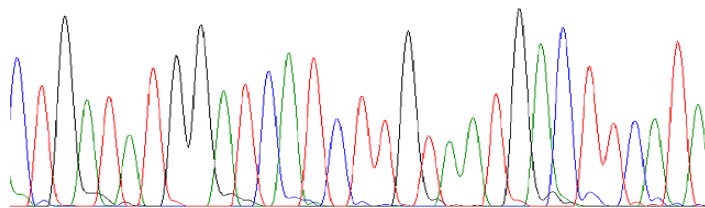


**IDENTIFICACIÓ DELS FACTORS GENÈTICS QUE
DETERMINEN LA VARIABILITAT DELS NIVELLS
DE FVII A LA POBLACIÓ ESPANYOLA**

Resultats del Projecte GAIT



MARIA SABATER LLEAL
2006

Tesi Doctoral
Departament de Genètica
Universitat de Barcelona

MATERIAL I MÈTODES

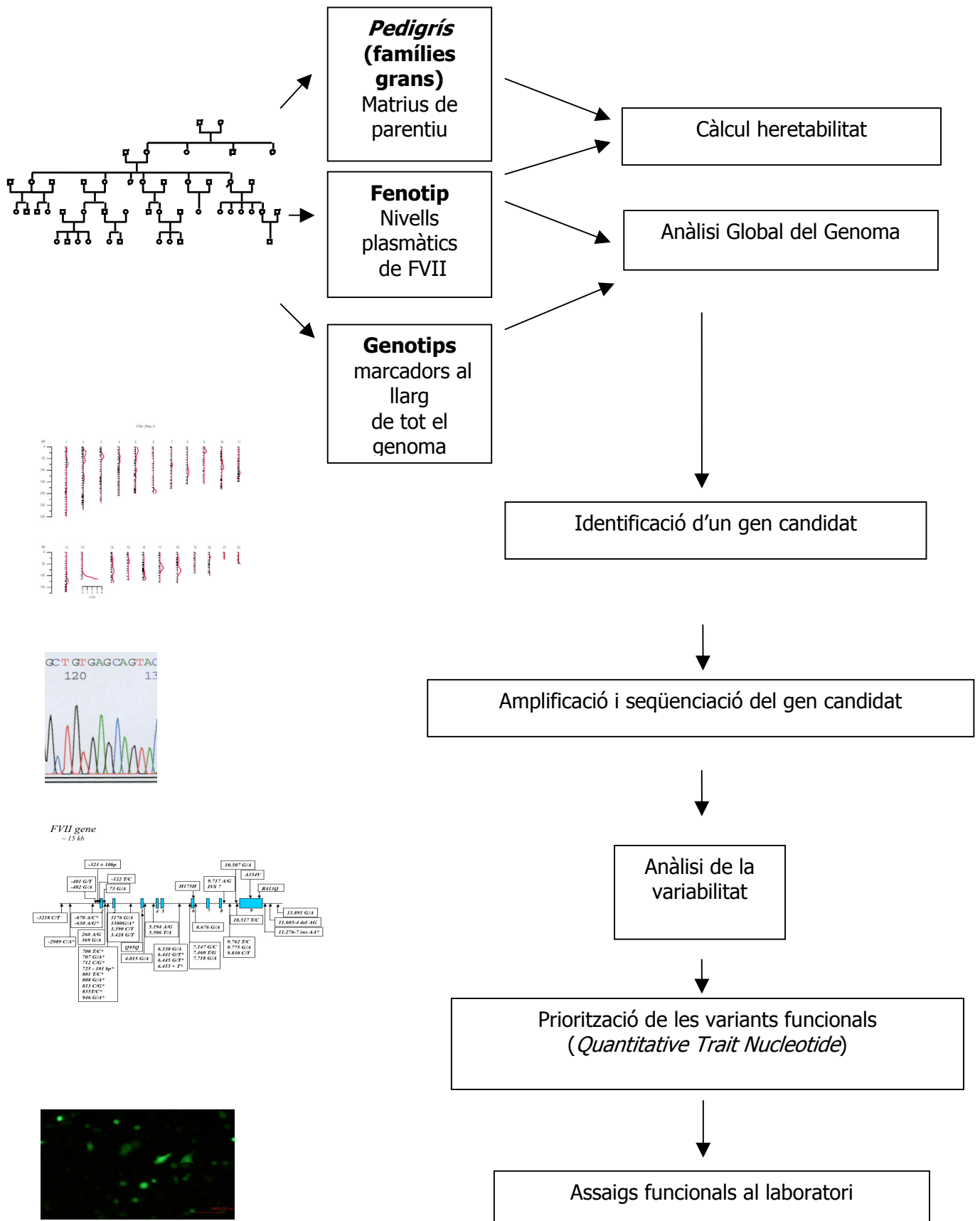


Figura 7 : Estratègia per la localització de polimorfismes que afecten un fenotip quantitatiu (QTNs).

1- LA MOSTRA GAIT

La mostra GAIT es compon de 21 famílies extenses (mínim de deu individus repartits en tres o més generacions). D'entre aquestes famílies, 12 van ser seleccionades a partir d'un *propositus* amb trombofília idiopàtica. Els criteris d'inclusió incloïen; o bé pacients que presentessin clínica trombòtica venosa abans dels 45 anys, o bé pacients amb trombosis recurrents, o bé pacients amb episodis trombòtics espontanis però amb història familiar de trombosi. La trombofília es va considerar idiopàtica per exclusió prèvia de les causes biològiques conegudes de trombofília durant el període de reclutament (de 1995 a 1997), és a dir, deficiències d'AT, PC, PS, i plasminogen, la presència de resistència a la PC activada (RPCa), la presència d'anticoagulant lúpic, anticossos antifosfolípids o disfibrinogenèmia.

Les 9 famílies control es van seleccionar amb els mateixos criteris de mida d'entre la població general i de forma aleatòria.

El reclutament dels membres de les famílies es va realitzar fonamentalment a Barcelona, encara que es va completar amb individus emparentats i residents a Lleida, Córdoba, Málaga, Cádiz i Albacete. Es van reclutar un total de 398 individus, dels quals 186 homes i 212 dones. Del total d'individus, 101 eren individus fundadors.

A tots els individus se'ls va recollir la història clínica i un consentiment informat per mitjà d'una entrevista personal amb els metges de la Unitat d'Hemostàsia i Trombosi de l'Hospital de Sant Pau. També es va recollir informació bàsica (edat, sexe, grau de relació amb el *propositus*), factors de risc cardiovascular (consum de tabac, hipertensió, dislipèmia, diabetis, varius, neoplàsia, obesitat), història reproductiva i utilització d'anticonceptius orals en dones per determinar els factors ambientals que podien influir en el risc trombòtic, i descripció minuciosa d'antecedents trombòtics venosos o arterials (localització, edat i circumstàncies associades). Finalment es va registrar la composició dels domicilis per valorar l'efecte d'influències ambientals compartides, com la dieta.

Nº de família	Nº d' individus	Homes	Dones	Tipus de família
1	21	15	7	Control
2	27	14	13	Control
3	22	10	12	Control
4	15	6	9	Control
5	14	3	11	Control
6	24	10	14	Control
7	17	11	6	Control
8	22	7	15	Control
9	20	9	11	Control
10	21	15	6	Trombofilia
11	38	19	19	Trombofilia
12	13	5	8	Trombofilia
13	11	6	5	Trombofilia
14	10	6	4	Trombofilia
15	33	10	23	Trombofilia
16	20	10	10	Trombofilia
17	14	5	9	Trombofilia
18	13	8	5	Trombofilia
19	14	7	7	Trombofilia
20	18	7	11	Trombofilia
21	9	3	6	Trombofilia
Total	398	186	211	9 Controls 12 Trombofilia

Taula 6. Distribució, per tipus de família, dels individus examinats del Projecte GAIT

2- FENOTIPS: DETERMINACIONS PLASMÀTIQUES

La sang es va obtenir per punció venosa en un tub de vidre siliconat (*Vacutainer, Becton Dickinson* ref 367702) amb citrat trisòdic a 0,129 M, en una proporció de 1:9 volum de citrat/sang. L'interval entre l'extracció i el processament al laboratori no va superar mai les dues hores.

Per la determinació de les proteïnes de la coagulació, es va centrifugar la sang a 3500 rpm (1600g) durant 20 minuts a temperatura ambient per tal d'obtenir plasma pobre en plaquetes. A continuació, aquest plasma es va separar del precipitat cel·lular i es va analitzar el mateix dia de l'obtenció de la sang, o bé es va congelar a $-40\text{ }^{\circ}\text{C}$ per ser analitzat abans d'un mes, o a $-80\text{ }^{\circ}\text{C}$ per ser emmagatzemat un temps més prolongat.

A partir d'aquí, es van realitzar 64 determinacions plasmàtiques relacionades amb fenotips de l'hemostàsia (taula 7), del metabolisme del ferro o dels lípids.

Risc de trombosi	% $\alpha_2\beta_0$ *
TTPA*	α_2 -antiplasmina*
Temps Protrombina*	Plasminogen*
Temps de protrombina en presència de trombomodulina (TM) (1)*	PAI-1*
Temps de protrombina modificat (sense TM) (2)*	t-PA*
Ratio entre (1) i (2)*	uPA*
Trombina*	TAFI Ag*
Fibrinogen*	TAFI func*
Precalicroïna *	Dímer D*
FV*	Homocisteïna*
FVII*	F1+F2*
FVIII*	Anticòs anticardiolipina IgG*
FIX*	Anticòs anticardiolipina IgM*
FX*	Anticòs antifosfatidilserina IgG*
FXI*	Anticòs antifosfatidilserina IgM*
FXII*	P-selectina
FXIII*	Folat en sang total
Factor Tissular*	Folat sèric
Factor vWF*	Vitamina B ₁₂
β_2 -GPI*	Sideremia
HRG*	Capacitat trans. Insat
TFPI*	Capacitat trans. Total
PC*	Saturació
Antitrombina*	Ferritina
TM*	Receptor Transferrina
RPCa*	Temps de Russell*
Cofactor II heparina*	Colesterol total
PS lliure*	HDL
PS func*	LDL
PS total*	VLDL
C4BP*	Lipoproteïna A
C4BP- β *	Triglicèrids

Taula 7: Fenotips determinats en el Projecte GAIT. * Indica els fenotips de l'hemostàsia

TTPA: Temps de Tromboplastina Parcial Activat

RPCa: Resistència a la proteïna C activada

HRG: Glicoproteïna rica en histidina

β_2 -GPI: beta 2 Glicoproteïna

Determinació plasmàtica del FVII

El FVII funcional es va mesurar el mateix dia de l'obtenció de la sang mitjançant un sistema deficient en FVII. En aquesta metodologia, es mesura el temps de coagulació en presència de tromboplastina i ions calci en una mostra on hi ha tots els factors necessaris per la coagulació en excés a excepció del factor VII, que l'aporta el plasma diluït del malalt.

Thromborel S: Tromboplastina càlcica humana, extreta de placenta, clorur càlcic i estabilitzadors.

Behring, ref. OUHP

Stago Deficient VII: Plasma deficient en Factor VII- Stago ref. 5000743

Unicalibrator: Plasma calibrador pel test de coagulació. Stago ref. 00625.

Taula 8: Reactius utilitzats en la determinació plasmàtica del FVII

El FVII antigènic es va quantificar per les famílies amb deficiència de FVII mitjançant un mètode ELISA (*Enzyme-linked immunosorbent assay*) amb el Kit comercial *Asserachrom® VII:Ag* (*Diagnostica Stago*, Asnieres, France).

3- OBTENCIÓ DEL DNA

Per l'obtenció de DNA utilitzat en les anàlisis genètiques, es van seguir els protocols de rentat de sang, lisi d'eritròcits, i extracció del DNA de leucòcits amb el mètode *Salting out procedure* (Miller, et al., 1988). Un cop obtinguda la mostra de DNA, es va conservar a 4 °C o es va congelar indefinidament a -20 °C.

Rentat de la sang i lisi dels eritròcits

El rentat de sang total es va realitzar afegint sèrum fisiològic salí (Braun) al total del volum de la sang (de 5 a 20 ml de sang perifèrica) en un tub cònic fins a un volum final de 50 ml i centrifugant a 2800 rpm durant 10 minuts a 4 °C i sense fre per separar el precipitat cel·lular (format per eritròcits, leucòcits i plaquetes) del plasma de la mostra.

Un cop obtingut el precipitat cel·lular, es va procedir al procés de lisi d'eritròcits mesclant el precipitat amb un tampó de lisi d'eritròcits (taula 9) fins a un volum final de 50 ml, per obtenir només els leucòcits. La mescla s'ha de mantenir en gel durant 20-30 minuts i centrifugar a 3000 rpm durant 15 minuts a 4 °C i sense fre. Es va descartar el sobrenadant i es va repetir el procés fins obtenir un precipitat de leucòcits net (generalment, dos o tres rentats). El precipitat es va congelar a -20 °C o es va prosseguir amb l'extracció de DNA pel mètode *Salting out procedure* (Miller, et al., 1988) que consisteix en tres passos successius: lisis de leucòcits, precipitació de proteïnes, i finalment, precipitació del DNA.

