Hum Mol Genet.*, **5**: 843-847.

Shaffer LG, Jackson-Cook CK, Meyer JM, Brown JA, Spence JE, (1991). A molecular genetic approach to the identification of isochromosomes of chromosome 21. *Hum Genet.*, 86: 375-382.

Shaffer LG, Jackson-Cook CK, Stasiowski BA, Spence JE, Brown JA, (1992). Parental origin determination in thirty de novo Robertsonian translocations. Am J Med Genet., 43: 957-963.

Shapiro BL, (1994). The environmental basis of the Down syndrome phenotype. Dev Med Child Neurol., 36: 84-90.

Shapiro BL, (1997). Whither Down syndrome critical regions?. Hum Genet., 99: 421-423.

Sherman SL, Petersen MB, Freeman SB, Hersey J, Pettay D, Taft L, Frantzen M, Mikkelsen M and Hassold TJ, (1994). Non-disjunction of chromosome 21 in maternal meiosis I: evidence for a maternal agedependent mechanism involving reduced recombination. *Hum Mol Genet.*, **3**: 1529-1535.

Sherr CJ, (1993). Mammalian G1 cyclins. Cell, 73: 1059-1065.

Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, (1995). Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*, **375**: 754-760.

Siddique T, Figlewicz DA, Pericak-Vance MA, Haines JL, Rouleau G, Jeffers AJ, Sapp P, Hung WY, Bebout J, McKenna-Yasek D, (1991). Linkage of a gene causing familial amyotrophic lateral sclerosis to chromosome 21 and evidence of genetic-locus heterogeneity. *N Engl J Med.*, **324**: 1381-1384.

Siegfried E, Ambrosio L, Perrimon N, (1990). Serine/threonine protein kinases in Drosophila. Trends Genet., 6: 357-362.

Skovby F, Krassikoff N, Francke U, (1984). Assignment of the gene for cystathionine beta-synthase to human chromosome 21 in somatic cell hybrids. *Hum. Genet.*, 65: 291-294.

Smith DJ, Zhu Y, Zhang J, Cheng JF, Rubin EM, (1995). Construction of a panel of transgenic mice containing a contiguous 2-Mb set of YAC/P1 clones from human chromosome 21q22.2. Genomics, 27: 425-434.

Smith DJ, Stevens ME, Sudanagunta SP, Bronson RT, Makhinson M, Watabe AM, O'Dell TJ, Fung J, Weier HU, Cheng JF and Rubin EM, (1997). Functional screening of 2 Mb of human chromosome 21q22.2 in transgenic mice implicates *minibrain* in learning defects associated with Down syndrome. *Nat Genet.*, 16: 28-36.

Song WJ, Chung SH and Kurnit DM, (1997). The murine Dyrk protein maps to chromosome 16, localizes to the nucleus, and can form multimers. *Biochem Biophys Res Commun.*, 231: 640-644.

Sago H, Carlson EJ, Smith DJ, Kilbridge J, Rubin EM, Mobley WC, Epstein CJ and Huang TT, (1998). Ts1Cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. *Proc Natl Acad Sci U S A*, 95: 6256-6261.

Stallings RL, Ford AF, Nelson D, Torney DC, Hildebrand CE and Moyzis RK, (1991). Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. *Genomics*, 10: 807-815.

Stene J, (1970). Statistical inference on segregation ratios for D/G-translocations, when the families are ascertained in different ways. Ann Hum Genet., 34: 93.

Stewart EA, McKusick KB, Aggarwal A, Bajorek E, Brady S, Chu A, Fang N, Hadley D, Harris M, Hyssain S, Lee R, Maratukulam A, O'Connor K, Ferkins S, Piercy M, Qin F, Reif T, Sanders C, She X, Sun W-L, Tabar, P, Woyticky S, Cowles S, Fan J-B, Mader C, Quackenbush, J, Myers RM and Cox DR, (1997). An STS-based radiation hybrid map of the human genome. *Genome Res.*, 7: 422-433.

