



UNIVERSITAT DE
BARCELONA

New insights into the colorectal carcinogenesis: from early precursor lesions to the role of aneuploidy

Isabel Quintanilla Leo



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- SenseObraDerivada 3.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - SinObraDerivada 3.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NoDerivatives 3.0. Spain License.**



UNIVERSITAT^{DE}
BARCELONA

**New insights into the colorectal carcinogenesis: from early
precursor lesions to the role of aneuploidy**

Thesis presented by

Isabel Quintanilla Leo

to qualify for the degree of Doctor in Medicine

by the University of Barcelona

Under the supervision of:

Dr. Jordi Camps Polo

Dr. Maria Pellisé Urquiza

Barcelona, July 2017

A mis padres

A mis tíos

To Sam

"You're crazy and I'm out of my mind"

(All of me, John Legend)

AGRADECIMIENTOS

El momento que parecía tan lejano por fin ha llegado, dejándome más que nunca con la sensación de que el tiempo pasa volando. Aún me acuerdo de mi primer día en el lab OGiP, cuando Jenny me recibió con los brazos abiertos y me embarqué en la aventura de hacer el doctorado. Desde entonces muchas cosas han cambiado y tengo mucho que agradecer a toda la gente que ha estado a mi lado durante este largo camino.

Primero de todo, me gustaría agradecerles a mis directores, Maria y Jordi, que hayan hecho posible el desarrollo de esta tesis y hayan confiado en mí desde el primer momento. Durante este tiempo he crecido mucho científicamente hablando y no hubiera sido posible sin ellos. Además, también quería agradecer a Francesc y a Toni toda la ayuda que me han prestado durante estos años. Gracias Francesc por ser tantas veces mi “tercer director”, guiarme y aconsejarme siempre que lo he necesitado. Gracias Toni por hacer los miles de trámites relacionados con mi doctorado, que no han sido pocos, y contribuir a que llegara este momento.

También me gustaría dar las gracias a toda la gente que ha formado y forma parte del lab OGiP, y que tan a gusto me han hecho sentir durante esta etapa. Gracias Sergi, Txell, Eva, Lety, Laila, Claruchi, Claudia, Keyvan, Lorena, Elena, Irene y Maria. Y bueno, como no mencionar a los componentes del área “bioinformática” del grupo...muchísimas gracias Marcos por transmitir siempre esa energía positiva, gracias Sebas por escuchar mis preocupaciones tantas veces y darme siempre ánimos, y gracias María Vila por ser mi compañera de sufrimiento con nuestra queridísima FPU y por todos esos análisis estadísticos que tanto han aportado a mi tesis, eres una crack. Además, también tengo mucho que agradecer a Elena Asensio por toda su ayuda y horas en el microscopio, sin ti no hubiera conseguido acabar muchísimas cosas.

Y pasando al terreno tanto profesional como personal....llegan las que han sido las protagonistas de mi doctorado: Jenny, Esther y Saray. Muchas gracias Jenny por haberme acompañado durante todo este tiempo, me has enseñado muchísimas cosas y sin ti nada de esto hubiera sido posible. Gracias por toda tu ayuda, tu paciencia y tus consejos. Gracias por ser el alma del lab y apoyarme en todo, no tengo palabras para

agradecerte todo lo que me has aportado. Esther, gracias por ser la mami de todas, por cuidarme, apoyarme y enseñarme también tantísimas cosas. Me has ayudado a crecer dentro y fuera del lab, y para mí solo sumas en positivo. Y gracias Saray por todas esas conversaciones tan enriquecedoras, por entenderme tanto y siempre estar ahí cuando más lo he necesitado. Somos muy parecidas en muchos aspectos y creo que nos hemos ayudado a intentar superar nuestros dichosos miedos. Cree más en ti misma, que vales millones. Mis queridas mafiosas, aunque suene cursi lo peor de acabar esta etapa es que significa no veros cada día, aunque espero que sigáis formando parte de mi vida.

También me gustaría dar las gracias a Neus y Amaia por ayudarme en mi entrada en el mundo del “estrés replicativo”. Y gracias también a Miriam, que siempre ha estado dispuesta a sacarme una sonrisa y hacer unas birras.

I would also like to thank Thomas and Dara for their collaboration and all their help and kindness during all these years. They are an important part of this thesis too.

Moreover, I would also like to thank all the people who were part of my Virginia Tech experience. I’ve learnt from everything and my time there made me a better person and scientist. It was a once in a lifetime experience that I truly appreciate now.

Thank you Sam, thank you for being the best I’ve got from my time in Blacksburg. Somehow you are the soul of this thesis. Thank you for all your time teaching me proper English. Thank you for making me a better person and always being by my side. Thank you for being my life partner. I am looking forward to the new experiences and challenges with you.

Gracias también a mis amigos, los que han “sufrido” estos años conmigo: Samantha, Jose, Silvia, Yuste,....Y especialmente, gracias a Erica por estar a mi lado, entenderme y apoyarme tantísimo durante este proceso. Lo que unió Biotec, no lo separará nadie!

A mi familia. A mis padres. A mis tíos. Ellos son los que siempre han estado, están y estarán a mi lado. Gracias por todo vuestro apoyo y ánimos. Y especialmente, esta tesis va para ti. Mi interés por el mundo de la ciencia empezó por ti, y espero que estés donde estés, estés orgullosa de mí.

INDEX

ABBREVIATIONS	15
INTRODUCTION	21
1. <i>Colorectal cancer: an overview</i>	<i>23</i>
2. <i>Precursor lesions of colorectal cancer.....</i>	<i>25</i>
2.1. Aberrant crypt foci	26
2.2. Adenomas.....	28
2.2.1. Conventional adenomas	28
2.2.2. Serrated lesions.....	28
2.2.3. Malignant or progressed adenomas.....	29
3. <i>Molecular pathogenesis in colorectal cancer</i>	<i>30</i>
3.1. Key genes in CRC	31
3.1.1. Adenomatous Polyposis Coli.....	31
3.1.2. KRAS.....	32
3.1.3. BRAF	32
3.1.4. TP53.....	32
3.1.5. SMAD4.....	33
3.2. Epigenetics of CRC	33
3.2.1. DNA methylation in cancer	34
3.2.1.1. LINE-1 hypomethylation	37
3.2.1.2. MGMT hypermethylation	37
3.3. Abnormal karyotypes in cancer	37
4. <i>Aneuploidy and chromosomal instability.....</i>	<i>38</i>
4.1. Causes of numerical CIN	38
4.1.1. Spindle assembly checkpoint defects	38
4.1.2. Cohesion defects.....	39
4.1.3. Supernumerary centrosomes	39
4.1.4. Kinetochore-Microtubule attachments.....	41
4.1.5. Microtubule dynamics	42
4.1.6. Polyploidization.....	43
4.2. Causes of structural chromosome instability	45
4.3. Consequences of aneuploidy and chromosome missegregation	46
4.3.1. Immediate effects of chromosome missegregation	46
4.3.2. Cellular responses to aneuploidy.....	47
5. <i>Tumor heterogeneity.....</i>	<i>48</i>
6. <i>The role of replication in cancer.....</i>	<i>51</i>
6.1. Replication in normal conditions.....	51
6.2. Replication stress	51
6.3. Replication stress, genomic instability and cancer.....	52
HYPOTHESIS	55

OBJECTIVES	59
MATERIALS AND METHODS	63
1. <i>Patients</i>	65
1.1. ACF and methylation studies (Objective 1 and 2)	65
1.2. CRC evolution study (Objective 3)	65
1.3. Tetraploidy study (Objective 4)	66
2. <i>Endoscopic assessment of aberrant crypt foci</i>	66
3. <i>Pathological assessment of aberrant crypt foci</i>	67
4. <i>Preparation of cytopins from FFPE specimens</i>	67
5. <i>Cell culture</i>	69
5.1. Cell lines	69
5.2. Tetraploid cell lines generation	69
6. <i>Volume and area quantification</i>	69
7. <i>Nucleic acid extraction and quantification</i>	70
8. <i>Mutational analysis</i>	71
9. <i>Bisulfite pyrosequencing and methylation analysis</i>	72
10. <i>Microarrays analysis</i>	73
10.1. Array comparative genomic hybridization	73
10.2. Gene expression	73
11. <i>Real-Time PCR</i>	74
12. <i>Protein analysis</i>	75
12.1. Western blot	75
12.2. Immunostaining	76
12.3. Microscopy analysis	76
13. <i>Cytogenetic procedures</i>	77
13.1. Chromosome harvesting and cytokinesis-block micronucleus assay	77
13.2. Fluorescence in situ hybridization and spectral karyotyping	77
13.3. FISH data analysis	78
13.4. Modeling of tumor progression and analysis of node depth	79
14. <i>Flow cytometry</i>	81
15. <i>DNA ploidy measurement</i>	82
16. <i>Viability, growth and colony formation assays</i>	82
17. <i>Migration and invasion assays</i>	83
18. <i>Live cell imaging</i>	83
19. <i>Statistical analysis</i>	84
RESULTS	85
1. <i>Chapter I: Rectal aberrant crypt foci in humans are not surrogate markers for colon cancer risk</i>	87
1.1. Patient baseline characteristics	89
1.2. Endoscopic features of ACF as a surrogate marker for CRC risk	90
1.3. ACF histology as a surrogate marker for CRC risk	92