- La **lisi de leucòcits** es va realitzar mitjançant la mescla de 6 ml de tampó de lisi de leucòcits, 0.4 ml SDS al 10% (dodecilsulfat sòdic, SERVA) com a detergent, i 1 ml (o més, en funció de la mida del precipitat cel·lular) de solució de proteïnasa K a 2 mg/ml (taula 9) al precipitat cel·lular. La mescla es va mantenir en agitació suau durant 16 hores a 37 °C.
- Per la **precipitació de proteïnes** es va afegir 1 ml (o la quantitat proporcional necessària) de NaCl (Merck) sobresaturat a 5.5 M i es va agitar durant 15-30 segons fins obtenir una emulsió completa. Es va centrifugar a 3400 rpm 15-20 minuts a 4 °C i es va recuperar el sobrenadant en un altre tub cònic. El procés es va repetir les vegades necessàries fins eliminar completament totes les restes de sals de NaCl (normalment tres vegades).
- Per la **precipitació del DNA** es van afegir 20 ml (en el cas de 6 ml de lisi de leucòcits) d'etanol absolut (Merck). Seguidament es va anar agitant suaument fins que es va observar la formació del precipitat del DNA en forma de medusa. Es va recollir el DNA amb una pipeta Pasteur de vidre estèril i es van eliminar les restes de sals per immersió en etanol al 70%. Finalment, el precipitat de DNA es va resuspendre en TRIS/EDTA 10 mM/0.2 mM (0.5-3 ml, depenent de la mida de la medusa de DNA) i es va mantenir en agitació suau durant una hora a 37 °C i després a 4 °C .

Sèrum fisiològic BRAUN. Solució isotònica de clorur sòdic al 0,9% (500 ml).

Tampó de lisi d'eritròcits: s'ajusta amb HCl a pH 7,5

Tris HCl 20 mM (Tris: hidroximetil-aminometan; HCl: àcid clorhídric, Merck).

MgCl₂ 5 mM (Clorur de Magnesi, Merck)

Tampó de lisi de leucòcits: s'ajusta pH a 8,2 amb HCl

Tris (Merk) 0,01M

NaCl (Clorur sòdic) (Merck) 0.4 M

EDTA (Etilendiaminotetracètic Merck) 2 mM

Proteïnasa K:

Proteïnasa K (Roche) 2mg/ml.

SDS 1%.

EDTA 2mM

Tris-EDTA 10 mM/0,2 mM:

Tris (Merk) pH 7,5; 10 mM

EDTA (Merck) 2 mM

Taula 9: Reactius i solucions de treball per l'extracció de DNA de leucòcits

Per analitzar la puresa i la concentració del DNA es van fer dues lectures en l'espectrofotòmetre: una a densitat òptica (DO) de 260 nm (DO₂₆₀), que correspon al DNA bicatenari present en la mostra, i una altra a DO de 280 nm (DO₂₈₀) que correspon a la concentració de proteïnes contaminants. La puresa de la mostra ve determinada per la relació

$$\frac{DO_{260}}{DO_{280}}$$

que ha de quedar compresa entre 1.6 i 2. Una relació inferior a 1.6 indica contaminació de proteïnes, mentre que una relació superior a dos indica contaminació per sals.

La concentració del DNA es calcula mitjançant la lectura a DO₂₆₀, tenint en compte la dilució realitzada en la mostra i el coeficient d'extinció molar del DNA bicatenari (50).

$$\mu\text{g/ml de ADN} = DO_{260} \times 50 \times \text{dilució}$$

4- MÈTODES ESTADÍSTICS PER A LA LOCALITZACIÓ DE GENS

4.1- Càlcul de l'heretabilitat

Els elements necessaris per calcular l'heretabilitat d'un fenotip són una mostra d'individus relacionats en famílies, els seus graus de parentiu, i els valors quantitius del fenotip que es vulgui analitzar. Quan major sigui el grau de parentiu entre dos individus, major quantitat d'al·lels compartiran en promig en el global del genoma. Si els valors del fenotip estudiat tendeixen a ser més semblants en els individus emparentats que en els no emparentats, i més semblants quan més proper és el parentiu, s'estableix que els gens tenen una influència en la variabilitat del fenotip analitzat.

Tipus de parentiu	Grau de parentiu	Fracció del genoma que comparteixen en promig	Nombre de parelles en la mostra GAIT
Individus (398)	0	1	-
Bessons monozigots	0	1	1
Pare-fill	1	1/2	470
Germans	1	1/2	340
Avi-nét	2	1/4	225
Oncle-nebot	2	1/4	693
Mig germans	2	1/4	13
Besavi-besnét	3	1/8	13
Oncle avi-nét	3	1/8	137
Cosins germans	3	1/8	547
Oncle besavi-besnét	4	1/16	9
Oncle segon-nebot	4	1/16	233
Cosins segons	5	1/32	63
TOTAL			2744

Taula 10: tipus de parentiu, grau de relació, fracció del genoma compartit en promig, i nombre de parelles en la mostra GAIT. En total tenim 2744 parelles diferents d'individus amb diferents graus de parentiu, nombre que representa la mida real de la mostra estadística.

Per calcular l'heretabilitat utilitzem un model matemàtic poligènic. Un model poligènic intenta explicar la variabilitat fenotípica d'un caràcter complex regulat per múltiples gens i estima l'efecte que fan el global dels gens sobre el fenotip determinat. La funció matemàtica que el determina és la següent:

$$Y = \mu + \Sigma cov + \sigma^2$$

És a dir, que el valor del fenotip en cada individu és igual a la mitjana del valor del fenotip en la població més l'efecte d'unes covariables conegudes (sexe, edat, ús d'anticonceptius orals, consum de tabac, ...) més una variabilitat deguda a l'efecte de variables no conegudes.

Un cop determinat el percentatge de la variància degut a les covariables controlades, aquesta es treu de l'equació, i obtenim

$$Y = \mu + \sigma^{2*}$$

$$(\sigma^{2*} < \sigma^2)$$

Aquesta nova variància és deguda a tots els factors desconeguts que afecten el nostre fenotip, i el model matemàtic utilitzat per estimar-la es basa en l'anàlisi dels components de la variància (Almasy and Blangero, 1998). Aquest model postula que la variància total observada en un fenotip (σ^{2*}) és el resultat de la suma de les variàncies degudes a factors genètics (σ_g^2) i ambientals (σ_e^2) i descriu la dispersió dels valors del fenotip en els diferents individus al voltant de la mitjana de la població.

$$\sigma^{2*} = \sigma_g^2 + \sigma_e^2$$

$$h^2 = \sigma_g^2 / \sigma_p^2$$

El càlcul de l'**heretabilitat** (h^2) no és més que el quocient entre la variància causada pels factors genètics i la variància causada pels factors ambientals. El programa utilitzat per calcular-la, integrat dins el paquet informàtic SOLAR (Almasy and Blangero, 1998), analitza la correlació que hi ha entre el grau de parentiu (percentatge d'al·lèles que comparteix cada parella de la mostra) i els valors del fenotip. Com més alta sigui aquesta correlació, més alta serà l'heretabilitat del fenotip.

4.2- Anàlisi de gens candidats i Anàlisi Global del Genoma (*Genome Wide Scan*)

Per a la localització dels gens de susceptibilitat en la malaltia tromboembòlica, en aquest treball es van desenvolupar dues estratègies complementàries; en un primer moment es va fer un estudi de gens candidats en el qual es van analitzar 70 marcadors de tipus microsatèl·lits o SNP localitzats dins o molt a prop de gens candidats relacionats amb l'hemostàsia (taula 11). Els marcadors es van analitzar al Centre de Genotipació de París (França).

En segon lloc, degut a la millora de la tecnologia de genotipació, i per tal d'ampliar la cerca també a gens desconeguts, es va contemplar la realització de la primera Anàlisi Global del Genoma dirigida a la cerca de QTLs relacionats amb l'hemostàsia.

L'Anàlisi Global del Genoma realitzada en aquest treball ha consistit en una anàlisi de lligament genètic entre el QTL causant del fenotip d'interès i tota una bateria de marcadors altament polimòrfics situats al llarg de tot el genoma, amb l'objectiu de poder localitzar el *loci* responsable de la variació en el fenotip en una regió concreta del genoma.

Es van usar inicialment 363 marcadors de tipus microsatèl·lit separats 9,5 cM (unitats de recombinació genètica) entre ells. Els microsatèl·lits utilitzats van ser genotipats al Centre de Genotipació de París (França). Es va realitzar l'anàlisi de lligament implementat en el programa informàtic SOLAR (Almasy and Blangero, 1998) com es detalla a continuació.

Marcador	Gen Candidat
D20S101	Trombomodulina
D11S986	Protrombina
C677T	Metiltetrahidrofolat Reductasa
A455V	Trombomodulina
F8C	Factor VIII
D2S2271	Proteïna C
DXS52	Factor VIII
D1S194	Factor V
D1S318	Factor V
Hpa 5	Glicoproteïna plaquetària de membrana
PT20210	Protrombina
FV-Leiden	Factor V
ABO	Grup Sanguini
Hpa 1-2-3	Glicoproteïna plaquetària de membrana
D3S1271	Proteïna S
D1S2635	Locus Factor V cromosoma 1
A34V	Factor XIII
PS Heerlem	Proteïna S
D2S1895	IL-1
D2S1896	IL-1
D1S215	Locus Factor V cromosoma 1
His1299Arg (4070A/G)	Factor V
IVS16	Factor V
Arg485Lys (1628 G/A)	Factor V
IVS 11	Factor V
R353Q	Factor VII
Ser739Ser (2391A/G)	Factor V
-323ins0/10	Factor VII
His115His	Factor VII
Ser333Ser	Factor VII
-122 T/C	Factor VII
-401 G/T	Factor VII
-402 G/A	Factor VII
Met1736Val (5380A/G)	Factor V
2298 C/T	Factor V
2325 T/C	Factor V
2833 A/T	Factor V
A455V	TM
ABO	ABO
D11S986	PT
D1S318	FV
D20S101	TM
F13A	FXIII
FV-Leiden	FV
C677T	MTHFR
DXS52	FVIII
F8c	FVIII
F13B	FXIII
D1S194	FV
D2S2271	PC
G20210A	PT
D22S280	HCII
D13S285	X/VII
Hpa1	GPIIIa
Hpa2	GPIbalfa
Hpa3	GPIIb
Hpa5	GpiA
A34V	FXIII
D3S1271	PS
D1S215	FV
D1S2635	FV
D2S2215	PC
D17S807	beta2GP
D4S171	FXI
D5S625	FXII

Taula 11: Relació dels marcadors utilitzats i els gens candidats analitzats en el Projecte GAIT.

Matrius Identical By Descent (IBD)

Amb la informació dels marcadors i de les famílies es van calcular les matrius IBD. Per a cada marcador, es va calcular el nombre d'al·lels que compartia cada parella d'individus de la mostra, és a dir, com de genèticament semblants eren les parelles d'individus per a cada marcador (0, 0.5 ó 1).

A partir d'aquests resultats, es van estimar les **matrius multipoint**, en les quals es calcula la semblança genètica entre parelles d'individus per a cada centimorgan (cM) en concret del genoma, a partir de les dades dels marcadors genotipats. Per a estimar les matrius *multipoint* es va utilitzar el paquet informàtic SOLAR (Almasy and Blangero, 1998) amb el que es va obtenir una matriu per a cada cM del genoma.

L'anàlisi de lligament es basa en el següent model

$$Y = \mu + \Sigma_{cov} + h^2q_1 + h^2r + e$$

és a dir, per a cada una de les matrius *multipoint* (per a cada cM) estima l'efecte *d'aquell locus* concret versus l'efecte de la resta dels gens, més un error que inclou l'efecte dels factors ambientals, l'efecte *household* (efecte de la dieta i altres factors compartits pels membres d'un mateix domicili), i l'error de mesura. En aquest sentit, el càlcul no considera la semblança genètica del *global* del genoma sinó la semblança genètica existent *per cada cM en concret*, i ho compara amb la hipòtesi nul·la en què el cM no tingui cap efecte sobre el fenotip.

El paràmetre utilitzat en estadística per mesurar el grau de lligament és l'escala de LOD, o logaritme de l'odds ratio entre la probabilitat que ambdós *loci* estiguin lligats respecte de la probabilitat que no ho estiguin. Es considera que un LOD és significatiu quan és superior a 3, és a dir, quan la probabilitat que ambdós *loci* estiguin lligats és 1000 vegades superior a la probabilitat que no ho estiguin. Aquesta assumpció ens dona una p (probabilitat d'error de tipus I) de 0.0001 (Lander and Kruglyak, 1995).

$$\text{LOD} = \log_{10} \frac{\text{Probabilitat que ambdós loci estiguin lligats}}{\text{Probabilitat de no lligament}}$$

Quan es calcula el LOD per a cada cM concret, obtenim una probabilitat de lligament en cada punt del genoma, i per tant, ens permet localitzar el QTL. La probabilitat de localitzar un QTL en una anàlisi determinada és funció de la mida de la mostra, la complexitat del pedigrí i la importància relativa del QTL en la població estudiada (Blangero, et al., 2003).