Straub RE, Lehner T, Luo Y, Loth JE, Shao W, Sharpe L, Alexander JR, Das K, Simon R, Fieve RR, Lerer B, Endicott J, Ott J, Gilliam T.C. and Baron M, (1994). A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. *Nature Genet.* 8: 291-296.

Suetsugu M and Mehraein F, (1980). Spine distribution along the apical dendrites of the pyramidal neurons in Down's syndrome. A quantitative Golgi study. Acta Neuropathol, 50: 207-210.

Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A and Kola I, (1996). Down's syndrome-like skeletal abnormalities in *Ets2* transgenic mice. *Nature*, **379**: 534-537.

\mathcal{T}

Tagle DA and Collins FS, (1992). An optimized Alu-PCR primer pair for human-specific amplification of YACs and somatic cell hybrids. *Hum Mol Genet.*, 1: 121-122.

Tagle DA, Swaroop M, Lovett M and Collins FS, (1993). Magnetic bead capture of expressed sequences encoded within large genomic segments. *Nature*, **361**: 751-753.

Tanzi RE, Haines JL, Watkins FC, Stewart GD, Wallace MR, Hallewell R, Wong C, Wexler NS, Conneally PM and Gusella JF, (1988). Genetic linkage map of human chromosome 21. *Genomics*, 3: 129-136.

Tanzi RE, Watkins FC, Stewart GD, Wexler NS, Gusella JF and Haines JL, (1992). A genetic linkage map of human chromosome 21: analysis of recombination as a function of sex and age. Am J Hum Genet., 50: 551-558.

Tassone F, Xu H, Burkin H, Weissman S and Gardiner K, (1995). cDNA selection from 10 Mb of chromosome 21 DNA: efficiency in transcriptional mapping and reflections of genome organization. *Hum. Mol. Genet.*, **9**: 1509-1518.

Taylor SS, Knighton DR, Zheng J, Sowadski JM, Gibbs CS and Zoller MJ, (1993). A template for the protein kinase family. *Trends Biochem Sci.*, 18: 84-89.

Tejedor F, Zhu XR, Kaltenbach E, Ackermann A, Baumann A, Canal I, Heisenberg M, Fischbach KF, Pongs O, (1995). minibrain: a new protein kinase family involved in postembryonic neurogenesis in Drosophila. Neuron, 14: 287-301.

Teller JK, Russo C, DeBusk LM, >, (1996). Presence of soluble amyloid beta-peptide precedes amyloid plaque formation in Down's syndrome. Nat Med., 2: 93-95.

Thase ME, (1982). Longevity and mortality in Down's syndrome. J. Ment Defic Res., 26: 177.

Tjio JH and Levan A, (1956). The chromosomes of man. Acta Genet., 6: 264-266.

Uberbacher EC and Mural RJ, (1991). Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc Natl Acad Sci USA*, 88: 11261-11265.

V

Valdes JM, Tagle DA and Collins FS, (1994). Island rescue PCR: a rapid and efficient method for isolating transcribed sequences from yeast artificial chromosomes and cosmids. Proc Natl Acad Sci USA, 91: 5377-5381.

Van Ness J and Hahn WE, (1980). Sequence complexity of cDNA transcribed from a diverse mRNA population. Nucleic Acids Res., 8: 4259-4270.

Veske A, Oehlmann R, Younus F, Mohyuddin A, Muller-Myhsok B, Mehdi SQ and Gal A, (1996). Autosomal recessive non-syndromic deafness locus (*DFNB8*) maps on chromosome 21q22 in a large consanguineous kindred from Pakistan. *Hum Mol Genet.*, 5: 165-168.

Virtaneva K, D'Amato E, Miao J, Koskiniemi M, Norio R, Avanzini G, Franceschetti S, Michelucci R, Tassinari CA, Omer S, Pennacchio LA, Myers RM, Dieguez-Lucena JL, Krahe, R, de la Chapelle A and Lehesjoki AE, (1997). Unstabel minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1. Nat. Genet., 15: 393-396.

Vogelstein B and Gillespie D, (1979). Preparative and analytical purification of DNA from agarose. Proc Nat Acad Sci USA, 76: 615-619.