1.4.	Molecular analysis of ACF.....	93
1.4.1.	APC mutation and microsatellite instability analysis	93
1.4.2.	BRAF and KRAS mutations	95
1.4.3.	MGMT methylation status analysis	95
2.	<i>Chapter II: LINE-1 hypomethylation is neither present in rectal aberrant crypt foci nor associated with field defect in sporadic colorectal neoplasia</i>	<i>97</i>
2.1.	Patients baseline characteristics	99
2.2.	Rectal ACF do not display LINE-1 hypomethylation	101
2.3.	LINE-1 methylation levels in normal colorectal mucosa are similar in all CRC risk groups	104
2.4.	LINE-1 methylation levels in normal mucosa differ between colonic segments	106
3.	<i>Chapter III: The role of copy number alterations in the adenoma-to-adenocarcinoma transition of colorectal cancer</i>	<i>107</i>
3.1.	Sample baseline characteristics	109
3.2.	Chromosomal instability and clonal patterns.....	110
3.3.	Cells from ADK exhibit a greater potential for progression than cells from AD	115
3.4.	Consensus evolutionary trees reveal common ancestor clones and drivers of malignant transformation	116
3.5.	Whole genome duplication events play a main role in CRC progression	125
4.	<i>Chapter IV: Near-tetraploid cancer cells show chromosome instability triggered by replication stress and exhibit enhanced invasiveness</i>	<i>129</i>
4.1.	Generation and characterization of diploid and near-tetraploid isogenic clones	131
4.2.	Transcriptional profiling of 4N cells reveals a deregulation of genes associated with DNA replication and mitosis	135
4.3.	Near-tetraploid cells display increased replication stress and DNA damage	137
4.4.	Near-tetraploid cells show karyotypic heterogeneity and genomic instability	143
4.5.	Replication stress triggers genomic instability	150
4.6.	Near-tetraploid cells exhibit enhanced migratory and invasive capabilities in primary tumors.....	151
	DISCUSSION	157
1.	<i>Aberrant crypt foci and the need for colorectal cancer biomarkers.....</i>	<i>159</i>
1.1.	The importance of endoscopic and histological assessments of ACF.....	160
1.2.	Genetic and epigenetic findings in ACF	161
1.2.1.	KRAS mutations are commonly found in ACF but are not predictive of CRC	161

1.2.2. Enhanced <i>MGMT</i> hypermethylation is associated with <i>KRAS</i> mutations in ACF.....	162
1.2.3. LINE-1 hypomethylation does not occur in ACF and does not depict a field defect for CRC	162
2. The role of chromosome copy number alterations in the adenoma-to-adenocarcinoma sequence.....	163
2.1. Copy number alterations result in heterogeneity during CRC evolution ..	163
2.2. Specific patterns of chromosomal alterations are selected during CRC ..	164
2.3. Depicting the phylogenetic evolution of CRC.....	165
3. Tetraploidy and cancer.....	166
3.1. The role of p53 in the generation of tetraploidy.....	166
3.2. Chromosomal instability is a common feature of near-tetraploid cells ...	167
3.3. Near-tetraploid cells experience replication stress.....	168
3.4. Replication stress leads to both structural and numerical chromosome abnormalities.....	169
3.5. The benefit of tetraploidy on cancer cell behavior	169
CONCLUSIONS	171
BIBLIOGRAPHY	175

ABBREVIATIONS

2N: diploid

4N: near-tetraploid

ACF: aberrant crypt foci

aCGH: array comparative genomic hybridization

AD: adenoma region of progressed adenoma

ADK: adenocarcinoma region of progressed adenoma

APC: adenomatous polyposis coli

APC/C: anaphase promoting complex

ATM: ataxia-telangiectasia mutated

ATR: ataxia-telangiectasia mutated and Rad3 related

BIR: break-induced replication

CC: colon cancer

CDK2: cyclin dependent kinase 2

CE: chromoendoscopy

CIMP: CpG island methylation phenotype

CIN: chromosomal instability

CFSs: common fragile sites

CMS: consensus molecular subtypes

CNAs: copy number alterations

CRC: colorectal cancer

DDR: DNA damage response

DNMTs: DNA methyltransferases

DSB: double-strand breaks

ERFS: early-replicating fragile sites

FAP: familial adenomatous polyposis

FFPE: paraffin-embedded specimens

FISH: fluorescence *in situ* hybridization

H&E: hematoxylin and eosin

HDAC: histone deacetylases

HMPS: hereditary mixed polyposis syndromes

HNPCC: hereditary non-polyposis syndromes

KT: kinetochore

LINEs: long interspersed nuclear elements

LOH: loss of heterozygosity

LS: Lynch syndrome

MAPK: MAP kinase

MCM2-7: mini-chromosome maintenance complex 2-7

MGMT: O6-methylguanine-DNA methyltransferase

MMR: mismatch repair

MNi: micronuclei

MSI: microsatellite instability

MT: microtubule

N_AD: non-advanced adenoma

NBI: Narrow Band Imaging

NDM: normal descending mucosa

NRM: normal rectal mucosa

NTP: nucleotides

RNR: ribonuclease reductase

RS: replication stress

SAC: spindle assembly checkpoint

SCNA: somatic copy number alteration

SKY: spectral karyotyping

SSA/P: sessile serrated adenoma/polyp

ssDNA: single-stranded DNA

TSA: traditional serrated adenoma

VPI: vascular pattern intensity

INTRODUCTION

1. Colorectal cancer: an overview

Colorectal cancer (CRC) is the fourth most common cancer and the third leading cause of death in the world (*International Agency for Research on Cancer*, <http://www.globocan.iarc.fr>) (**Figure 1**). The incidence of CRC strongly increases with age, with 70 years old being the median age at the time of diagnosis¹. Although the highest incidence is reported in developed countries, important increases in incidence have been detected in previously low-risk countries due to the spread of the so-called western life-style². Conversely, the incidence in developed countries has stabilized or started to decrease likely due to screening programs, the colonoscopic polypectomy and the improved treatments²⁻⁴.

The prognosis of patients with CRC strongly depends on the stage at the time of diagnosis¹, but it has improved over the past several decades, reaching an average 5-year relative survival of 65% (ranging from 8 to 92% depending on the tumor stage)⁵.

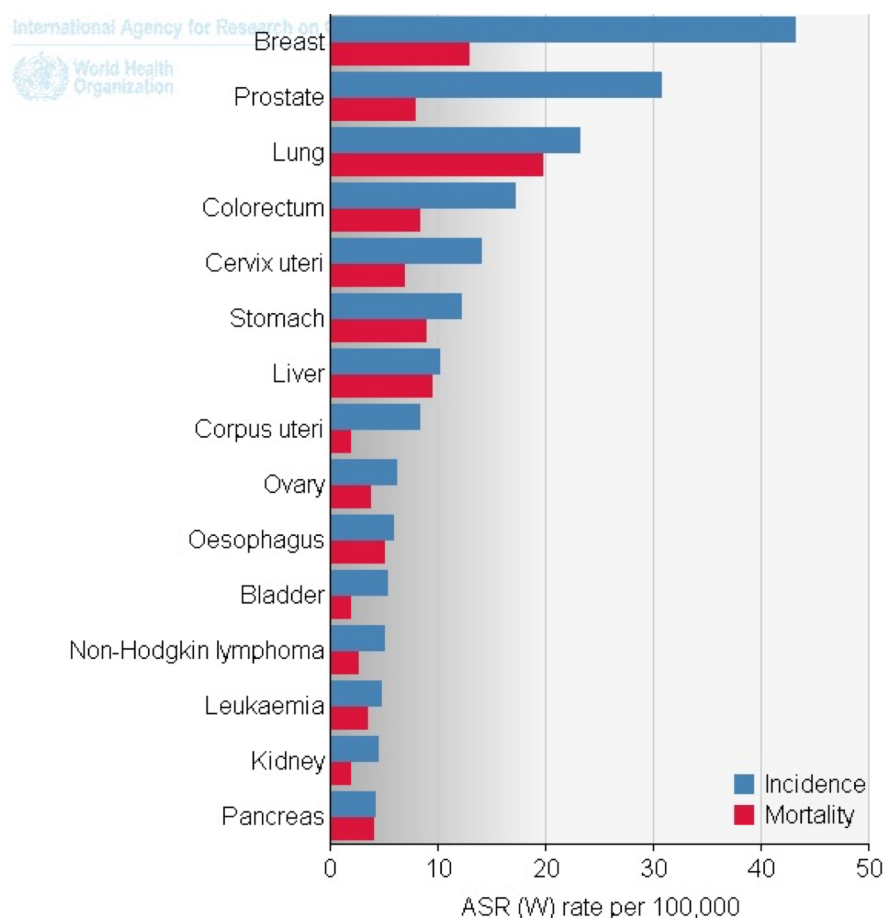


Figure 1. Estimated age-standardised incidence and mortality rates of the most common cancer types in the world. Extracted from Globocan 2012. *International Agency for Research on Cancer* (www.globocan.iarc.fr). ASR (W), age-standardized rate (weighed).

Several risk factors are responsible for most cases of CRC: family history⁶, inflammatory bowel disease⁷, smoking⁸, excessive alcohol consumption⁹, high consumption of red and processed meat¹⁰, obesity¹¹, and diabetes¹². Furthermore, several preventive factors have been described, such as physical activity and aspirin; however, endoscopic removal of precancerous lesions provides the strongest risk reduction¹³.

The vast majority (70%) of CRC patients exhibit sporadic disease without evidence of any inherited disorder, whereas 25% of patients have familial CRC, and about 5% of patients exhibit hereditary CRC without familial aggregation^{14, 15}. These hereditary forms of CRC can be subclassified into two main groups: the hereditary mixed polyposis syndromes (HMPS) and the non-polyposis syndromes (HNPCC). The former are the less common and basically include the hamartomatous polyposis syndromes, which are the Juvenile, Cowden and Peutz-Jeghers syndromes. On the other hand, HNPCC mainly include the familial adenomatous polyposis (FAP), caused by mutations in *APC* or *MYH*, and the Lynch Syndrome (LS), caused by mutations in the DNA mismatch repair (MMR) genes (**Figure 2**). Moreover, the polyposis associated with mutations in the *MUTYH* gene is also included in HNPCC, although its incidence is rather low in comparison with FAP and LS.