4.3- Localització d'un gen candidat

Arribats en aquest punt, el camí a seguir dependrà de si es té un gen candidat, és a dir, un gen conegut susceptible de ser el responsable del senyal de lligament observat (en aquest cas perquè codifiqui per una proteïna implicada en l'hemostàsia). En aquesta fase d'identificació de gens candidats dintre d'una zona de lligament genètic, el Projecte Genoma Humà pot servir de gran ajuda, ja que aporta un catàleg complet dels gens ubicats en la regió cromosòmica d'interès, i un resum de les seves funcions.

Si tenim un gen candidat es pot passar al pas següent, que és la identificació de les variants al·lèliques d'aquest gen (polimorfismes) que influeixen en els fenotips analitzats i que, per tant, puguin contribuir al risc trombòtic (*QTNs*, *Quantitative Trait Nucleotide*).

En aquest treball, el *locus* lligat a la variació de nivells plasmàtics de FVII funcional comprenia precisament el gen estructural del FVII, és a dir, el gen (*F7*) que codifica per la seva seqüència d'aminoàcids. Per tant, *a priori* teniem el gen candidat obvi.

5- AMPLIFICACIÓ I SEQÜENCIACIÓ DEL GEN F7

5.1- La reacció en cadena de la polimerasa (PCR)

La tècnica de la PCR és un mètode extremadament útil en el camp de la genètica molecular, que es basa en l'amplificació selectiva de regions específiques de DNA *in vitro* (Saiki, et al., 1988). La tècnica, posada a punt el 1985 per K. Mullis, consisteix en utilitzar l'activitat de l'enzim Taq polimerasa que, partint d'una cadena motlle i d'un petit oligonucleòtid inicial (*primer*), és capaç d'estendre una nova cadena complementària en presència de nucleòtids en excés i magnesi. Una reacció de PCR consta de varis cicles (generalment 30-35) que van oscil·lant contínuament entre tres temperatures

- una temperatura de desnaturalització de les cadenes de DNA (94-95 °C)
- una temperatura d'hibridació dels *primers* específics (50-65 °C)
- una temperatura d'extensió per l'enzim Taq Polimerasa (72-74 °C)

Les noves molècules sintetitzades serveixen també de nous motlles, de manera que en cada cicle el nombre de molècules de DNA creix de forma exponencial i a més es limiten entre els dos *primers*, determinant així la regió específica de DNA que volem amplificar.

La temperatura d'hibridació dels *primers* (T) depèn de la seva composició de bases, i es calcula en °C mitjançant la fórmula següent:

$$T = [4 \times (G+C) + 2 \times (A+T)] - 4$$

La temperatura d'hibridació d'una PCR ens serveix per controlar les seves condicions d'astringència, és a dir, l'especificitat amb la qual s'uniran els *primers* a les cadenes motlle. A temperatures més baixes, els *primers* seran capaços d'unir-se a cadenes cada vegada menys homòlogues, mentre que a temperatures molt altes, les unions seran altament específiques.

El disseny dels *primers* és determinant per l'especificitat de la reacció de PCR. Generalment, les característiques que s'han de controlar a l'hora de dissenyar els *primers* d'una PCR són els següents:

- llargada d'entre 17-28 pb
- composició d'entre 40-60% (G+C)
- és convenient que acabin amb G o C a l'extrem 3'
- temperatura d'hibridació entre 50-65 °C
- Evitar regions complementàries dins el *primer* o amb l'altre *primer*

Amplificació per PCR del gen *F7*

Per localitzar els polimorfismes funcionals que determinen la variabilitat dels fenotips analitzats és imprescindible seqüenciar tot el gen candidat en un nombre suficientment ampli d'individus com per assegurar que tots els polimorfismes presents en la població amb una prevalència superior al 5% hi siguin representats i es puguin detectar. Aquesta seqüenciació ha d'incloure la zona promotora, exons, introns i regió 3' no traduïda, per assegurar-nos que no obviem regions reguladores que modulin l'expressió del gen.

En aquest treball es van reseqüenciar completament un total de 40 individus de la mostra GAIT. El criteri de selecció per tal d'augmentar la probabilitat de detectar variants funcionals va consistir en triar 22 dels individus amb nivells alts de FVII (almenys una desviació estàndard per sobre la mitjana) i 18 individus amb nivells baixos de FVII (almenys una desviació estàndard per sota la mitjana), el qual donava una probabilitat major del 99% de detectar qualsevol polimorfisme amb una freqüència igual o superior 0.05.

A més, també es van seqüenciar els membres de dues famílies amb deficiència pràcticament total de FVII. Finalment també es van seqüenciar els DNAs corresponents a dos individus no relacionats de cadascuna de les espècies de primats següents: goril·la (*Gorilla gorilla*), orangutà (*Pongo pygmaeus*), ximpanzè (*Pan troglodytes*) i bonobo (*Pan paniscus*) per les anàlisis especificades en l'apartat 6.2. En tots els casos es va utilitzar la mateixa estratègia d'amplificació i seqüenciació que es detalla a continuació.

Degut a les limitacions de la PCR, la nostra estratègia va consistir en amplificar fragments de 1 a 2 Kb. El gen *F7* va ser dividit en 10 fragments amplificables per PCR amb els *primers* descrits en la taula 12. La Figura 8 mostra un esquema del gen amb els fragments que han estat utilitzats per la seva amplificació. Els fragments se solapen totalment a excepció d'una regió de l'intró 2 que no va poder ser amplificada degut al seu gran contingut en GCs.

Les amplificacions per PCR es van dur a terme mitjançant *GeneAmp PCR System 9700 (PE Applied Biosystems)*. Els productes de PCR es van generar mitjançant reaccions de 50 µl que contenien 200 ng de DNA genòmic, 0.5 U de Taq DNA polimerasa (*Biotaq DNA Polymerase. Boline*), els *primers* (taula 12) a una concentració de 0.5 µM cada un, dNTPs a una concentració de 0.05 mM cada un, 1 mM o bé 1.5 mM de MgCl₂ (taula 12) i en alguns casos, 5% DMSO (taula 12) en tampó *Boline PCR buffer*.

El programa de PCR començava a 5 minuts a 94 °C per la desnaturalització inicial i era seguit per 30 cicles d'amplificació, consistents en 1 minut a 94 °C, 1 minut a la temperatura d'unió dels *primers* (taula 12) i 2 minuts a 72 °C. En l'últim cicle, l'extensió va ser de 72 °C 10 minuts. Els fragments amplificats es van visualitzar en una electroforesi en gel d'agarosa normal a l' 1%.

Fragment	Primers	Seqüències 5' 3'	Condicions de PCR			
			TH	cicles	MgCl ₂	DMSO
1 3119 pb	71 (-3355 -3337) 72 (-237 -256)	AGGCTTTTCCTCGCCAGTG AGAGCGGACGGTTTTGTTGC	61 °C	30	1,5 mM	
2 1830 pb	73 (-402 -382) 74 (1588 1567)	CGGTCTTGAGATTTGACTCGC CACACGATTATCTGGAAGGAAC	57 °C	35	1,5 mM	5%
3 1468 pb	77 (2899 2917) 78 (4367 4349)	CGCGGGCTGAGGCAGGTTC ACCACGTCCCTTCTGCGAG	59 °C	30	1,5 mM	5%
4 1000 pb	79B (5317 5338) 710B (6317 6297)	CCTCATGCTCAAAGAAGCCTCA CCTGTCAAAGACCTCAGACTG	59 °C	30	1,5 mM	
5 705 pb	79C (6138 6157) 710.1 (6862 6843)	CCCACTTTGGGTCCCATATT GTGCCTGGTCAGCTGGGTCT	59 °C	30	1 mM	
6 1099 pb	711.3 (6784 6803) 712.1 (7883 7864)	GGGCTCAATGACATAGACCCA GCCGTTCTCGTTCACACAGA	57 °C	30	1,5 mM	5%
7 702 pb	711.1 (7737 7757) 712.2 (8439 8420)	ACCTTCCAGGCAGAACACCAC CCCTGCTTTTGGGAAGTGCAG	59 °C	30	1,5 mM	5%
8 1165 pb	711.2 (8313 8331) 712 (9478 9450)	CGGCACAGCCAATGTCTGTA TCTAGCCGAGACGTGCTCTTG	59 °C	30	1,5 mM	5%
9 1701 pb	713 (9310 9328) 714 (9706 9725)	CGAGTTGTCACGTCGTCCTC ACTGTCCCCCTTGCAGGAGT	61 °C	30	1,5 mM	
10 1510 pb	715 (10809 10827) 716 (12319 12298)	TTTCATTGGTCAGCGGCT GGGTTTCATTCAGTGATGTTGA	59 °C	30	1,5 mM	5%

Taula 12: Primers i condicions utilitzades per l'amplificació dels diferents fragments de PCR del gen *F7*. TH és la Temperatura d'Hibridació.

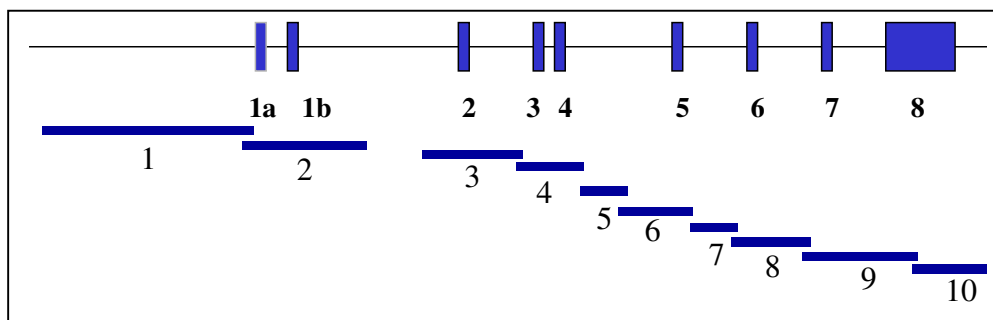


Figura 8: Esquema del gen *F7* amb els fragments utilitzats per la seva amplificació.

5.2- Seqüenciació del DNA

La tècnica de seqüenciació utilitzada està basada en la metodologia descrita per Sanger *et.al.* (1977). El punt clau d'aquesta metodologia és la utilització de dideoxinucleòtids (ddNTPs). Els ddNTPs són un tipus de nucleòtids modificats en el seu extrem 3' de manera que no poden unir-se a cap altre nucleòtid per

aquest extrem. Això provoca que un cop incorporat un ddNTP en la cadena de DNA ja no es pugui incorporar res més, i per tant la síntesi de DNA s'aturi.

La reacció de seqüència és una reacció de PCR asimètrica a partir d'un sol *primer*, i amb una proporció dNTPs/ddNTPs de 10:1. Cada ddNTP, a més, va marcat amb un fluorocrom d'un color diferent segons la base nitrogenada que porta: blau per la citosina, negre per la guanina, verd per l'adenina i vermell per la timina.

Com en el cas de la reacció de PCR, la polimerasa va incorporant nucleòtids dNTPs a partir del *primer* per complementarietat de seqüència, fins que, per atzar, s'incorpora un ddNTP que fa que s'aturi la reacció.

Com que estem parlant de milions de molècules, el resultat és una població de molècules de totes les mides possibles de la seqüència (mida del *primer* +1, mida del *primer* +2, ...) i cada grup de molècules de la mateixa mida, a més, van marcades amb el color de l'última base que s'ha incorporat, que és el ddNTP marcat amb el fluorocrom.

Aquesta reacció se sotmet a una electroforesi capil·lar en el seqüenciador automàtic, on les molècules s'ordenen per mida de menor a major i són estimulades per un làser a mesura que van passant a través del capil·lar. L'emissió de color és detectada pel seqüenciador automàtic i traduïda a la base nitrogenada corresponent.