W

Wang and Bellugi, (1994). Evidence from two genetic syndromes for a dissociation between verbal and visual-spatial short-term memory. J Clin Exp Neuropsychol, 2: 317-322.

Wang D and Smith CL, (1994). Large-scale structure conservation along the entire long arm of human chromosome 21. Genòmics, 3: 441-451.

Warburton D, (1985). Genetic factors influencing aneuploidy frequency, in Dellarco VL, Voytek PE, Hollaender A (eds): Aneuploidy, Etiology and Mechanisms. New York, Plenum, p133.

Warren AC, Chakravarti A, Wong C, Slaugenhaupt SA, Halloran SL, Watkins PC, Metaxotou C and Antonarakis SE, (1987). Evidence for reduced recombination on the nondisjoined chromosomes 21 in Down syndrome. *Science*, 237: 652-654.

Warren AC, Slaugenhaupt SA, Lewis JG, Chakravarti A and Antonarakis SE, (1989). A genetic linkage map of 17 markers on humanchromosome 21. Genomics, 4: 579-591.

Weber JL and May PE, (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet., 44: 388-396.

Weber JL, Botsein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment lenght polymosphisms. Am J. Hum. Genet. 32: 314-331.

Weis S, (1991). Morphometry and magnetic resonance imaging of the human brain in normal controls and Down's syndrome. Anat Rec., 231: 593-598.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau F, Vaysseix G and Lathrop M, (1992). A second-generation linkage map of the human genome. *Nature*, 359: 794-801.

White K and Kankel DR, (1978). Fatterns of cell division and cell movement in the formation of the imaginal nervous system in *Drosophila melanogaster*. Dev Biol Aug. 65: 296-321.

Willard HF, (1990). Centromeres of mammalian chromosomes. Trends Genet, 6: 67-112.

Williams CA, Frias JL, McCormick MK, Antonarakis SE and Cantu ES, (1990). Clinical, cytogenetic, and molecular evaluation of a patient with partial trisomy 21 (21q11-q22) lacking the classical Down syndrome phenotype. Am J Med Genet., Suppl 7: 110-114.

Williams JD, Summitt RL, Martens FR and Kimbrell RA, (1975). Familial Down syndrome due to t(10;21) translocation: evidence that the Down phenotype is related to trisomy of a specific segment of chromosome 21. Am J Hum Genet., 27: 478-485.

Williams Rs and Matthyse S, (1986). Age-related changes in Down syndrome brain and the cellular pathology of Alzheimer disease. *Prog Brain Res.*, **70**: 49.

Wilson T and Treisman R, (1988). Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3'AU-rich sequences. *Nature*, 336: 396-399.

Wisniewski KE, (1990). Down syndrome children often have brain with maturation delay, retardation of growth, and cortical dysgenesis. Am J Med Genet., Suppl 7: 274-281.

Х

Xu H, Wei, H, Tassone F, Graw S, Gardiner K and Weissman S, (1995). A search for genes from the dark band regions of human chromosome 21. Genomics, 27: 1-8.

Y

Yaspo M-L, Gellen L, Mott R, Dorn B, Nizetic D, Poustka AM and Lehrach H, (1995). Model for a transcript map of human chromosome 21: isolation of new coding sequences from exon and enriched cDNA libraries. *Hum. Mol. Genet.*, 8: 1291-1304.

Yoon PW, Freeman SB, Sherman SL, Taft LF, Gu Y, Pettay D, Flanders WD, Khoury MJ, Hassold TJ, (1996). Advanced maternal age and the risk of Down syndrome characterized by the meiotic stage of chromosomal error: a population-based study. Am J Hum Genet., 58: 628-633.

Zucman J, Melot T, Desmaze C, Ghysdael J, Plougastel B, Peter M, Zucker JM, Triche TJ, Sheer D, Turc-Carel C, (1993). Combinatorial generation of variable fusion proteins in the Ewing family of tumours. *EMBO J.*, **12**: 4481-4487.