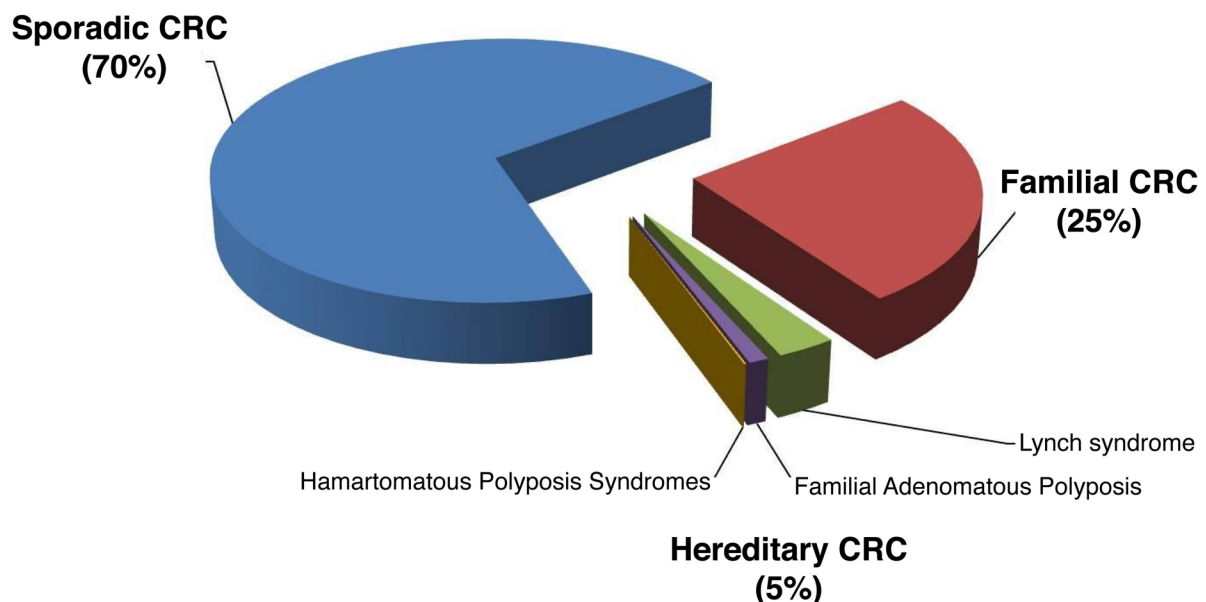


Figure 2. Classification and frequency of the main CRC types depending on family risk. Adapted from Burt, R.W., *Gastroenterology* 2000.

2. Precursor lesions of colorectal cancer

CRC develops slowly, in some cases for over more than 10 years, with dysplastic adenomas being the most common type of premalignant lesions¹⁶. In an effort to characterize the stepwise progression of CRC, Vogelstein and Fearon proposed the adenoma-carcinoma sequence as a genetic model of colorectal tumorigenesis, beginning from the transformation of normal colorectal epithelium into adenoma and eventually leading to the invasive and metastatic tumor stage¹⁷. As such, the commonly accepted paradigm is that specific genetic alterations accumulate in a stepwise manner with the initial genetic changes appearing in the early adenoma stage and progressively increase throughout the transformation into carcinoma (**Figure 3**).

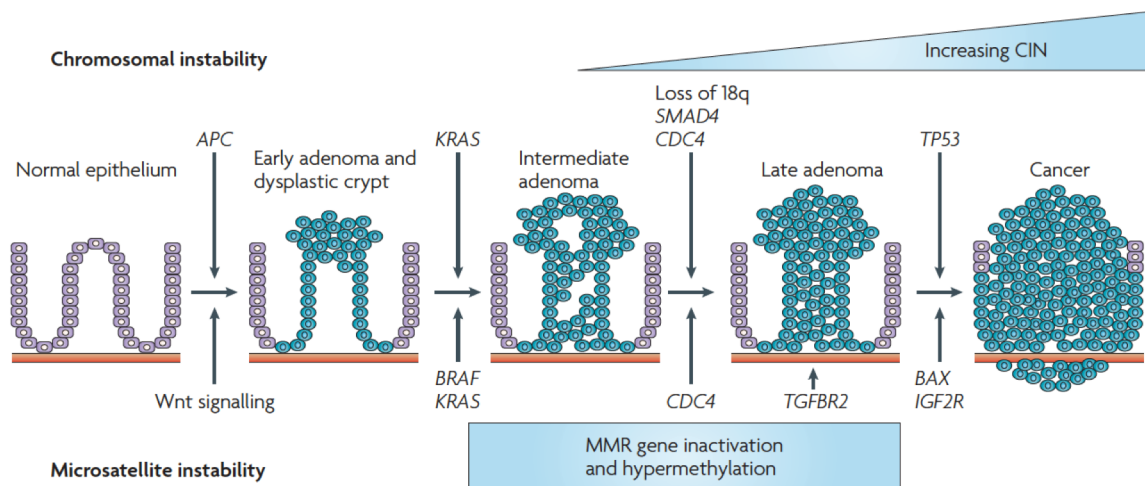


Figure 3. Genetic model for the adenoma-carcinoma sequence. *Extracted from Walther, A. et al., Nature Reviews 2009.* MMR, mismatch repair; and CIN, chromosomal instability.

Moreover, aberrant crypt foci (ACF) have been proposed as the earliest precursor lesions for CRC, appearing even before the adenoma formation (**Figure 4**). To this day, the adenoma-carcinoma sequence still represents the basis of our fundamental understanding of this disease; however, it is considered to be an oversimplification of a much more complex process.

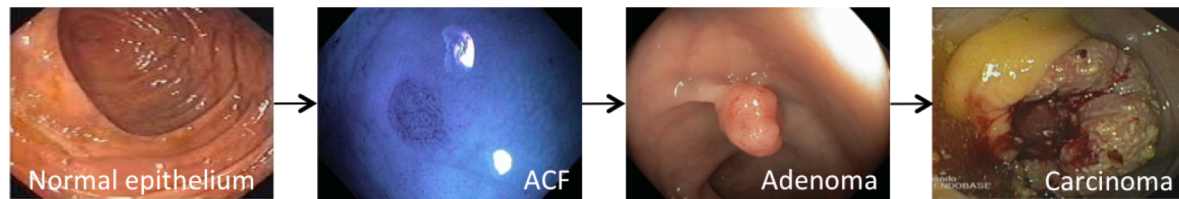


Figure 4. Stepwise progression from normal epithelium to invasive colorectal carcinoma. ACF, aberrant crypt foci.

2.1. Aberrant crypt foci

ACF are the earliest visible lesions in the colorectum and potential precursors of adenomas and CRC. They are defined as clusters of colonic crypts that are microscopically raised above the normal mucosa and exhibit thicker epithelia, altered luminal openings and a larger size compared to the normal adjacent crypts¹⁸. ACF were first detected in the colon of rodents treated with colon-specific carcinogens¹⁹ and later were identified in human patients at high risk for CRC²⁰. Moreover, although these lesions can arise in both the proximal and distal colon, they are mostly observed in the distal colon and rectum²¹. ACF are identified by either magnification chromoendoscopy or high-resolution chromoendoscopy^{22, 23} and are mainly used as a biological end-point for testing the efficacy of chemopreventive agents in animals²⁴.

In humans, several studies observed that the number and size of ACF were related to CRC risk²⁵⁻²⁸, although others argued against these findings²⁹⁻³². In addition, dysplasia, monoclonality and some of the genomic and epigenomic alterations present in adenomas and carcinomas have been identified in ACF^{33, 34}. Nevertheless, the frequency of alterations detected in ACF varies substantially between studies, being *KRAS* mutation the first and only alteration consistently detected in these lesions³⁵⁻³⁸ (**Table 1**).

From a histological perspective, ACF can be classified as hyperplastic or dysplastic, according to the World Health Organization³⁹. The main differential characteristics are shown in **Table 2**. The frequency of dysplastic ACF is low; however, they initially seemed to exhibit more potential of degeneration because they have been frequently described in patients at higher risk for CRC and carry more genomic alterations^{23, 40}. Nonetheless, other

studies suggest that hyperplastic ACF might also have malignant potential though the serrated type neoplasia⁴¹. Conversely, some authors reported that most ACF might remain dormant or regress and disappear. In fact, rectal ACF are difficult to re-identify after an initial endoscopy^{30, 42-44}.

Table 1. Rectal ACF genomic and epigenomic alterations and prevalence rates. Reviewed in Gupta, K. et al., *Clin Gastroenterol Hepatol* 2007.

Aberration	Prevalence
<i>KRAS</i> mutations	13-82%
<i>APC</i> mutations	0-5%
<i>p53</i> mutations	None
<i>BRAF</i> mutations	2-3%
β -catenin gene mutations	None
Microsatellite instability	0-11%
Chromosomal instability	16-23%
<i>MGMT</i> hypermethylation	12-53%

Table 2. Differential characteristics of hyperplastic and dysplastic ACF. Adapted from Renehan, A.G. et al., *Colorectal Dis* 2002.