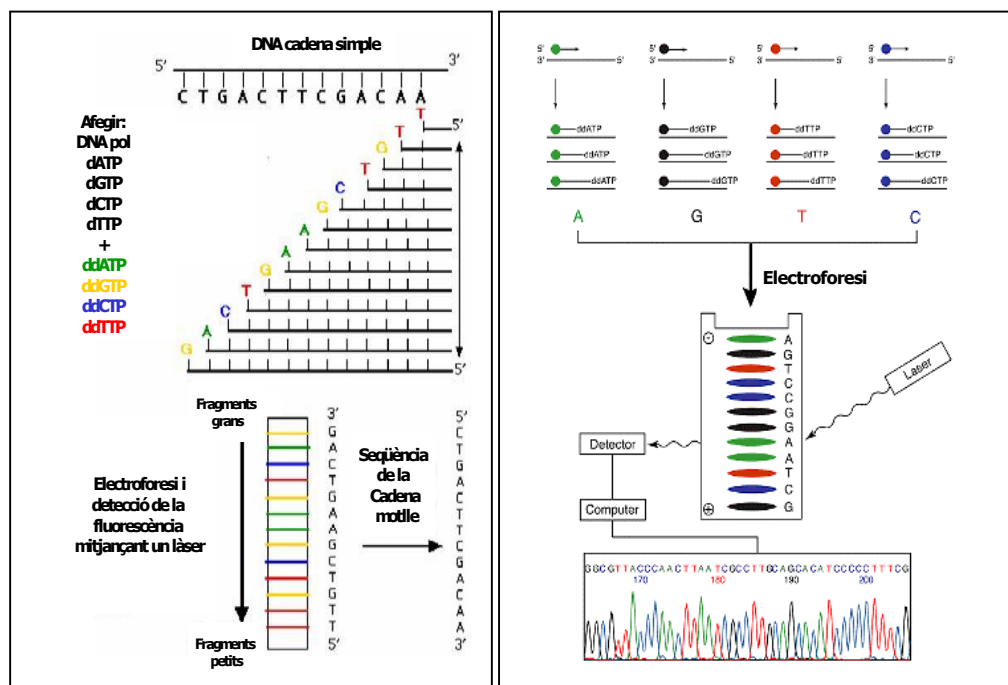


Figura 9: Mètode de seqüenciació a) imatge de Pharmacia Biotech Inc., b) imatge de Columbia Genome Center

Seqüenciació del gen *F7*

Els productes de PCR es van purificar amb el kit de purificació de *Quiagen 'QIAquick PCR Purification Kit* per ser seqüenciats. Amb aquesta purificació s'elimina l'excés de nucleòtids i de *primers* de la reacció de PCR.

La reacció de seqüència es va realitzar utilitzant un termociclador GeneAmp PCR system 9700, en un volum de 10 µl de reacció que contenia 3 µl del DNA purificat, 4 µl del kit de seqüenciació *DNA Sequencing Kit BigDye Terminator Cycle Sequencing Ready Reaction (Applied Biosystems)*, 5% de dimetilsulfòxid (DMSO vol/vol) i 0,32 µM del *primer* de seqüenciació (taula 13). El programa de seqüència començava amb 3 minuts a 94 °C i anava seguit de 25 cicles consistents en 10 segons a 96 °C, 5 segons a la temperatura d'unió dels *primers* (taula 13) i finalment 4 minuts a 60 °C; a excepció del *primer 73.6* amb el que es va usar un programa especial per seqüències riques en GC que consistia en 3 minuts de desnaturalització a 97.5 °C seguits per 25 cicles consistents en 30 segons a 96 °C i 4 minuts a 50 °C.

Les seqüències van ser realitzades amb el seqüenciador automàtic *ABI PRISM 3100 Genetic Analyzer*.

Fragment	Primers de seqüència	Seqüències 5' 3'	Hibridació
1	71 (-3355 -3337)	AGGCTTTTCCTCGCCAGTG	50 °C
	71.2 (-340 -360)	CGTAGATTCAAAATTCCTTATGGG	50 °C
	71.3 (-1190 -1166)	AAAGAAGTCATCAGCAATGCAACT	50 °C
	71.4 (-772 -753)	GGTGACGTGCACCTGTGGTC	50 °C
	71.8 (-1963 -1943)	CTCAGGGAGGTATGGAGCGG	50 °C
	72 (-237 -256)	AGAGCGGACGGTTTTGTTC	52 °C
	72.1 (-340 -360)	CAGATTTAAACTGCAGATGA	50 °C
	72.2 (-1166 -1190)	GTTGCATTGCTGATGACTTCTTT	50 °C
2	73 (-402 -382)	CGGTCTTGAGATTTGACTCGC	52 °C
	73.1 (86 101)	CCATAAACTTGGTGGAAGGGC	52 °C
	73.2 (415 434)	AGCACAGGTAGGGGACGGTG	52 °C
	73.6 (565 585)	GGCCTGGGAAACAGCATGTGG	50 °C
	74 (1588 1567)	CACACGATTATCTGGAAGGAAC	52 °C
	74.1 (1101 1080)	TGGCGTGTGGGGATGGCGAGT	52 °C
	74.2 (1080 1101)	ACTCGCCATCCCCACGCCA	52 °C
	3	77 (2899 2917)	CGCGGGCTGAGGCAGGTTC
77.1 (3213 3233)		AGGTCTGGAGCTCTCAGGGGT	50 °C
77.2 (3116 3136)		CCTGTTTGTCCACAAGTAAC	50 °C
77.3 (3830 3849)		CAGTCTTCGTAACCCAGGAG	50 °C
78 (4367 4349)		ACCACGTCCCTTCTGCGAG	50 °C
78.1 (3988 3969)		TCTCCGCGTCCTTGAAGATC	52 °C
4		79B (5317 5338)	CCTCATGCTCAAAGAAGCCTCA
	710B (6317 6297)	CCTGTCAAAGACCTCAGACTG	50 °C
5	79C (6138 6157)	CCCCTTTGGGTCCCATATT	50 °C
	710.1 (6862 6843)	GTGCTGGTCAGCTGGGTCT	50 °C
	710G (6295 6314)	TCCAGTCTGAGGTCTTTGAC	50 °C
6	711.3 (6784 6803)	GGGCTCAATGACATAGACCCA	50 °C
	711.5 (7195 7214)	TGGTCATCTGGGTCCAGAAT	50 °C
	712.1 (7883 7864)	GCCGTTCTCGTTCACACAGA	50 °C
7	711.1 (7737 7757)	ACCTTCCAGGCAGAACACCAC	50 °C
	712.2 (8439 8420)	CCCTGCTTTTGGAAAGTGCAG	50 °C
8	711.2 (8313 8331)	CGGCACAGCCAATGTCTGTA	50 °C
	711.4 (8716 8736)	CCTGACCATTGTCTCCTCAG	50 °C
	712 (9478 9450)	TCTAGCCGAGACGTGCTCTTG	50 °C
9	713 (9310 9328)	CGAGTTGTACGTCGTCCTC	50 °C
	713.2 (10503 10522)	AGCCCCTGCAGACCTAGAAA	52 °C
	713.3 (9706 9725)	TGATCAACACCATCTGGGTG	52 °C
	714 (9706 9725)	ACTGTCCCCCTTGCAGGAGT	50 °C
	714.1 (10522 10503)	TTTCTAGGTCTGCAGGGGCT	50 °C
	714.2 (10201 10181)	TGGGCTCTTGGTCAAGTGAG	52 °C
	10	715 (10809 10827)	TTCTCATTTGGTCAGCGGCT
715.1 (11259 11278)		ATCCCATATATTCTTCTGCA	50 °C
716 (12319 12298)		GGGTTCAATTCAGTGATGTTGA	50 °C
716.1 (12808 12789)		GTGCGTGCATCCATGTGTAT	50 °C

Taula 13: Primers i condicions utilitzades per les seqüències dels fragments d'amplificació.

El producte de la reacció de seqüenciació es va purificar mitjançant precipitació per etanols per eliminar les restes de *primer* i nucleòtids no incorporats. Per això, es van afegir 10 µl d'H₂O als 10 µl del volum de la reacció i a continuació es va afegir etanol (Merck) a la concentració adequada per tal que la concentració a volum final fos del 60%.

Es va centrifugar 20 minuts a 14.000 rpm en una microcentrífuga i es va decantar l'etanol.

Els precipitats es van rentar amb 200 µl d'etanol al 70% i es van tornar a centrifugar 5 minuts a 14.000 rpm. Es va decantar la solució d'etanol, els precipitats es van assecar al màxim i es van resuspendre en 15 µl de formamida desionitzada just abans de procedir a l'electroforesi capil·lar en el seqüenciador automàtic ABI PRISM 3100, seguint les instruccions del proveïdor (Applied Biosystems, USA).

5.3-Identificació de les variants de seqüència

Les seqüències van ser alineades amb el programa Seqman (Lasergene, DNASTar) i comparades amb la seqüència publicada per O'Hara (1987). L'alineament de les 40 seqüències amb l'original va permetre detectar les variants al·lèliques així com els errors de la seqüència anteriorment publicada. Els resultats d'aquestes anàlisis de seqüències van ser sotmesos a GeneBank (nombre d'accés AY212252).

Les seqüències de primats es van alinear utilitzant el programa d'alineaments mVISTA (Mayor, et al., 2000). Els resultats d'aquestes anàlisis van ser sotmesos a GeneBank (nombres d'accés DQ142912-DQ142919).

5.4.- Detecció de mutacions al total de la mostra

Una vegada detectades totes les variants al·lèliques per seqüenciació dels 40 individus seleccionats, les variants van ser genotipades en els 398 membres de la mostra GAIT.

En alguns casos, es van dissenyar fragments de PCR més petits dirigits només a amplificar la zona del polimorfisme a genotipar (taula 14). En altres, la genotipació es va fer directament a partir del fragment original (Taula 12).

Fragment	Primers	Seqüències 5' 3'	condicions de PCR			
			TH	cicles	MgCl ₂	DMSO
A 404 pb	71 (-3355 -3337)	AGGCTTTTCCTCGCCAGTG	59 °C	30	1 mM	
	71.9A (-2952 -2971)	GTGATTTCCAGGCTCTTCCA				
B 536 pb	71.4 (-772 -753)	GGTGACGTGCACCTGTGGTC	57 °C	30	1 mM	5%
	72 (-237 -256)	AGAGCGGACGGTTTTGTTGC				
C 806 pb	75.2 (-393 -372)	GATTTGACTCGCATGATTGCT	57 °C	30	1 mM	5%
	76.2 (431 413)	CGTCCCCTACCTGTGCTCC				
D 819 pb	73.2 (415 434)	AGCACAGGTAGGGGACGGTG	62 °C	35	1 mM	5%
	76.3 (1214 1234)	GCATGTCCCCTGTTTCTCCTC				
E 193 pb	75191 (5230 5249)	CACAGGTCGCCACACCTCGC	57 °C	30	1 mM	
	710A (6295 6315)	AAAAATGCTAGGCATGACCATC				
F 291 pb	IVS7a (9628 9647)	AATGTGACTTCCACACCTCC	57 °C	30	1 mM	
	IVS7b (10129 10110)	GATGTCTGTCTGTCTGTGGA				
G 290 pb	H115a (7746 7765)	GCAGAACACCACTGCTGACC	59 °C	30	1 mM	
	H115b (8037 8018)	AGTGGGACAGGGACTGGTGT				

Taula 14: *Primers* i condicions utilitzades per l'amplificació dels fragments de genotipació. Els set fragments de genotipació s'anomenen amb lletres per tal de distingir-los dels fragments utilitzats per l'amplificació inicial. TH és la temperatura d'hibridació.

En aquests casos, les reaccions de PCR constaven de 30 cicles (35 pel fragment D) consistents en 30 segons a 94 °C (1 minut pels fragments C i D), 30 segons a la temperatura d'hibridació dels *primers* (1 minut pels fragments C i D), i 30 segons a 72 °C (1 minut pels fragments C i D) (Taula 14).

La majoria de les variants al·lèliques es van genotipar per seqüenciació directa, especialment si es podien genotipar dues o més variants al·lèliques amb un mateix *primer* de seqüenciació. Aquest va ser el cas dels següents polimorfismes:

- Les variants A-670C, A-630G, G-402A, G-401T i -323ins0/10 del promotor es van seqüenciar amb el primer 71.4 a partir del fragment B
- La variant T-122C del promotor, i les variants, G73A, A260G, G369A de l'intró 1 es van seqüenciar amb el primer 76.2 (hibridació 50 °C) a partir del fragment C.
- Les variants T700C, G707A, C712G, 725-826del, T801C, G808A, C813G, C835G, i G946A de l'intró 1 es van seqüenciar amb el primer 73.6 (taula 13) a partir del fragment D.
- Les variants G3176A, G3300A, C3390T i G3428T de l'intró 2 es van seqüenciar amb el primer 77.2 a partir del fragment 3 (taula 13).

- Les variants G3938A i G4015A de l'intró 3 es van seqüenciar amb el primer 77.3 a partir del fragment 3 (taula 13).
- Les variants G6330A, G6441T, G6445T, i 6455VNTR de l'intró 5 van ser seqüenciades amb el primer 710G a partir del fragment 5 (taula 13).
- Les variants G7147C, T7460G, i G7718A de l'intró 5 es van seqüenciar amb els primers 711.3, 711.5 i 712.1 respectivament, a partir del fragment 6 (taula 13).
- La variant G8676A de l'intró 6 es va seqüenciar amb el primer 711.2 a partir del fragment 8 (taula 13).
- Les variants A9717G, T9762C, G9775A i C9830T de l'intró 8 es van seqüenciar amb el primer 713.3 a partir del fragment 9 (taula 13).
- Les variants G10507A, T10517C de l'intró 8 i C10782T (A354V) de l'exó 9 es van seqüenciar amb el primer 714 a partir del fragment 9 (taula 13).
- Les variants 10900 (S393S) i G10959A (R413Q) de l'exó 9 es van seqüenciar amb el primer 715 a partir del fragment 10 (taula 13).
- Les variants 11276-7insAA i 11603-4delAG de la zona 3' no traduïda, es van seqüenciar amb el primer 715.1 a partir del fragment 10 (taula 13).
- Finalment, la variant G11891A de la zona 3' no traduïda, es va seqüenciar amb el primer 716 també a partir del fragment 10 (taula 13).

Anàlisi de fragments de restricció de longitud variable (*Restriction Fragment Length Polymorphisms, RFLPs*)

Un dels mètodes més utilitzats per la detecció de variants al·lèliques conegudes és l'anàlisi de fragments de restricció de longitud variable. Es basa en la capacitat dels enzims de restricció de reconèixer seqüències específiques del DNA. Per a genotipar una determinada mutació, es busquen enzims de restricció que reconeguin específicament un al·lel i no l'altre i es visualitzen en un gel d'agarosa que separa els fragments segons la mida.

En el cas que la mutació no modifiqui cap seqüència de restricció, es poden utilitzar *primers* de mutagènesi en el qual s'introdueix una diana de restricció artificialment mitjançant el *primer*.

- La variant C-2989A del promotor es va genotipar amplificant el fragment A (taula 14) i digerint-lo amb l'enzim *ApoI* a 50 °C tota la nit.
- La variant G5194A es va genotipar introduint una diana de restricció per *BstUI* amb el *primer* directe (fragment E, taula 14) i digerint a 60 °C tota la nit.
- La variant T5506A es va genotipar amplificant el fragment 4 (taula 13) i digerint-lo amb l'enzim *SfaNI* a 37 °C durant 4 hores.
- La variant C7869T (H175H) es va genotipar amplificant el fragment G (taula 14) i digerint-lo amb l'enzim *NlaIII* a 37 °C tota la nit.

Per últim, la variant IVS7, en ser un polimorfisme de repetició (VNTR) de 37 pb, es va poder genotipar separant el producte de PCR (fragment F) per mida. De tota manera, per facilitar la visualització, el fragment es va digerir amb l'enzim *RsaI* tota la nit. Aquest enzim talla pels extrems de la repetició i facilita la distinció dels fragments en un gel d'electroforesi.

Tots els polimorfismes detectats es van sotmetre a la base de dades de polimorfismes del National Center for Biotechnology Information (NCBI) (dbSNPs) amb els números de submissió ss7844209-ss7844256.

5.5- Anàlisi de les seqüències de les diferents espècies de primats

Les seqüències de primats es van alinear respecte a la seqüència de referència humana sotmesa a partir d'aquest treball (AY212252) per localitzar les posicions polimòrfiques detectades en humans. L'anàlisi de la divergència entre seqüències es va analitzar amb el programa DnaSP (Rozas and Rozas, 1999).

5.6- Construcció dels haplotips

Els haplotips per tots els cromosomes independents de la mostra GAIT es van construir utilitzant la informació de les famílies per cada individu fundador, mitjançant el paquet informàtic MRH (*Minimum Recombinant Haplotype*, v.0.2) (Qian and Beckmann 2002).

6- MÈTODES ESTADÍSTICS PER A L'ANÀLISI DE LA VARIABILITAT

6.1- Bayesian Quantitative Trait Nucleotide

Quan totes les variants al·lèliques han estat identificades, cal saber quines són les funcionals. Mitjançant avançats mètodes d'estadística genètica (*Bayesian Quantitative Trait Nucleotide*) i a partir de la informació completa de tots els polimorfismes genotipats en el total de la mostra, es pot determinar estadísticament quin polimorfisme o polimorfismes identificats són amb més probabilitat els responsables de la variabilitat dels nivells de la proteïna que s'està analitzant. Per fer-ho, es consideren tots els polimorfismes del gen i s'estimen tots els models possibles considerant totes les possibles combinacions dels polimorfismes amb la intenció d'obtenir el model que expliqui millor les variacions dels nivells de la proteïna (Almasy, et al., 2001). Si només considerem els efectes additius dels polimorfismes, hi ha 2^m models possibles, on m és el nombre de polimorfismes considerats. Un cop estimats els models, es calcula una mitjana ponderada de l'efecte de cada polimorfisme en els diferents models.

$$Y = \mu + \sum \alpha_i s_i + \sum q_k + \sum cov + g + e$$

És a dir, que la variació del fenotip és deguda a una mitjana, més la variació deguda a un polimorfisme en concret, més la variació deguda a la resta dels polimorfismes, més la variació deguda a les covariables controlades, més l'efecte d'altres factors genètics i ambientals no controlats.

A partir d'aquests models, s'estima la probabilitat posterior de cada variable de ser funcional. Els procediments per calcular els models han estat incorporats al paquet informàtic SOLAR, disponible a <http://www.sfbr.org>.

Un cop identificades les variants genètiques amb més probabilitat de ser funcionals, es va realitzar una anàlisi de lligament condicionada a les determinacions genotípiques funcionals, és a dir, considerant aquestes variants com a covariables controlades. En tal cas, si el conjunt de les variants identificades expliquen tota la variabilitat en el fenotip no s'hauria d'observar lligament genètic. Això és degut a que tota la variància genètica que és deguda al QTL ha estat eliminada en forma de covariable.

6.2- Anàlisi de la diversitat del DNA

Les anàlisis de la diversitat del DNA, tant dins de la nostra espècie com entre espècies diferents, ens ajuden a comprendre la història evolutiva que ha patit la nostra espècie, així com la història evolutiva

dels gens concrets sota estudi, ampliant el nostre coneixement sobre els fenòmens genòmics i poblacionals que han contribuït a la variació constant dels nostres gens.

6.2.1- Variabilitat intraespecífica

La majoria de canvis en el genoma són selectivament neutres, és a dir, no provoquen un canvi en l'eficàcia biològica de l'individu (Kimura, 1983). En aquests casos, les noves variants al·lèliques que s'introdueixen en població són eliminades o fixades per deriva genètica, mantenint-se fins llavors com a polimorfismes. El temps que triga una variant al·lèlica nova a fixar-se per deriva depèn de la mida efectiva de la població. No obstant, els al·lells que estiguin afectats per la selecció provocaran un canvi addicional en les freqüències al·lèliques en la població, que afectarà tant a la freqüència dels *loci* seleccionats com en els que estiguin en desequilibri de lligament amb ells (*genetic sweep*).

L'estudi de la variabilitat intraespecífica, doncs, ens serà indicativa dels tipus de forces selectives que han actuat sobre el gen estudiat, i de l'evolució molecular que ha patit.

Una de les mesures importants a l'hora d'estudiar la variabilitat intraespecífica és la mesura del grau de polimorfisme observat respecte l'esperat, i es quantifica mitjançant el **Test de Tajima (D) o test de neutralisme**. Aquesta mesura ens dóna un patró de la distribució dels polimorfismes en una determinada regió.

$$D = k - \theta$$

- El grau de polimorfisme observat (k) mesura el nombre mitjà de diferències nucleotídiques comparant les seqüències de dues en dues.
- El grau de polimorfisme esperat (θ), per contra, estima la variació en una població amb taxa de mutació μ i en equilibri mutació-deriva per seqüència i per generació amb N_e com a mida efectiva de la població.

Segons la teoria neutralista (Kimura, 1983) s'estima que el temps que triga una mutació neutra des que apareix fins que es fixa és de $4N_e$ generacions. La mida efectiva d'una població és la mida de la població biològicament activa.

- D és positiu quan $k > \theta$, és a dir, quan observem molt més polimorfisme de l'esperat. Hi ha un excés polimòrfic de freqüències intermitges que ens indica l'efecte de selecció equilibradora, probablement degut a un avantatge dels heterozigots.

- D és negatiu quan $k < \theta$, és a dir, quan observem molt menys polimorfisme de l'esperat. Això pot ser degut a una selecció direccional o bé a un efecte de coll d'ampolla o d'expansió. Per diferenciar aquests dos fenòmens s'han de comparar diferents gens. L'efecte d'una selecció es donarà en el gen seleccionat determinat mentre que l'efecte de coll d'ampolla o d'expansió s'observarà de forma generalitzada en tot el genoma.
- Finalment, $D = 0$ quan no hi ha cap força selectiva que afecti les seqüències estudiades i per tant es confirma la hipòtesi nul·la de la teoria neutralista.

Una altra mesura interessant per analitzar el grau de polimorfisme observat en la mostra és la **diversitat nucleotídica** (π), que es defineix com l'heterozigositat mitjana per posició nucleotídica, és a dir, és la probabilitat que dues seqüències a l'atzar difereixin en una posició concreta. Aquest paràmetre no depèn de la mida de la seqüència per la qual cosa es pot usar per comparar mostres de diferent longitud.

Mesura del desequilibri de lligament

El coeficient de desequilibri de lligament (D) és la diferència entre les freqüències haplotípiques observades *versus* les esperades tenint en compte les freqüències al·lèliques. Per dues variants al·lèliques concretes, la freqüència haplotípica esperada serà el producte de les freqüències al·lèliques. Per estandaritzar aquesta mesura, una de les mesures de desequilibri de lligament més utilitzades és D' (Lewontin, 1964), que és el quocient entre D i el desequilibri de lligament màxim possible donades les freqüències al·lèliques determinades (D'). D'aquesta manera es poden fer comparacions de desequilibri de lligament entre *loci*.

$$D = \text{freq obs} - \text{freq esp}$$

$$D' = D / D_{\text{max}}$$

D' té valors entre -1 i 1, essent positiu quan els al·lèls estan lligats i negatiu quan estan en repulsió. Un valor = 0 indica equilibri de lligament.

Anàlisi dels haplotips

L'estudi dels haplotips, i la seva agrupació segons la semblança, ens indica quins d'ells són evolutivament més propers, i per tant afegeix informació sobre la història evolutiva del gen i ens indica com s'ha anat generant la diversitat observada en la població.

A partir de la semblança dels haplotips entre ells, es poden construir *networks* filogenètics, que són esquemes gràfics que connecten els haplotips entre ells amb el menor nombre global de canvis possible, entenent com a canvis les diferències nucleotídiques. Els punts representen els haplotips i la seva mida és proporcional a la freqüència de l'haplotip en la població, i les línies els uneixen en funció de la proximitat, essent més llargues quan més canvis nucleotídics hi ha entre els haplotips que uneixen. Cal destacar que les recombinacions provocaran greus distorsions en el *network*.

Els *networks* filogenètics ajuden a explicar la història evolutiva del gen, i a més ens permeten estimar l'edat dels grups d'haplotips en funció de les mutacions acumulades en cada branca respecte de la seqüència ancestral.

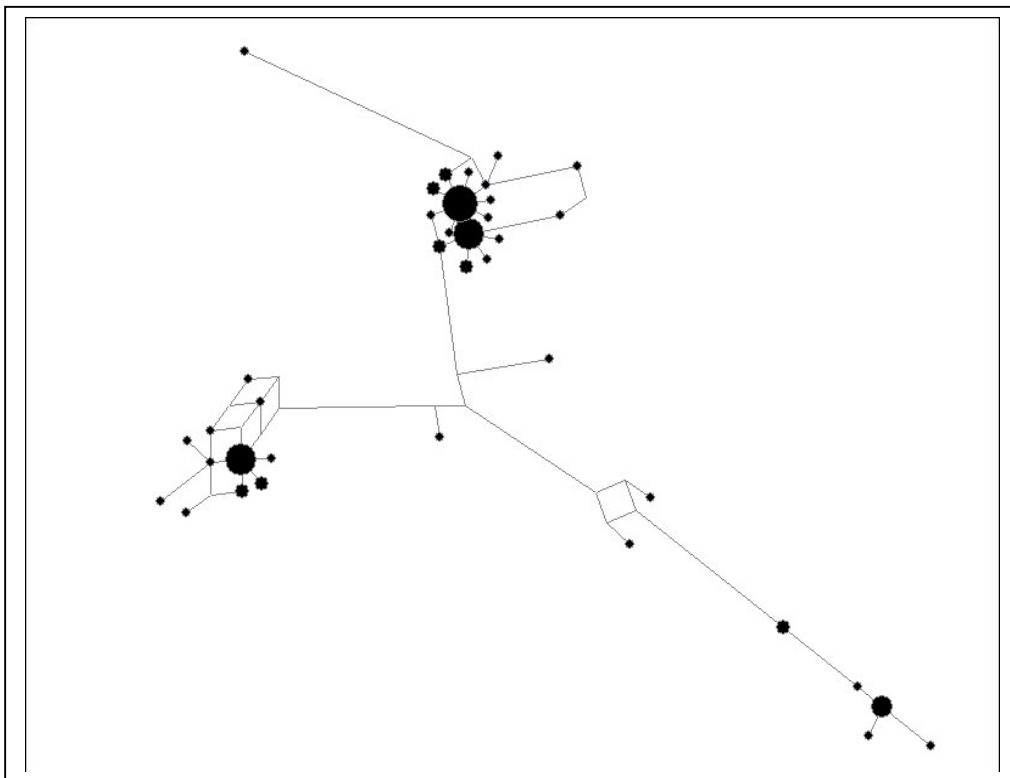


Figura 10: Representació d'un Network filogenètic

6.2.2- Variabilitat interespecífica

L'estudi de la variabilitat entre seqüències homòlogues d'espècies properes ens és molt útil per estudiar la filogènia del nostre gen.