	Hyperplastic ACF	Dysplastic ACF
Frequency	Most common	Uncommon
Clonality	Present	Present
<i>KRAS</i> gene mutation	Present	Absent
<i>APC</i> gene mutation	Absent	Present
Architecture	Larger and longer crypts, with apical branching and serration	Smaller crypts without serration
Nuclei	Enlarged, round or oval, not stratified	Larger than hyperplastic, elongated and polymorphic
Mucin depletion	Partially present	More prominent
Proliferation	Expansion of proliferative crypt compartment	Expansion to upper and middle crypt compartments
Maturation	Disordered, upper crypt not involved	Disordered, upper crypt involved

Collectively, the data reported about ACF is controversial and varies between studies. Accordingly, their natural history and role in the adenoma-carcinoma sequence are not fully understood.

2.2. Adenomas

2.2.1. Conventional adenomas

Conventional adenomas are the most common neoplastic lesions in the large intestine⁴⁵, and are defined by the presence of dysplastic epithelium⁴⁶. Most adenomas are polypoid or pedunculated, although a small number are flat or sessile⁴⁷. Moreover, adenomas are recognized as precursor lesions of CRC in the traditional colorectal carcinogenesis and they normally exhibit *APC* mutations, thought to be the earliest genetic event in colorectal tumorigenesis⁴⁸ (**Figure 3**). In fact, about 95% of adenocarcinomas arise from adenomas^{43, 49-51} and polypectomy has been demonstrated to reduce the risk of CRC development⁵²⁻⁵⁴.

Interestingly, although the incidence of adenomas is 40-60%, merely 5% of the population with adenomas develops CRC, indicating that only a small proportion of adenomas progress to carcinoma⁵⁵. Accordingly, some characteristics of adenomas have been associated with cancer development, including: size, villous architecture, as well as the degree of dysplasia^{54, 56, 57}.

2.2.2. Serrated lesions

On the other hand, approximately one third of colorectal cancers develop through the recently described serrated pathway of carcinogenesis⁵⁸. For this carcinogenesis pathway, the precursor lesions are the sessile serrated adenoma/polyp (SSA/P) and the traditional serrated adenoma (TSA), which preferentially arise in the proximal colon^{45, 59}. These lesions are characterized by a serrated architecture of the epithelial component and, in contrast with conventional adenomas; they initially exhibit *BRAF* or *KRAS* mutations and CpG island methylation phenotype (CIMP) (see next section)^{60, 61}.

2.2.3. Malignant or progressed adenomas

Malignant adenomas are an early form of CRC defined as adenomas with a focus of carcinoma that has invaded by direct continuity through the muscularis mucosa into the submucosa (**Figure 5**)⁶². The incidence of these lesions is reported to be approximately from 0.2% to 2% of all endoscopically removed colonic polyps^{63, 64}, and their management depends entirely on their risk of lymph node metastasis⁶⁵. In fact, several histological features are associated with lymph node metastasis and local cancer recurrence, including poorly differentiated histological features, vascular or lymphatic invasion, depth of tumor invasion into the submucosa, presence of tumor budding, and incomplete resection^{66, 67}. Endoscopic resection alone is adequate for the treatment of low-risk malignant adenomas, whereas subsequent oncological surgical resection should be performed for high-risk lesions^{68, 69}.

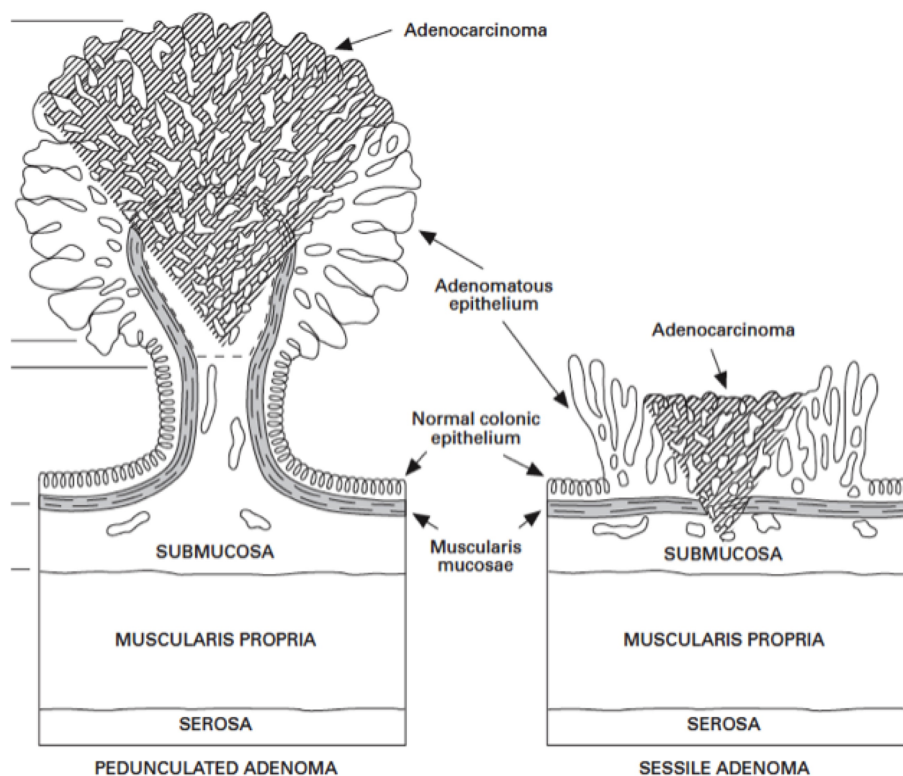


Figure 5. Representation of progressed adenomas, both pedunculated and sessile forms. *Extracted from Nivatvongs, S. et al., Surg Clin North Am, 2000.*

3. Molecular pathogenesis in colorectal cancer

As stated before, in the 1990s, a genetic model for colorectal carcinogenesis was proposed in which genome alterations occur in a sequence that parallels the clinical progression of the tumor (**Figure 3**)⁷⁰. Accordingly, three molecular carcinogenesis pathways have been described for CRC⁷¹.

The first and most commonly identified in sporadic CRC is the **chromosomal instability (CIN) pathway**, accounting for approximately 65% of the cases⁷². It is characterized by large genetic aberrations, such as alterations in the number and structure of chromosomes, termed aneuploidy, as well as genetic mutations of oncogenes and tumor suppressor genes, such as *APC* mutations. The second pathway is the **microsatellite instability (MSI) pathway**, representing approximately 15% of sporadic CRC cases. Microsatellites are DNA tandem repeats distributed throughout the genome that are particularly prone to errors during DNA replication. Such errors are usually repaired by the MMR system, composed by proteins coded by the genes *MSH2*, *MLH1*, *PMS1*, *PMS2* and *MSH6*⁷³. Inactivation of the MMR system causes mutations not only in the microsatellites but also in gene coding regions⁷⁴, leading to hypermutated tumor phenotypes. The third pathway is the **CpG island methylation (CIMP) or epigenetic instability pathway** and, as previously mentioned, it is preferentially associated with serrated neoplasias. This pathway is characterized by the hypermethylation of various CpG island loci, which consequently alters the expression of tumor suppressor genes or tumor-related genes.

More recently, however, there has been consensus that it is the accumulation of alterations rather than their specific order that is most important in determining the progression and thus prognosis of CRC. In this sense, a recent study identified and described **the consensus molecular subtypes (CMS)** for CRC. This novel taxonomy integrates different key biological features of CRC, such as mutations, gene expression, copy number alterations (CNAs), methylation, microRNA expression, proteomics, as well as patient outcome in order to accurately classify CRC⁷⁵. Interestingly, CMS have shown that there is an overlap between the originally described molecular pathways and revealed that tumors harboring the same driver events for CRC vary substantially in their biology, highlighting the poor genotype-phenotype correlations for this disease. In fact, none of the described CRC groups were

defined by an individual event, and no genetic aberration was specific for a specific subtype. Conversely, it was gene expression and proteomics analyses what offered valuable insight into the biological understanding of CRC. Consequently, the proposed classification included four distinct groups: CMS1, which exhibits MSI and immune phenotype; CMS2, the epithelial subtype with activation of Wnt pathway and *MYC* oncogene; CMS3, which displays metabolic dysregulation; and the CMS4, which represents the mesenchymal phenotype. The detailed characteristics and frequency of each CMS are described in **Table 3**.

Table 3. Characteristics of the different consensus molecular subtypes for CRC. *Extracted from Guinney, J. et al., Nature Medicine 2015.*

CMS1 MSI immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermutation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
<i>BRAF</i> mutations		<i>KRAS</i> mutations	
Immune infiltration	Wnt and <i>MYC</i> activation	Metabolic deregulation	Stromal infiltration, TGF- β activation, angiogenesis
Worse survival			Worse relapse-free and overall survival

CIMP, CpG island methylator phenotype; MSI, microsatellite instability; and SCNA, somatic copy number alterations.

3.1. Key genes in CRC

3.1.1. Adenomatous Polyposis Coli

Adenomatous polyposis coli (*APC*) is a tumor suppressor gene located on chromosome 5q22.2⁷⁶ that plays a major role in the Wnt signaling pathway. *APC* mutations are mainly found in the exon 15 of the gene and are known to occur early in the adenoma-carcinoma sequence. In fact, their frequency accounts for approximately 30 and 70% of conventional sporadic adenomas and CRC, respectively⁷⁷. *APC* protein normally blocks the progression from G1 to S phase in the cell cycle by interacting with several proteins, including β -catenin. Specifically, *APC* induces the degradation of β -catenin via the ubiquitin-proteasome pathway, therefore acting as a negative regulator of the Wnt signaling pathway⁷⁸. Accordingly, mutated *APC* leads β -catenin to accumulate, increasing the

transcriptional activity of different downstream genes⁷⁹, including important oncogenes such as *MYC* and *CCND1*⁸⁰.