La **divergència (Dxy)** és el paràmetre utilitzat per analitzar com de diferents són les seqüències de les espècies estudiades, i bàsicament mesura el nombre de diferències nucleotídiques entre les seqüències agafades de dues en dues.

Amb la divergència, es pot fer un esquema de la filogènia del gen estudiat. La taxa en què una seqüència ha anat acumulant variació en el temps se suposa constant; per tant, la mesura de la divergència ens dóna una estima dels temps de divergència de les diferents espècies, de manera que es poden establir arbres filogenètics en funció del nombre de diferències nucleotídiques observades.

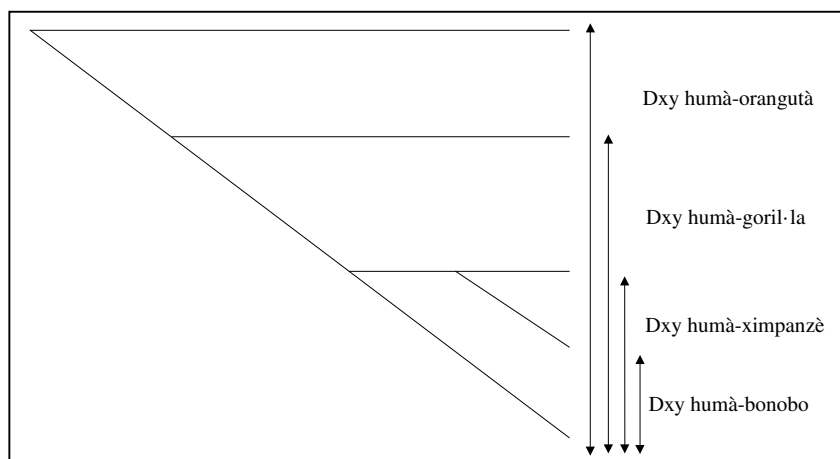
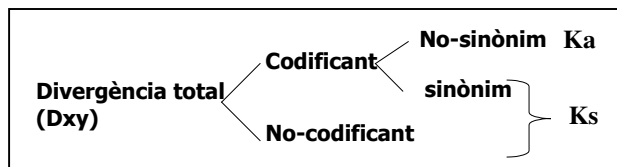


Figura 11: Exemple de l'esquema de la filogènia d'un gen calculada a partir de les divergències entre espècies.

A més de la divergència total, una altra informació interessant l'obtenim amb el desglossament d'aquesta divergència en dos grups, segons si el canvi nucleotídic en qüestió provoca o no un canvi d'aminoàcid.

En aquest sentit, es considera

- **Ka** com la divergència en les zones codificants que provoquen un canvi d'aminoàcid.
- **Ks** com la divergència de les zones no codificants, o bé en les zones codificants però que no provoquen un canvi d'aminoàcid.



La comparació d'aquests dos paràmetres ens dona una estimació del tipus de selecció que ha patit el gen estudiat :

- **Ka = Ks** indica que els canvis nucleotídics s'han produït aleatòriament al llarg del gen i, per tant, que no hi ha hagut cap tipus de pressió de selecció.
- **Ka > Ks** indica que hi ha més canvis que produeixen un efecte en el fenotip que canvis neutres i, per tant, que existeix una pressió de selecció positiva que afavoreix el canvi de la proteïna.
- **Ka < Ks** indica que hi ha més canvis neutres que canvis que tinguin un efecte en el fenotip i, per tant, que existeix una pressió de selecció purificadora que elimina la variació que produeix un canvi en el fenotip, probablement perquè la proteïna en qüestió té una funció essencial per a l'organisme que no li permet canvis.

7- ESTUDIS FUNCIONALS

Finalment, quan l'estadística genètica prioritza un grup de polimorfismes que són els que demostren tenir una probabilitat més alta de tenir un efecte sobre el fenotip estudiat (nivells plasmàtics de FVII), l'últim pas són els estudis funcionals, per acabar determinant amb seguretat quins són els polimorfismes que afecten els nivells d'expressió i quins els mecanismes fisiopatològics que intervenen en aquests canvis.

En aquest treball, s'han analitzat les variants al·lèliques localitzades a la zona promotora del gen *F7*. Per això, s'ha utilitzat un sistema *reporter*, que consisteix en la utilització *in vitro* d'una construcció genètica que conté el promotor d'interès seguit d'un gen de fàcil detecció i quantificació al laboratori. En aquest cas, s'ha utilitzat el gen de la Proteïna Verda Fluorescent (*GFP*, acrònim de l'anglès *Green Fluorescent Protein*). Aquest, codifica per una proteïna de 28 KDa aïllada de medusa que emet fluorescència, per la qual cosa és fàcilment quantificable amb un luminòmetre (Morise, et al., 1974; Prasher, et al., 1992).

7.1- Clonatge

La tècnica del clonatge consisteix en la introducció d'un fragment de DNA d'interès en un bacteri, en aquest cas *Escherichia coli* (*E. coli*), per tal que el bacteri repliqui el fragment en múltiples còpies idèntiques i puguem generar així suficient material per realitzar els experiments d'interès.

El primer pas en la tècnica de clonatge, doncs, és obtenir el fragment de DNA d'interès mitjançant PCR a partir de material genòmic.

El següent pas, és introduir aquest fragment en una molècula de DNA extracromosòmica capaç d'autoreplicar-se dins el bacteri: **el plàsmid**. Per això, tant el plàsmid com el fragment d'interès s'han de digerir amb els mateixos enzims de restricció, per tal que puguin ser lligats per formar la molècula recombinant.

Finalment, aquesta molècula s'ha d'introduir dins el bacteri per un procés anomenat transformació. A partir d'aquí, el bacteri replicarà el plàsmid amb el fragment d'interès de forma infinita.

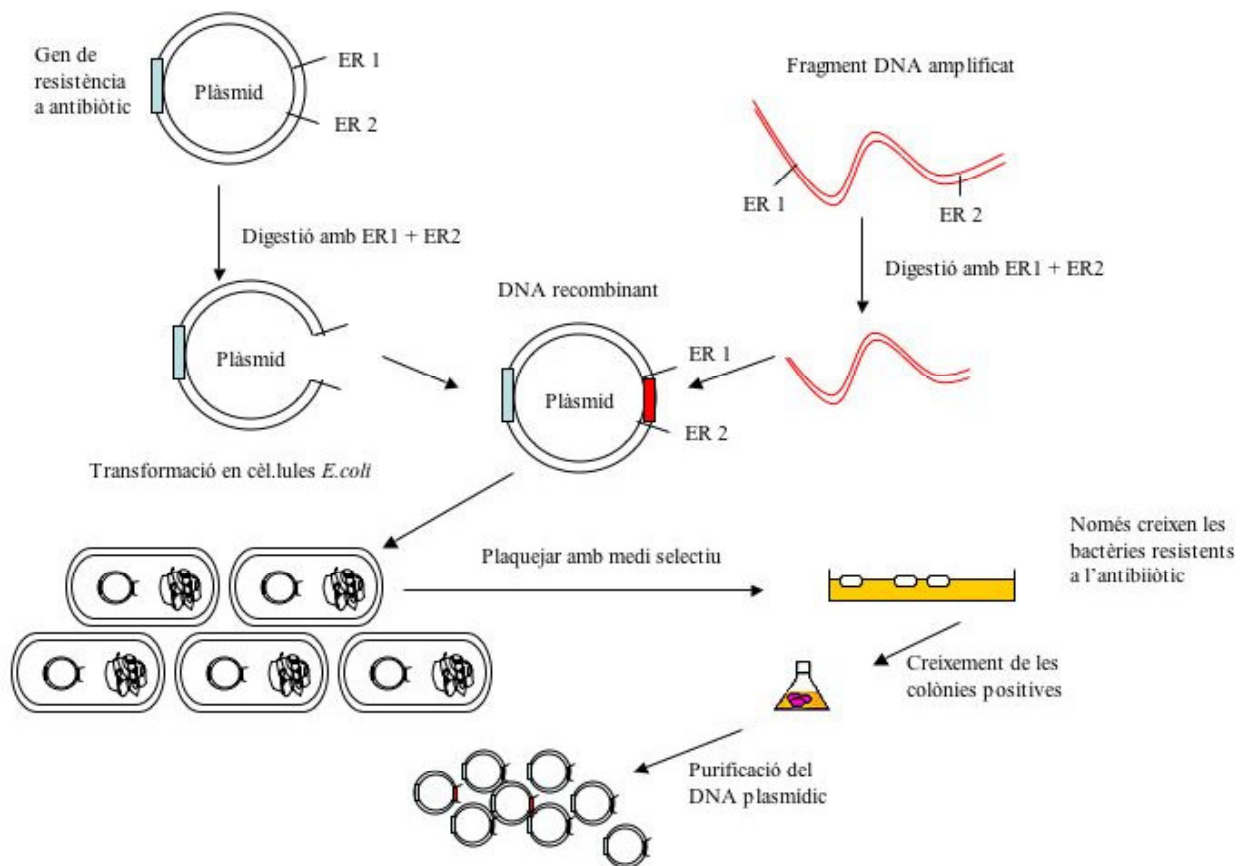


Figura 12: Esquema del procés de clonatge

ER: Enzim de Restricció

Existeixen diferents soques bacterianes amb característiques diferents. La soca bacteriana utilitzada per realitzar les anàlisis funcionals dels polimorfismes del promotor va ser la TOP10 de *E.coli*.

Creixement bacterià

Quan es fa créixer *E. coli* en un cultiu líquid, s'inoculen algunes cèl·lules en medi líquid estèril, on, després d'un període de temps anomenat *lag*, les cèl·lules començaran a dividir-se. Si el medi és ric (conté aminoàcids, vitamines, precursors nucleotídics i altres metabolits apart de l'aport de carboni i sals), les cèl·lules entren en fase exponencial i es dupliquen cada 20-30 minuts fins arribar a una densitat d'uns $1-2 \times 10^9$ cèl·lules/mL.

El creixement bacterià en aquest treball es va realitzar en medi de cultiu *Luria Broth* (LB), tant en estat líquid com sòlid (+2% agar) esterilitzat en autoclau (121 °C 20 minuts).

LB

10 g Triptona/l

5 g llevat/l

5 g NaCl/l, ajustat a pH 7 amb NaOH

El Plàsmid

Els plàsmids són molècules de DNA de doble cadena circular extracromosòmic capaces d'autoreplicar-se, que es troben en gairebé totes les espècies bacterianes i contenen informació addicional per la vida del bacteri (resistència a antibiòtics, sensibilitat a mutàgens, producció d'enzims de restricció o d'aminoàcids rars, toxines...).

Ja als anys 70, es va pensar en utilitzar aquests plàsmids (sobretot plàsmids d'*E.coli*) com a vectors per propagar i manipular seqüències de DNA d'interès (Bolivar, et al., 1977a; Bolivar, et al., 1977b; Jackson, et al., 1972; Kahn, et al., 1979).

Els tres elements essencials dels vectors plasmídics són els següents:

- un **inici de replicació** per tal que mantingui la capacitat autoreplicativa un cop introduït al bacteri i un fragment de DNA que contingui els gens necessaris per la replicació del plàsmid.
- un gen que li confereixi **resistència a un antibiòtic** amb el qual poder seleccionar aquells bacteris que hagin incorporat el vector.

- un **lloc de clonació múltiple** (*Multiple Cloning Site*, o *Polylinker*) amb múltiples dianes de restricció úniques per introduir el fragment d'interès sense que interfereixi en l'habilitat de replicació del plàsmid. Per tal d'agilitar la detecció de l'insert, la majoria de llocs de clonació múltiple estan dissenyats específicament per tal que disruptin un gen *LacZ*, que produeix un enzim anomenat beta-galactosidasa que resulta en la formació de colònies blaves en plaques que contenen IPTG (isopropil-beta-D-thiogalactopiranosid). Aquest component indueix l'activitat de la beta-galactosidasa mitjançant la seva unió i inhibició del repressor *lac* (Sambrook, 2001). La presència de l'insert disrupta l'expressió del gen i resulta en colònies blanques.

El vector utilitzat per la clonació i seqüenciació dels fragments de promotor amplificats per PCR va ser pGEM-T easy vector (PGEM-T Easy Vector System [Promega, Cat.#A1360]), un vector de 3 Kb que conté un lloc múltiple de clonació amb sistema de discriminació de colònies blanc/blau, i resistència a Ampicil·lina (figura 13).