3.1.2. *KRAS*

KRAS is an oncogene from the *RAS* gene family located on chromosome 12q12.1. It is a member of the MAP kinase (MAPK) pathway, a major cell proliferation signal transduction pathway. *KRAS* mutations are mainly found in codons 12, 13 and 61. Moreover, such mutations are present in approximately one-third of sporadic CRC and are considered an early-event in the adenoma-carcinoma sequence⁸¹. The activation of the MAPK cascade is dependent on the binding of growth factors to the epidermal growth factor receptor (EGFR); however, *KRAS* mutations lead to constitutive activation of the cascade and promote cell proliferation and survival independent of the EGFR receptor⁸². In fact, *KRAS* mutation status is a predictive marker for the treatment with EGFR inhibitors⁸³.

3.1.3. *BRAF*

BRAF is an oncogene from the *RAF* gene family located on chromosome 7q34, which also mediates cellular responses to growth signals through the MAPK pathway. In fact, *BRAF* protein acts downstream of *KRAS* in the MAPK cascade and mutations in these genes seem to be mutually exclusive but have similar consequences; the constitutive activation of the cascade. Mutations in *BRAF* have been found in nearly 10% of sporadic CRC and are associated with the CIMP pathway and serrated lesions. Furthermore, the vast majority of these mutations is represented by the V600E hotspot⁸² and are also determinants of resistance to EGFR-specific therapies⁸³.

3.1.4. *TP53*

TP53 is located on chromosome 17p13.1 and is the tumor suppressor gene most frequently altered in human cancers^{84, 85}. In CRC, *TP53* mutations occur relatively late, being found in 50-70% of CRC but rarely in the early adenomas⁷⁰. The p53 protein induces G1 cell cycle arrest when there is DNA damage and triggers DNA repair prior to replication. Nevertheless, if damaged DNA repair is unsuccessful, p53 induces apoptosis. Alterations in this gene deregulate cell proliferation and apoptotic mechanisms, leading to genomic

instability and tumor progression⁸⁶. Missense mutations in the DNA-binding domain of *TP53* are the most common source of its inactivation; however, allelic loss of the chromosome 17p due to aneuploidy also causes the loss of the protein⁸⁷.

3.1.5. *SMAD4*

The tumor suppressor gene *SMAD4* is a member of the transforming growth factor- β (TGF β) signaling pathway, which regulates cell growth, differentiation and apoptosis^{88, 89}. *SMAD4* is located on chromosome 18q21.2, the loss of which is the most common cytogenetic abnormality in CRC⁹⁰. Thus, conversely to what happen to other genes implicated in colorectal carcinogenesis, the *SMAD4* alteration is mainly due to aneuploidy instead of an individual gene mutation. Interestingly, the loss of 18q is a marker of CIN⁹¹ and is associated with CRC poor prognosis⁹².

3.2. Epigenetics of CRC

Epigenetics are heritable changes in gene expression that do not arise as a consequence of alterations in the DNA sequence⁹³. The best-known epigenetic marker is DNA methylation, the covalent attachment of a methyl-group to the 5' position of cytosine residues in cytosine-guanine (CG) dinucleotides. Nevertheless, the chromatin structure and histone modifications also contribute to transcription regulation⁹⁴. The CG dinucleotides susceptible to methylation are not randomly distributed throughout the genome; instead they are clustered in regions known as CpG islands, which span the 5' end of up to 70% of all gene promoters⁹⁵. Additionally, methylation can occur in non-promoter regions, the so-called CpG shores, which are located approximately 2kbp away from the CpG islands and are also considered to be regulatory regions of gene activity^{96, 97}.

DNA methyltransferases (DNMTs) are enzymes that regulate the methylome by catalyzing the addition of the methyl-groups to the cytosine residues. Three DNMTs have been described in mammalian cells: DNMT1, DNMT3a and DNMT3b⁹⁸. DNMT1 is responsible for the maintenance of methylation by incorporating methyl groups into newly replicated DNA, whereas DNMT3 proteins are responsible for *de novo* methylation of CpG islands and have been related to aberrant methylation phenotypes observed in CRC⁹⁹. Moreover, as the state of histone acetylation is important for chromatin structure and gene transcription,

DNMTs are also important for the recruitment of histone deacetylases (HDACs) and other chromatin-binding proteins. Accordingly, unmethylated CpG islands, in which the histones are acetylated and the nucleosomes are in a more open configuration, allow access for transcription factors and gene expression. Conversely, when CpG islands are hypermethylated the histones are maintained deacetylated by DNMTs and the nucleosomes become tightly compacted, as in heterochromatin, thus inhibiting transcription.

There are different approaches for determining DNA methylation; however, the majority of them use the bisulfite conversion of DNA and subsequent PCR-based method as it allows for a more rapid determination of DNA methylation-status and requires little material¹⁰⁰. Specifically, the bisulfite treatment converts cytosine residues to uracil, leaving the 5-methylcytosine residues unaffected, and thus allowing for the easy identification of methylation-status by sequencing.

3.2.1. DNA methylation in cancer

DNA methylation maintains tissue-specific expression patterns, defines imprinting and regulates genome integrity. Nevertheless, neoplastic transformation changes the normal epigenomic landscape. In the early 1980s, Feinberg and Volgestein first associated changes in the DNA methylation status to cancer¹⁰¹. Since then, many cancers have been found to exhibit aberrant epigenomes, including CRC. Such disruption of methylation can be classified as either global hypomethylation or localized hypermethylation.

The **global hypomethylation** of tumors was one of the first epigenetic alterations detected in human cancers¹⁰¹, and is caused by the loss of methylation in repetitive DNA sequences as well as in coding regions and introns¹⁰². The decreased levels of global methylation are related to colorectal carcinogenesis and associated with the CIN phenotype as it favors mitotic recombination causing chromosome deletions and translocations as well as chromosome rearrangements^{103, 104}. Also, transposable and repetitive DNA sequences such as long interspersed nuclear elements (LINEs) and Alu elements are reactivated by global hypomethylation, which in turn triggers genomic instability¹⁰⁵. Interestingly, the degree of DNA hypomethylation increases with the neoplastic progression¹⁰⁶.

On the other hand, the **hypermethylation** of tumor suppressor and other cancer-related genes is also a major event in cancer. Many of these silenced genes are involved in cell cycle, DNA repair, metabolism of carcinogens, cell-to-cell interaction, apoptosis and angiogenesis. Unlike hypomethylation, the profiles of hypermethylation are specific to the type of cancer¹⁰⁷. For CRC, patterns of aberrant methylation have been identified throughout the adenoma-carcinoma sequence, even in preneoplastic lesions such as ACF¹⁰⁸.

Table 4 shows a list of aberrantly methylated genes at different stages of CRC development. CIMP was defined for those colorectal tumors that exhibit a high methylation frequency at three or more specific gene loci such as *CDKN2A*, *MINT1*, *MINT2*, *MINT31* and *MLH1*¹⁰⁹. Interestingly, CIMP was found to be associated with *BRAF* mutations and MSI since one of the markers for methylation is the MMR gene *MLH1*.

Table 4. Aberrantly methylated genes in CRC progression. Reviewed in Carmona, F.J. et al., Mut Res 2010.

Histological grade	Gene	Role in colorectal tumorigenesis
Precursor lesion	<i>CDH13</i>	Aberrant methylation of <i>CDH13</i> was reported in ACF, and significant hypermethylation is consistently observed in adenomas. It is a regulator of the dynamic cellular adhesion–deadhesion processes, and its inactivation through hypermethylation contributes to the dissemination of cancer cells.
	<i>DAPK</i>	<i>DAPK</i> hypermethylation was previously observed in the colon mucosa adjacent to intraepithelial neoplasia or carcinoma. Its silencing contributes to the early steps of tumor progression in the colorectal carcinoma through the inactivation of gamma-interferon induced programmed cell death.
	<i>IGF2</i>	Loss of imprinting at this locus is a risk factor for CRC predisposition. It enhances the expression of VEGF and affects cell proliferation.
	<i>MGMT</i>	Involved in repairing DNA damage caused by alkylating agents. Cancer-associated hypermethylation contributes to colorectal field cancerization, and correlates with G-to-A mutation in the <i>KRAS</i> oncogene. Hypermethylation-induced inactivation is observed in histologically normal-appearing mucosa of cancer patients.
	<i>WIF1</i>	Tumor suppressor gene whose inactivation contributes to the aberrant activation of Wnt signaling program. It has been proposed as a valuable biomarker in the plasma for early detection of CRC.
	<i>HOXD1</i>	Sequence-specific transcription factor that is involved in differentiation. Its inactivation occurs in premalignant lesions and accumulates during carcinogenesis.

Table 4. Aberrantly methylated genes in CRC progression. Reviewed in Carmona, F.J. et al., Mut Res 2010 (Continued).

Histological grade	Gene	Role in colorectal tumorigenesis
Precursor lesion	<i>SFRP1</i>	SFRP family genes are soluble modulators involved in Wnt signaling. Epigenetic silencing of the <i>SFRP</i> genes occurs in precursor lesions leading to aberrant activation of the Wnt pathway.
	<i>SLC5A8</i>	Hypermethylation of this sodium transporter was common in colon adenomas, and could be also detected in microscopic colonic ACF.
Adenoma	<i>APC</i>	Tumor suppressor gene, antagonist of the Wnt signaling pathway. Involved in cell migration and adhesion, transcriptional activation, and apoptosis.
	<i>COX-2</i>	Involved in inflammation and mitogenesis, tumor angiogenesis acting through VEGF, and metastasis.
	<i>Cyclin A1</i>	Binds to important cell cycle regulators, such as Rb family proteins, transcription factor E2F-1, and the p21 family proteins.
	<i>GATA4</i>	Potential CRC biomarker found in stool. Suppresses proliferation, migration, invasion, and anchorage-independent growth of CRC cells.
	<i>HIC-1</i>	Modulates the transcriptional stimulation of genes regulated by Wnt/ β -catenin signaling. Inactivation in cancer leads to abnormal cell proliferation at early stages.
	<i>hMLH1</i>	Implicated in DNA repair, its inactivation causes microsatellite-unstable tumors, mainly in sporadic cases. A higher frequency of aberrant hypermethylation is observed in patients with a family history of CRC.
	<i>p16INK4a</i>	Tumor suppressor gene, negative regulator of cell growth and proliferation in the G1 phase of the cell cycle, and is an important regulator of the angiogenic switch.
	<i>RARB2</i>	Tumor suppressor gene that mediates the growth-inhibitory action of retinoic acid. Its silencing promotes cell growth and tumor progression.
Adenocarcinoma	<i>DKK-1</i>	Extracellular Wnt pathway inhibitor hypermethylated in advanced colorectal neoplasms.
	<i>HLTF</i>	Encodes a chromatin remodeling factor with growth-suppressive effect in cancer cells.
	<i>miR-34b/c</i>	Tumor suppressor microRNA that down-regulates oncogenic target genes such as <i>C-MYC</i> , <i>E2F3</i> and <i>CDK6</i> . Inactivation of this microRNA correlates with lymph node metastasis.
	<i>miR-148a</i>	Tumor suppressor microRNA that targets <i>TGIF2</i> oncogene. Its inactivation correlates with lymph node metastasis.

3.2.1.1. *LINE-1 hypomethylation*

Long interspersed nuclear element-1 (LINE-1) are transposable elements that comprise about 18% of the human genome and are considered to be a surrogate marker for genome-wide DNA methylation levels¹¹⁰. While in normal cells LINE-1 is heavily methylated¹¹¹, cancer cells display LINE-1 hypomethylation, which leads to genome instability¹¹². This phenomenon occurs early in colorectal carcinogenesis, increases with disease progression and can be even detected in the serum of CRC patients^{113, 114}. In fact, the levels of LINE-1 methylation have been associated with increased mortality in CRC patients¹¹⁵. Interestingly, recent studies have reported that the hypomethylation of these elements also constitutes a field-defect for CRC, suggesting that LINE-1 methylation status in normal mucosa could be used to stratify CRC risk and tailor preventive strategies¹¹⁶.

3.2.1.2. *MGMT hypermethylation*

The O6-methylguanine-DNA methyltransferase (*MGMT*) is a DNA repair gene involved in the removal of the mutagenic methyl group from O-6-alkylated guanines¹¹⁷. The *MGMT* promoter is frequently methylated in CRC as result of CIMP, which leads to a loss/reduction in protein function and an increase in genome instability. Furthermore, the hypermethylation of the *MGMT* promoter occurs during the early stages of CRC, such as ACF, it is associated with *KRAS* mutations, and increases in a stepwise manner during the adenoma-carcinoma sequence^{118, 119}. Indeed, *MGMT* methylation has also been reported to define a field-defect for CRC¹²⁰.

3.3. Abnormal karyotypes in cancer

Chromosome abnormalities are ubiquitous characteristics of most solid tumors. Aneuploidy, the state of harboring an abnormal number of chromosomes, was first associated with cancer more than a century ago¹²¹ and it is now recognized as a hallmark of cancer. Chromosome segregation errors due to the loss of mitotic fidelity are the main cause of aneuploidy. Furthermore, cancer cells also display numerical CIN, a form of genomic instability defined as high rates of chromosome missegregation leading to the continuous gain or loss of whole chromosomes¹²². Indeed, CIN is recognized as a common feature of many aneuploid cancer cells, driving karyotypic evolution and tumor

heterogeneity^{123, 124}. Although CIN is often cited as the most common pathway to aneuploidy, other studies have reported that aneuploidy, *per se*, leads to CIN¹²⁵⁻¹²⁹. Additionally, tumor cells may also exhibit structural CIN, defined as the continuous gain or loss of chromosome fragments as well as chromosome rearrangements (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>).

4. Aneuploidy and chromosomal instability

4.1. Causes of numerical CIN

4.1.1. Spindle assembly checkpoint defects

During mitosis, chromosomes attach to the microtubule (MT) spindle at structures known as kinetochores (KTs), a set of proteins that assemble onto centromeric chromatin. The spindle assembly checkpoint (SAC), also known as the mitotic checkpoint, ensures the accurate segregation of chromosomes by delaying the metaphase-to-anaphase transition until all chromosomes form proper attachments in order to undergo bipolar cell division¹³⁰. Initially, the molecular components of the SAC were identified in genetic screens of yeast and, later, their homologs in higher eukaryotes were identified. The six proteins necessary for SAC function are: Mad1, Mad2, BubR1, Bub1, Bub3 and Mps1¹³¹⁻¹³⁵. SAC remains active in mitosis as long as unattached KT's are present. Specifically, Mad1 and Mad2 localize to unattached KT's¹³⁶ causing a conformational change in Mad2 that leads to the binding of Cdc20, the activator of the anaphase promoting complex (APC/C)¹³⁷. Furthermore, Bub1 and BubR1 bind to the KT and interact with Mad2, forming the mitotic checkpoint complex, which inhibits APC/C¹³⁸. Conversely, when all KT's are attached, SAC is inactivated and APC/C uses different substrates to promote the completion of mitosis and the anaphase onset. For example, known APC/C substrates, such as Securin and Cyclin B, allow sister chromatid separation and mitotic exit, respectively¹³⁹.

Impairment of the mitotic checkpoint increases the rate of chromosome missegregation and CIN (**Figure 6A**)^{140, 141}. Nevertheless, complete inactivation of the mitotic checkpoint is lethal to cultured cells and mice due to massive chromosome missegregation^{142, 143}. In fact, mutations in the SAC genes are not commonly found in

cancer^{144, 145} and it is their altered expression, such as Mad2 overexpression, what is occasionally observed^{146, 147}.

4.1.2. Cohesion defects

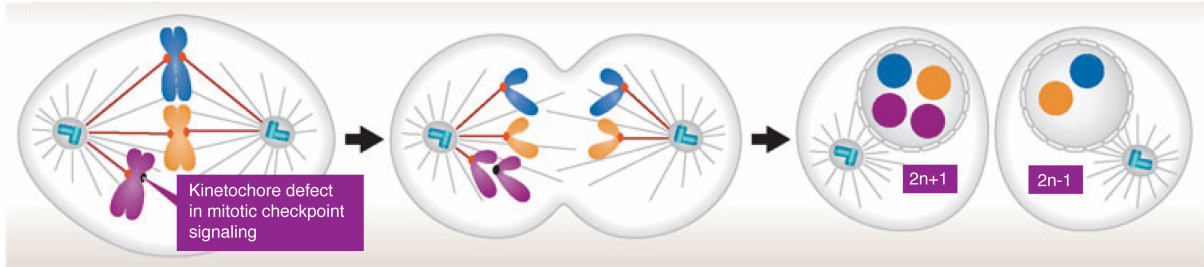
The separation of sister chromatids during mitosis is controlled by sister chromatid cohesion¹⁴⁸. As a result of a sequencing strategy, somatic mutations in the genes that regulate such sister chromatid cohesion were identified in one study¹⁴⁹. Interestingly, although the functional consequences of these mutations on chromosome segregation were not assessed, their results suggested that defects in the machinery that controls sister chromatid cohesion might cause CIN (**Figure 6B**). Moreover, mutations in *STAG2*, which encodes a subunit of the cohesion complex, were recently associated with aneuploidy in primary tumors and cancer cell lines¹⁵⁰. Additionally, acute depletion of the cohesion regulators Sgo1 and Separase increases the generation of tetraploid cells, indicating that defects in cohesion might lead to more broad defects in chromosome segregation rather than single chromosome missegregation^{151, 152}. Nevertheless, cohesion genes are barely mutated in tumor genomes as revealed by other large-scale sequencing studies^{145, 153}.

4.1.3. Supernumerary centrosomes

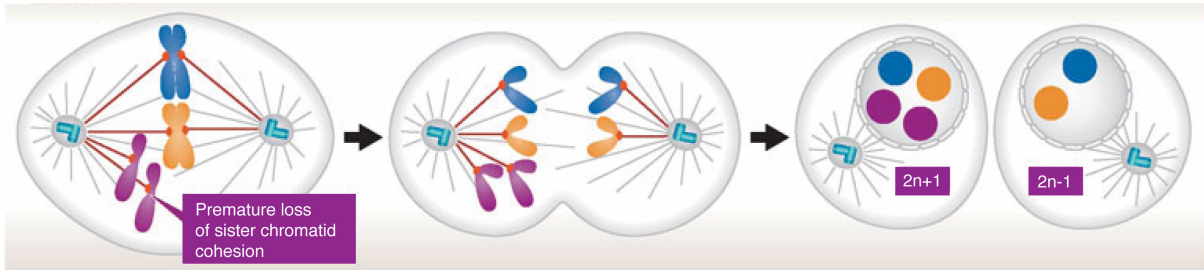
Centrosomes are the main MT organizing centers of cells responsible for the establishment of the bipolar MT spindle apparatus, in which chromosomes are segregated during mitosis. Centrosome amplification can arise through different mechanisms, including cytokinesis failure¹⁵⁴, mitotic slippage¹⁵⁵, centriole overduplication¹⁵⁶, centriole fragmentation¹⁵⁷, aberrant centriole cohesion¹⁵⁸, and cell fusion¹⁵⁹. It was over one hundred years ago when Theodor Boveri first proposed that centrosome amplification contributes to tumorigenesis¹⁶⁰. Since then, several studies have reported the presence of extra centrosomes in cancer cells and primary tumors^{161, 162} and it has been widely accepted as an early event in the development of most cancers^{163, 164}. The presence of extra centrosomes leads to multipolar mitotic spindles and the segregation of chromosomes into more than two daughter cells. However, multipolar spindles are infrequently observed during anaphase. Indeed, multipolar divisions generate daughter cells that can exhibit cell cycle arrest or postmitotic cell death¹⁶⁵. Conversely, cancer cells cluster their extra centrosomes to produce bipolar spindles and avoid unbearable missegregation events before

anaphase¹⁶⁶. Despite this clustering, cells with supernumerary centrosomes undergo a transient multipolar spindle stage during prometaphase that leads to high rates of chromosome missattachments, particularly merotelic (see next section), and produces CIN and aneuploidy (**Figure 6C**)^{165, 167}.