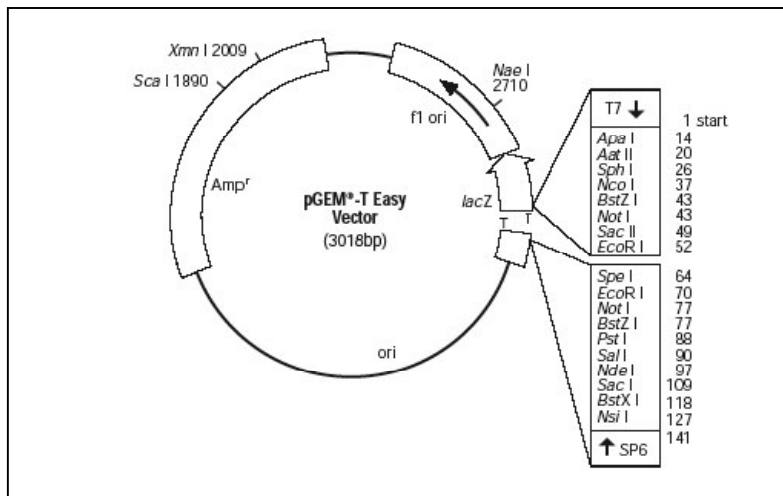


Figura 13: Representació esquemàtica del Vector de clonació pGEM-T-Easy Vector extret del catàleg de Promega.

Un cop comprovats els fragments amplificats, per analitzar els polimorfismes del promotor es va dissenyar una construcció que consistia en tres fragments de DNA clonats en el plàsmid comercial pBluescript (pBS) II KS – (Stratagene Cat.#212208). La seqüència del plàsmid es pot consultar a Genbank (#X52329).

pBluescript KS – és un vector comú de 3 Kb que conté un lloc múltiple de clonació amb sistema de discriminació de colònies blanc/blau, i resistència a Ampicil·lina (figura 14).

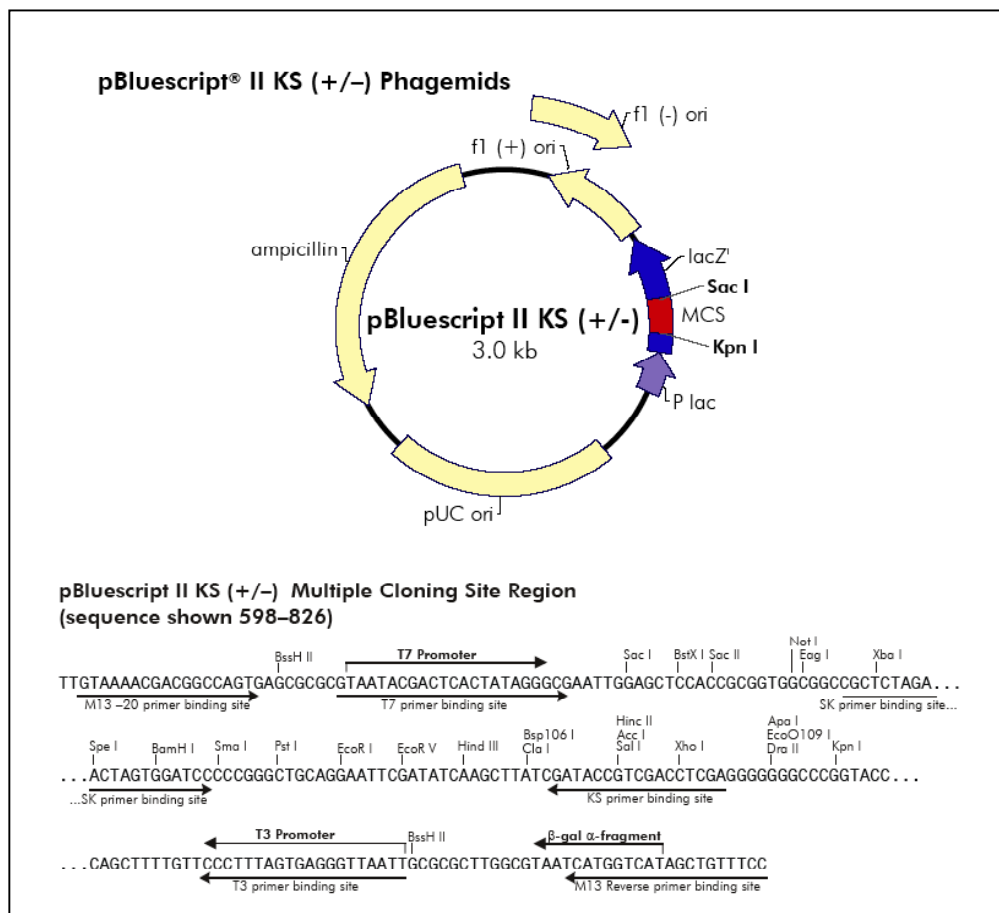


Figura 14: Representació esquemàtica del vector de clonació pBluescript II KS, extret del catàleg de Stratagene

Transformació

La introducció del DNA plasmídic dins la cèl·lula d'*E.coli* rep el nom de transformació. Per tal de poder transformar les cèl·lules, cal que aquestes adquireixin l'habilitat de poder introduir DNA exogen, estat en que s'anomenen **cèl·lules competents**.

Hi ha dos tipus bàsics de cèl·lules competents: les electrocompetents, i les quimiocompetents. Les cèl·lules electrocompetents són més eficients, però no permeten plàsmids més grans de 10 Kb.

Per preparar cèl·lules **electrocompetents** les cèl·lules es deixen créixer en medi LB líquid fins a una densitat òptica $OD_{600} = 0,6-0,9$ i es concentren per centrifugació. Per la transformació s'aplica un corrent elèctric d'alt voltatge a les cèl·lules que produeix uns porus transitoris per on entra el DNA (Shigekawa and Dower, 1988).

En el cas de les cèl·lules **quimiocompetents**, es parteix d'un cultiu *overnight* (12-16 hores de creixement) i es deixa créixer fins a una densitat òptica $OD_{600}=0,5$. El brou es concentra per centrifugació i es resuspèn en una solució que conté sals de calci (Mandel and Higa, 1970). Un derivat d'aquesta tècnica substitueix el calci per cations Rubidi (Kushner, 1978). En aquest cas, es resuspèn primer en un tampó amb 100 mM de RbCl, 50 mM de $MnCl_2$, 30 mM d'Acetat Potàssic, 10 mM de $CaCl_2$ i 15% de glicerol a pH 5,8. A continuació, es torna a concentrar per centrifugació i es resuspèn en un tampó amb 10 mM de RbCl, 10 mM de MOPS, 75 mM de $CaCl_2$ i 15% de glicerol a pH 6,8. L'exposició als ions dóna a les cèl·lules l'habilitat de poder introduir el DNA.

Per la transformació, s'uneix el DNA plasmídic a uns 100 μ l de cèl·lules quimiocompetents. La barreja és sotmesa a un xoc tèrmic a 42 °C, fet que permet que el DNA entri a la cèl·lula.

Les cèl·lules es deixen créixer aproximadament una hora en medi LB líquid sense antibiòtic per tal de permetre la seva recuperació i la síntesi de les proteïnes del plàsmid, i a continuació es sembren en plaques d'LB amb antibiòtic (en el nostre cas, 100 mg/L Ampicil·lina) per permetre la identificació de les colònies que han introduït el plàsmid.

Un cop realitzada la transformació, les plaques es deixen incubant tota la nit a 37 °C fins que s'observa la formació de colònies resistents a l'Ampicil·lina.

Extracció del DNA plasmídic

Aquestes colònies, es posen a créixer individualment en 2 mL de medi LB amb Ampicil·lina (Sigma) a una concentració de 125 mg/L. Passades 12-16 hores, es pot procedir a l'extracció del DNA plasmídic pel mètode de la lisi alcalina.

El cultiu saturat se centrifuga i el precipitat es resuspèn en un tampó amb Rnasa, TrisCl 0,5M i EDTA. A continuació es lisen les cèl·lules per incubació en NaOH 1M i SDS al 10% durant 5 minuts, i es neutralitza amb acetat potàssic, que fa precipitar les membranes i restes cel·lulars. Després d'una segona

centrifugació, es recupera el sobrenadant que conté el DNA plasmídic i aquest es precipita amb Isopropanol. El precipitat de DNA es neteja amb etanol al 70% i es resuspèn amb aigua o el tampó d'interès.

7.2- Construccions realitzades del promotor del gen *F7*

Com s'ha esmentat anteriorment, un cop comprovats els fragments amplificats, per analitzar els polimorfismes del promotor del gen *F7* es va dissenyar una construcció que consistia en tres fragments de DNA clonats en el plàsmid comercial pBluescript (pBS) II KS –

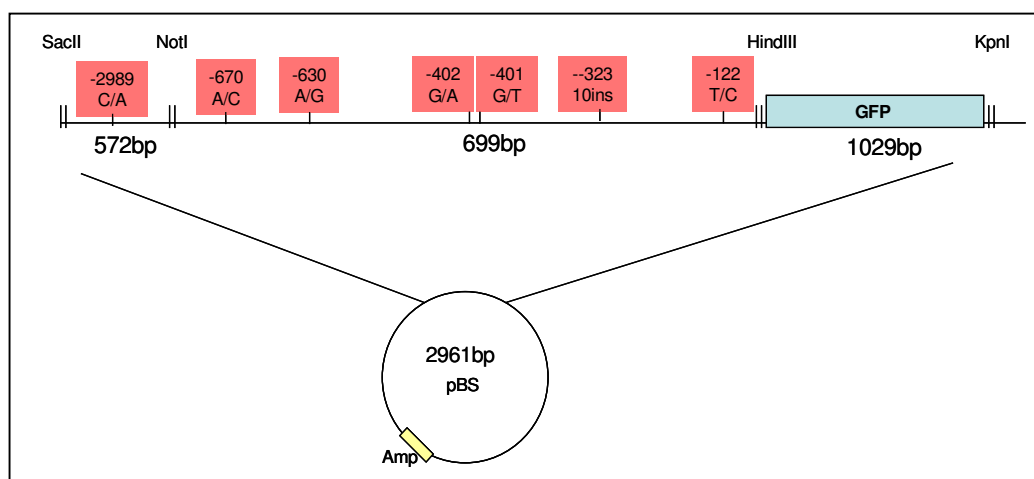
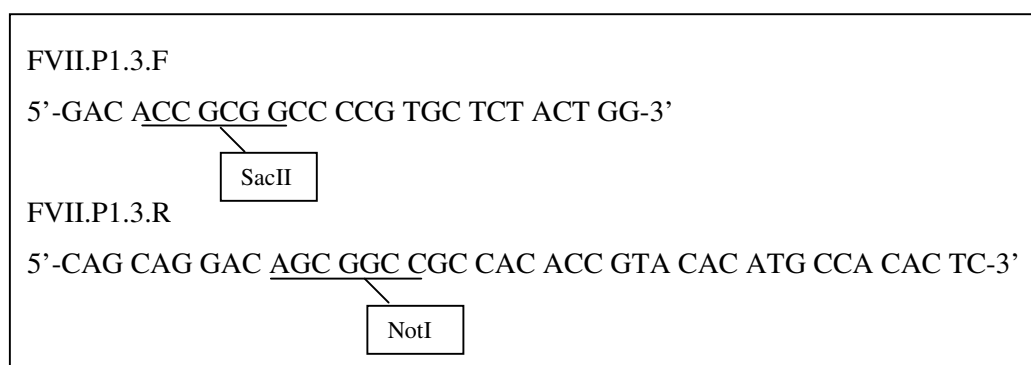


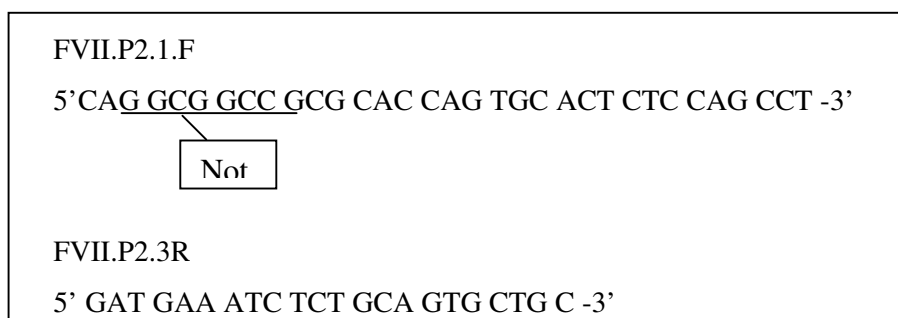
Figura 15: Esquema de les construccions realitzades. En vermell les posicions dels polimorfismes dels fragments del promotor. En blau, el gen reporter utilitzat.

Fragment 1: 565 pb del promotor distal que comprèn el polimorfisme –2989C/A, amplificat amb els següents primers:



La reacció d'amplificació es va dur a terme amb Taq Master Mix (Invitrogen) seguint les recomanacions del proveïdor, utilitzant un programa consistent en 15 minuts de desnaturalització a 95 °C, seguit per 30 cicles de 45 segons a 94 °C, 45 segons a 55 °C i 45 segons a 72 °C, i finalitzat en 10 minuts d'extensió a 72 °C.