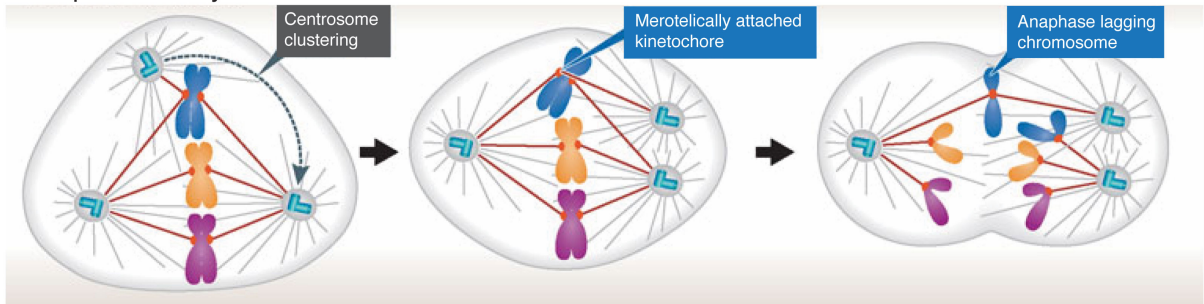
A. Spindle assembly checkpoint defects



B. Cohesion defects



C. Supernumerary centrosomes



D. Hyperstabilized kinetochore-microtubule attachments

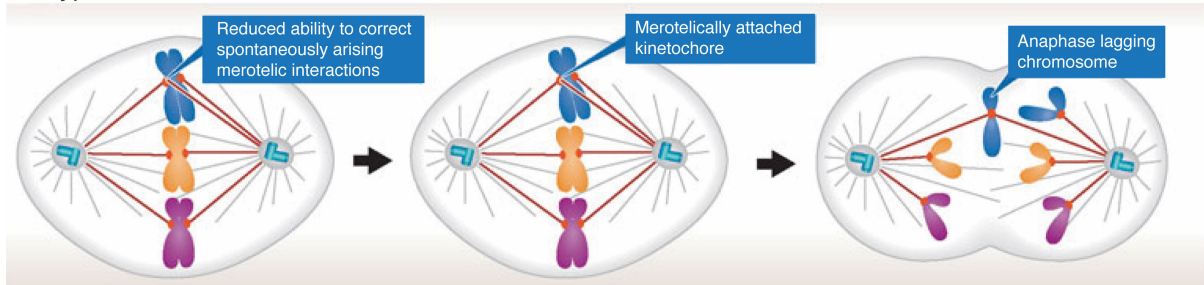


Figure 6. Pathways leading to aneuploidy. **(A)** Defects in the spindle assembly checkpoint. **(B)** Chromosome cohesion defects **(C)** Supernumerary centrosomes causing multipolar mitotic spindle **(D)** Hyperstable kinetochore-microtubule attachments. *Extracted from Holland, A.J. et al., EMBO Rep 2012.*

4.1.4. Kinetochore-Microtubule attachments

Correct chromosome segregation is ensured by the formation of amphitelic attachments, which occurs when the MT from opposite spindle poles attach to the two sister chromatids. However, other attachments can occur during mitosis, not all of which are recognized and corrected by the mitotic checkpoint (**Figure 7**).

Monotelic attachments occur when one of the sister KTs is attached to one spindle pole but the other sister KT is unattached. Although this type of attachment is very common early in mitosis, the mitotic checkpoint is able to recognize it and delay mitosis until it is corrected^{168, 169}.

Syntelic attachments are caused when MTs from the same spindle pole attach to two sister KTs on the same individual chromosome. This type of attachments have not been reported in cancer cells and they are not commonly observed in untreated cells, thus they may not have an important role in chromosome missegregation¹⁷⁰.

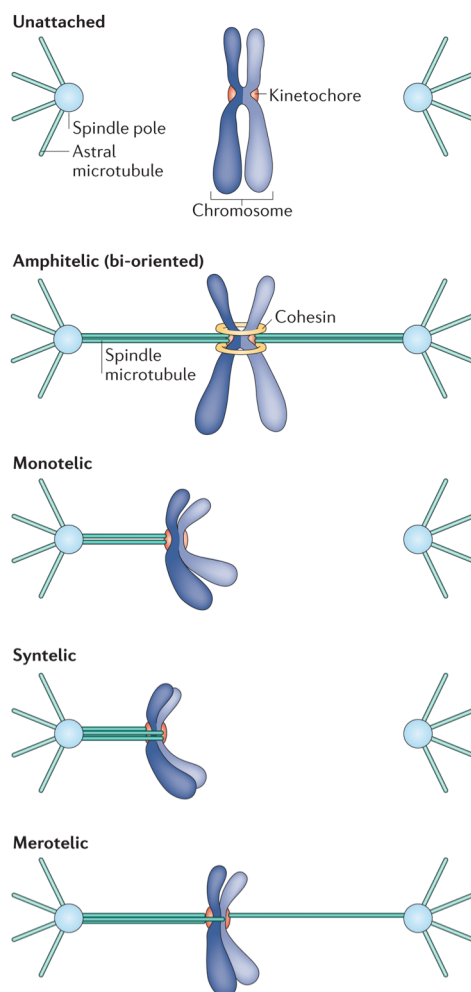


Figure 7. Types of kinetochore-microtubules attachments. *Extracted from London, N. et al., Nature Reviews 2014.*

Finally, **merotelic attachments** frequently occur in early mitosis when MTs anchored at both spindle poles attach to a single KT¹⁷¹. These attachments are possible because each KT has the capacity to bind 20-25 MTs in human cells and the main causes of them are supernumerary centrosomes and hyperstable KT-MT interactions. Although merotelic attachments can be corrected via an Aurora B-dependent mechanism before anaphase onset, they do not induce SAC-dependent mitotic delay and commonly persist into anaphase causing **lagging chromosomes**¹⁷². The fate of the lagging chromosomes depends on the relative size of the MT bundles bound to the individual KT. If the MT bundles have different sizes, the chromosome will be segregated to the side attached to the thicker bundle. Indeed, most merotelically attached chromosomes segregate correctly during anaphase because the smaller bundle of MTs is the one oriented to the wrong spindle pole. Consequently, the MTs oriented to the wrong spindle pole are detached, allowing the lagging chromosome segregate to the correct daughter cell¹⁷³. In contrast, when the two MT bundles are comparable in size the merotelically attached chromosome lags behind at the spindle equator and fails to segregate towards the spindles poles during anaphase, leading to aneuploid daughter cells¹⁷². These lagging chromosomes that fail to divide with the main chromosome masses undergo nuclear envelope reassembly and form micronucleus¹⁷³.

4.1.5. Microtubule dynamics

The high rates of anaphase lagging chromosomes observed in cancer cells are a consequence of increased rates of merotelic attachments formation or decreased rates of merotelic attachments correction^{165, 174}. The continuous association and disassociation of individual MT allows the efficient correction of incorrect KT-MT attachments. In fact, the reduction in the turnover rate of KT-bound MT causes missattachments and predisposes cells to CIN¹⁷⁵. The kinesin-13 proteins MCAK and KIF2B are two important microtubule depolymerases that regulate the microtubule turnover in an Aurora B-dependent manner. Loss of function of either MCAK or KIF2B results in higher rates of lagging chromosomes whereas increased levels of these proteins reduce the incidence of chromosome missegregation in CIN cells¹⁷⁵, indicating that they are crucial in the correction of erroneous KT-MT attachments. Therefore, perturbation of the proteins regulating microtubules dynamics creates more stable attachments and thus reduces the correction of erroneous KT-MT attachments and predisposes cells to CIN¹⁷⁶ (**Figure 6D**).

4.1.6. Polyploidization

Polyploidy is the condition of possessing extra haploid sets of chromosomes. The most common polyploidies in humans are triploidy and tetraploidy, which under normal physiological conditions are more tolerated than trisomies because they maintain the genomic content, in contrast to aneuploidy^{177, 178}. Indeed, polyploidization is even part of the developmental program in some tissues and polyploid cells are found in normal tissues such as liver and aorta, suggesting that polyploidy brings some advantages to certain cells^{179, 180}. In the case of cancer cells, polyploidy is a common event, accounting for a recurrence of over 50% for epithelial cancers, such as CRC¹⁸¹.

Tetraploid cells can arise through different mechanisms, including mitotic slippage, cell fusion, endoreduplication or cytokinesis failure. **Mitotic slippage** occurs when cells exit mitosis in the absence of chromosome segregation (**Figure 8**), which normally happens when cells are maintained in the presence of MT inhibiting drugs for long periods of time¹⁸². *In vivo*, this process has been observed in liver cells and it is suggested to confer resistance to cellular stresses¹⁸³. Indeed, although mitotic slippage has not been observed in cancer cells, it may play a role in the development of drug resistance.