Fragment 2: 682 pb del promotor proximal que comprèn els polimorfismes -670 A/C, -630 A/G, -402 G/A, -401 G/T, -323ins 0/10 i -122 T/C, amplificat amb els *primers* següents:



La reacció d'amplificació es va dur a terme amb Taq Master Mix (Invitrogen) seguint les recomanacions del proveïdor, utilitzant un programa consistent en 15 minuts de desnaturalització a 95 °C, seguit per 30 cicles d' 1 minut a 94 °C, 1 minut 55 °C i 1 minut a 72 °C, i finalitzat en 10 minuts d'extensió a 72 °C.

Ambdós fragments van ser prèviament clonats en pGEM mitjançant un kit comercial de lligació ràpida (PGEM-T Easy Vector System, Promega, Cat.#A1360) per poder ser seqüenciats amb facilitat. El DNA utilitzat com a motlle per amplificar aquests fragments va ser obtingut a partir d'individus de la mostra GAIT, cada un contenint un dels haplotips majoritaris del promotor. Com a resultat d'aquestes amplificacions es van obtenir les següents combinacions de mutacions:

- Pomotor WT (tots els al·lels normals)
- -670C/ -630G/ -402A
- -401T/ -323ins10/ -122C

Els fragments van ser verificats per seqüenciació. A continuació, el fragment 1 va ser introduït en pBS mitjançant digestió amb els enzims SacII i Not I (ambdós introduïts en el primer de la PCR), i relligació amb T4 DNA *ligase* (Biolabs) i el fragment 2 va ser introduït en pBS-fragment1 mitjançant digestió amb els enzims Not I i SpeI (el primer introduït en el *primer* de la PCR i el segon existent en el lloc de clonació múltiple del pGEM), i relligació amb T4 DNA *ligase* (Biolabs). Finalment, es va introduir el

fragment 3 que consistia en la regió codificant del gen GFP mitjançant digestió i relligació amb els enzims Hind III i Kpn I. La seqüència de GFP es pot consultar a Genbank (M62653).

Les reaccions de lligació es van dur a terme en tots els casos amb 50 ng de plàsmid pBS i una proporció d'insert 3:1 (insert:plàsmid) calculat segons la fórmula següent:

$$\text{ng insert} = \frac{\text{ng plàsmid} \times \text{mida insert}}{\text{mida plàsmid}} \times \frac{3}{1}$$

Els fragments clonats van ser transformats en cèl·lules *E.coli* TOP10 quimiocompetents per xoc tèrmic (45 segons a 42 °C). Les cèl·lules es van incubar en plaques de LB amb Agar i Ampicil·lina (LB +Amp) tota la nit a 37 °C. Les colònies individuals es van posar a créixer en medi líquid de LB i Ampicil·lina tota la nit a 37 °C i se'n va extreure el DNA plasmídic mitjançant el mètode detallat anteriorment. La introducció correcta dels fragments desitjats va ser comprovada per digestió.

7.3- Mutagènesi dirigida

Utilitzant aquesta construcció WT com a motlle, les variants -670 A/C, -630 A/G, -402 G/A, -401 G/T, i -122 T/C van ser introduïdes per mutagènesi dirigida mitjançant Dpn (Fisher and Pei, 1997) per acabar formant les 6 construccions diferents, cada una amb una de les variants introduïdes. La variant amb la inserció -323ins10 va ser obtinguda a partir de l'amplificació d'un individu amb l'haplotip -401T /-323ins10 /-122C al qual se li van introduir els al·lels normals en les posicions -401 i -122 per mutagènesi dirigida.

El mètode utilitza un vector de doble cadena amb l'insert d'interès i dos *primers* sintètics complementaris a cadenes oposades del plàsmid, i tots dos amb la mutació desitjada.

Amb aquests nucleòtids, es realitza una reacció de PCR, de manera que el nou plàsmid generat contindrà les mutacions introduïdes en els *primers*. A continuació, la reacció és digerida amb l'enzim DpnI. Aquest enzim reconeix la seqüència 5'-Gm⁶ATC-3', és a dir, és específic de DNA metilat i hemimetilat, de manera que digereix només la cadena parental que ha pogut ser metilada per *E.coli*. Com a resultat, s'obté una reacció on només hi ha plàsmid que ha incorporat la mutació. Aquesta reacció pot ser transformada en cèl·lules d'*E.coli* per produir noves colònies que contenen el plàsmid amb la mutació desitjada.

La reacció de mutagènesi es va dur a terme utilitzant la SuperMix Platinum PCR (Invitrogen, Cat.· 12532-016) segons les instruccions del proveïdor, amb els *primers* i el programa mostrats a continuació:

-2989.mutagen.F
5' ACC TGC CCG TAG ATT CAA ATT CCT TAT GGG TGT G 3'

-2989.mutagen.R
5' CAC ACC CAT AAG GAA TTT GAA TCT ACG GGC AGG T 3'

-670.mutagen.F
5' CTC TCC CGC CTG GCA ACA 3'

-670.mutagen R
5' CAG GCG GGA GAG TGC ACT 3'

-630.mutagen.F
5' AA CAA GAC AAA ACA AAC AAA AAG 3'

-630.mutagen.R
5' TTT GTC TTG TTT GTT TGA GAC A 3'

5' -402.mutagen.F
AAG ATA CAG TCT TGA GAT TTG 3'

-402.mutagen.R
5' CAA GAC TGT ATC TTG GGT G 3'

-401.mutagen.F
5' ATA CGT TCT TGA GAT TTG ACT 3'

-401.mutagen.R
5' TCA AGA ACG TAT CTT GGG TG 3'

-401WT.F
CAC CCA AGA TAC GGT CTT GAG ATT TGA CTC

-401WT.R
CGA GTC AAA TCT CAA GAC CGT ATC TTG GGT G

-122.WT.F
5' GGT GTT CAG AGG ACG CCT 3'

-122.WT.R
5' TCT GAA CAC CCC AGC TGA G 3'

-122.mutagen.F
5' GGT GCT CAG AGG ACG CCT 3'

-122.mutagen.R
5' TCT GAG CAC CCC AGC TGA G 3'

Taula 15: *Primers* utilitzats per la mutagènesi dels fragments del promotor.

Per a cada parella de *primers* de mutagènesi se'n van calcular dues temperatures d'unió (hibridació): una calculada pels primers cicles, en els quals la mutació introduïda forma un aparellament incorrecte (*mismatch*), i una altra calculada pels darrers cicles, en els qual la mutació ja s'ha introduït i totes les bases troben homologia (taula 16).

Programa:

5 minuts	95 °C					
30 segons	94 °C	}	x7	Primers	Hibridació 1	Hibridació 2
30 segons	hibridació 1			-2989	60°	63°
4 minuts	72 °C			-670	42°	50°
30 segons	94 °C	}	x28	-630	42°	50°
30 segons	hibridació 2			-402	42°	50°
4 minuts	72 °C			-401	42°	50°
10 minuts	72 °C			-401wt	45°	52°
5 minuts	25 °C			-122	48°	54°
∞	4 °C			-122wt	48°	54°

Taula 16: Temperatures d'hibridació per les mutagènesis dels diferents polimorfismes

A continuació es va degradar la cadena parental metilada amb l'enzim DpnI (10 U) per digestió a 37 °C durant 2 hores i es va transformar directament en 100 µl de cèl·lules d'*E.Coli* TOP10 quimiocompetents.

Un cop extret el DNA plasmídic, les construccions es van seqüenciar per tal d'assegurar que les mutacions havien estat introduïdes i que no s'havia introduït cap altra mutació no desitjada.

A més, mitjançant digestions i lligacions, es van realitzar diferents combinacions d'al·lells, de les quals es van obtenir les següents construccions

- -2989A/ -401T/ -323ins10/ -122C
- -2989A/-670C/-630G/-402A
- -2989A/ -122C
- -2989A/ -670C
- -2989A/-323ins10

Les 15 construccions resultants (figura 16) es van usar per transfectar de manera transitòria la línia de cèl·lules hepàtiques tumorals HepG2.

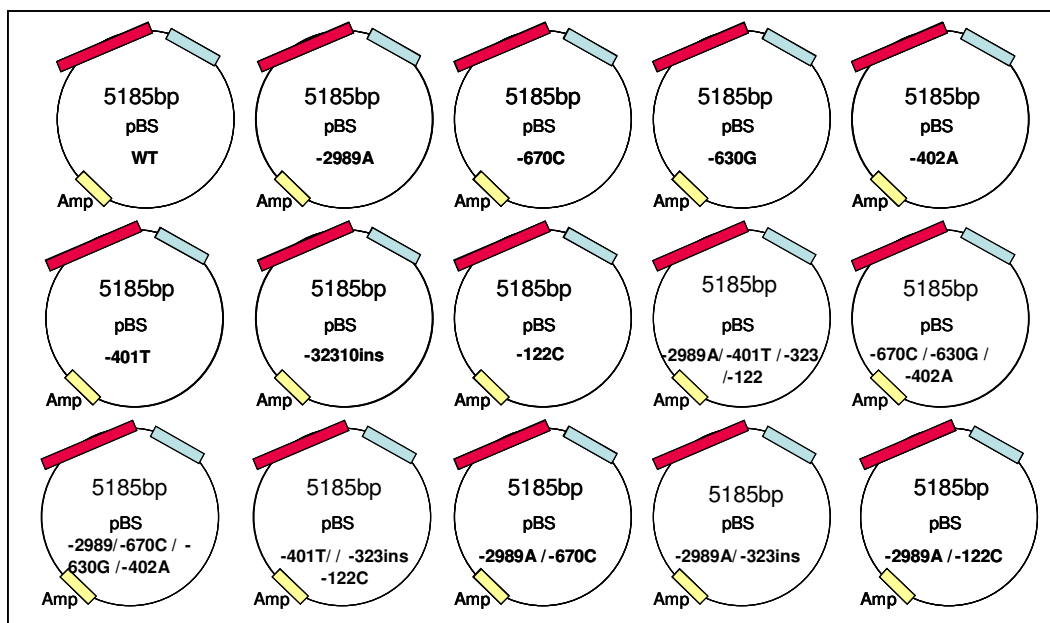


Figura 16: Esquema de les diferents construccions transfectades en les cèl·lules HepG2.

7.4- Transfecció

La transfecció és el procediment pel qual introduïm el DNA plasmídic en cèl·lules eucariotes.

Creixement i manteniment de les cèl·lules

Per la transfecció es van usar plaques de sis pous que contien cèl·lules hepàtiques humanes HepG2 en un 50-60% de confluència. Les cèl·lules es van fer créixer en medi DMEM amb 10% de sèrum fetal boví, 1% de Glutamina i 1% de Penicil·lina/Streptomicina (P/S).

Just abans de la transfecció, es van canviar els medis de les cèl·lules per DMEM +2% de sèrum fetal boví, 1% de Glutamina i P/S.

Cada construcció es va transfectar per triplicat.

La transfecció es va realitzar mitjançant la formació de complexos amb Polietilenimida de 25 KDa (PEI).

- 5 µg de cada construcció es van dissoldre en 200 µl de NaCl 150 mM
- s'hi van afegir 2,25 µl de PEI 10 mM per µg de DNA dissolt en 200 µl de NaCl 150 mM gota a gota sobre el DNA
- incubació durant 30 minuts a temperatura ambient
- es va afegir en cada un dels pous amb cèl·lules HepG2 en el medi amb 2% de sèrum fetal boví
- incubació durant 4 h a 37 °C, 5% CO₂
- aspiració del medi i es va afegir medi fresc amb 10% de sèrum fetal boví i 1% de Glutamina
- incubació durant 72 h a 37 °C, 5% CO₂

7.5- Quantificació de la fluorescència

Per tal de quantificar la fluorescència emesa per les cèl·lules, es va aspirar el medi i les cèl·lules es van desenganxar dels pous amb 1 ml de tampó salí. A continuació, es van fixar amb paraformaldehid al 2%.

La quantificació de les cèl·lules es va fer mitjançant un citòmetre de flux fluorescent (*Fluorescence Activated Cell Sorting, FACS*). El citòmetre aspira les cèl·lules de la mostra i les injecta d'una en una a través d'un raig làser que és capaç de contar el total de cèl·lules i distingir les que presenten fluorescència. Obtenim com a resultats el percentatge de cèl·lules fluorescents respecte les cèl·lules totals, i la intensitat de la fluorescència per cada experiment, dades que ens permeten comparar entre les diferents construccions.

7.6- Anàlisi estadística per comparar l'expressió de les diferents construccions

Per tal de normalitzar l'expressió de les diferents construccions dins de cada experiment, es va substreure el valor mitjà de les transfeccions de tres controls negatius (cèl·lules sense construcció) a cada mesura. A més, per comparar l'expressió entre experiments diferents, tots els experiments van ser normalitzats mitjançant l'expressió de la mitjana de la intensitat de fluorescència com a percentatge respecte la mitjana de la intensitat de fluorescència del WT de l'experiment en concret.

L'anàlisi de les diferències entre cada construcció es va realitzar mitjançant l'elaboració d'un model lineal mixt on es va considerar cada construcció com a variant fixa i l'efecte dels diferents experiments com a variant aleatòria. Totes aquestes anàlisis es van realitzar utilitzant el paquet estadístic R (The R Project for Statistical Computing, www.r-project.org).