Endoreduplication, the result of a round of DNA replication without the subsequent cell division¹⁸⁰ and **cell fusion** are important processes in the development of some normal tissues that result in terminally differentiated polyploid cells (**Figure 8**)^{184, 185}. Additionally, it has been reported that cell fusion can be induced by viral infection, which results in unstable tetraploid cells that are capable of transformation in the absence of p53^{186, 187}.

The most common mechanism to generate tetraploid cells is **cytokinesis failure**, which can be induced by defects in different cytokinetic proteins. Such defects lead to the failure of the actin-myosin cytokinetic ring to assemble or to initial furrow ingression followed by cleavage furrow regression¹⁸⁸. The result of cytokinesis failure is a single binucleated daughter cell that if progressed through the cell cycle would become a tetraploid cell with four centrosomes (**Figure 8**). In fact, cytokinesis failure is believed to be an occasional event in cancer cells because otherwise it would lead to an ever-increasing number of chromosomes.

Interestingly, a common feature of tetraploid cells is high rates of both numerical and structural chromosome aberrations, thus the CIN phenotype. The presence of extra centrosomes itself can lead to high rates of chromosome missegregation due to transient multipolarity or multipolar cell division, as previously described. Nevertheless, as tetraploid cells lose extra centrosomes overtime, other mechanisms may also lead to the high rates of CIN observed in tetraploid cells¹⁶⁵. Moreover, increased rates of DNA damage and less efficient DNA repair systems in tetraploid cells could also play roles in causing the reported structural chromosome aberrations¹⁸⁹.

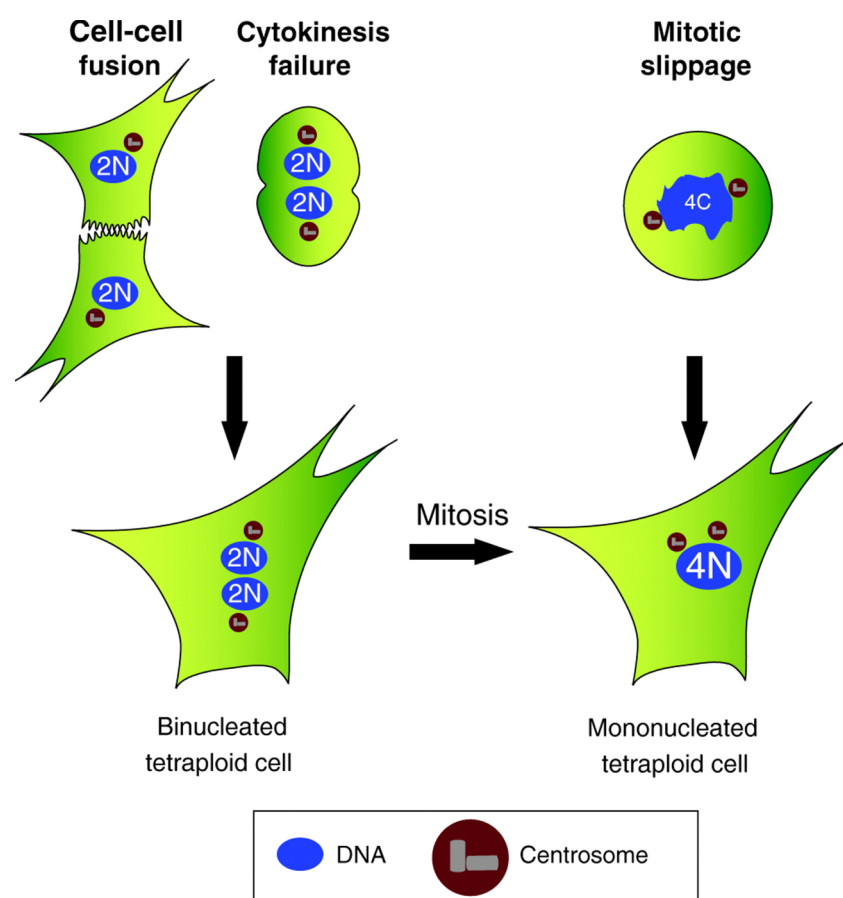


Figure 8. Pathways to tetraploidy. Cell fusion and cytokinesis failure generate binucleated cells with two centrosomes, which can form mononucleated tetraploid cells after successful passage through the next mitosis. Mitotic slippage is a cellular adaptation to persistent mitotic arrest that generates a tetraploid cell with a single nucleus but two centrosomes. 2N, diploid nucleus; 4N, tetraploid nucleus; and 4C, diploid nucleus with replicated chromosomes. *Reviewed in Storchova, Z. et al., J Cell Sci 2008.*

Furthermore, tetraploidy has been proposed to be an unstable intermediate on the road to aneuploidy and tumorigenesis. While a diploid cell with increased CIN would probably die after losing several chromosomes in an aberrant mitosis, a tetraploid cell might have a greater chance of surviving due to its redundancy in chromosome content¹⁹⁰. In fact, p53-deficient tetraploid mouse cells can initiate tumorigenesis in immune-compromised mice, whereas isogenic diploid cells do not¹⁵⁴. Also, ovarian aneuploid cells generated from a tetraploid intermediate form tumors in mice¹⁹¹. Furthermore, the presence of tetraploid or near-tetraploid cells has been described in the early stages of cancers, even preceding aneuploidy and cellular transformation, such as in Barrett's esophagus, cervical and colorectal cancer¹⁹²⁻¹⁹⁴.

In many cases, experimentally created tetraploid cells undergo a p53-dependent cell cycle arrest in G1 phase, raising the idea of a "tetraploid checkpoint"^{195, 196}. A recent study showed that oxidative DNA damage triggered by chromosome missegregation was responsible for the p53-dependent arrest in newly generated tetraploid cells¹⁹⁷. Moreover, it has been reported that extra centrosomes activate the Hippo tumor suppressor pathway¹⁹⁸ and abnormalities associated with tetraploidy formation can also limit cell proliferation through the activation of cell death pathways, such as apoptosis¹⁹⁹. Nevertheless, some spontaneously generated tetraploid cells are capable of proliferating and other studies have found that it is not necessary for them to undergo a p53-dependent arrest¹⁹¹. In fact, it has been reported that the arrest observed in some tetraploid cells may probably be a result of the drug treatment to generate these cells more than a consequence of tetraploidy or the presence of extra centrosomes *per se*^{200, 201}. Likewise, others observed no differences in the generation of tetraploidy when cells with different p53 status were treated to block cytokinesis¹⁵⁴. Nevertheless, the fact that cells that are p53^{-/-} or lack a functional apoptotic pathway spontaneously generate or are easier to induce tetraploidy indicates that there should be a mechanism of controlling tetraploidy formation and propagation, although this remains poorly understood.

4.2. Causes of structural chromosome instability

In addition to numerical chromosome alterations, cancer cells often exhibit structural chromosome abnormalities, including deletions, duplications, inversions and

translocations. Some of these structural alterations arise through the inappropriate repair of DNA double strand breaks (DSBs) or shortened telomeres, leading to the breakage and fusion of two chromosomes, and ultimately, the formation of dicentric chromosomes possessing two centromeres. During mitosis, dicentric chromosomes form **chromosome bridges** between daughter cells, which usually break during abscission, causing another round of DNA repair and chromosome fusion^{202, 203}. Cancer cells often activate the expression of telomerase in order to overcome this cycle, but after only a few rounds of chromosome breakage structural aneuploidy is evident. Moreover, mutations in cell cycle proteins may allow cells with DSBs to divide, thereby facilitating structural chromosome instability²⁰⁴. Conversely, other studies argued against the breakage of chromosome bridges during anaphase and showed that they can persist beyond completion of mitosis and delay abscission, form micronuclei and even cause cleavage furrow regression and lead to the formation of binucleated cells²⁰⁵⁻²⁰⁷.

Additionally, chromosome bridges can be caused by defects in proteins implicated in the decatenation and cohesion of sister chromatids. For example, the enzyme topoisomerase II decatenates the two DNA molecules after DNA synthesis, and defects in this protein lead to high frequencies of chromosome bridges and even cell division failure when the impairment is severe²⁰⁸⁻²¹⁰. Similarly, anaphase chromosome bridges and defective cell division can result due to defects in cohesin degradation^{211, 212}.

4.3. Consequences of aneuploidy and chromosome missegregation

4.3.1. Immediate effects of chromosome missegregation

On the one hand, anaphase lagging chromosomes can be trapped and damaged in the cleavage furrow during cytokinesis, leading to the activation of the **DNA damage response (DDR)**. Consequently, the DNA repair of the broken chromosome can itself cause translocations and deletions²¹³. Furthermore, lagging chromosomes that are not incorporated into the reforming nucleus of any of the daughter cells form their own micronuclei²¹⁴. These micronuclei undergo slow DNA replication due to stalled replication forks²¹⁵, which increases DNA damage and leads to extensive DNA rearrangements, such as chromotripsis²¹⁶.

And, on the other hand, problems during chromosome segregation result in **p53 activation and G1 arrest**¹²⁸, thus regulating the progression of aneuploid cells. Although the aspect of chromosome missegregation that activates p53 remains unknown, some possibilities have been explored. One study suggested that DNA damage during cytokinesis causes p53 activation²¹³, while another suggested that it is aneuploidy, *per se*, that activates p53¹²⁸. Other authors, however, have reported that reactive oxygen species (ROS) cause the activation of the DNA damage checkpoint kinase ataxia-telangiectasia mutated (ATM), which in turn activates p53²¹⁷.

4.3.2. Cellular responses to aneuploidy

As Boveri initially proposed²¹⁸, the gain or loss of hundreds of genes due to aneuploidy alters the cellular physiology and is typically detrimental for the cell survival under normal conditions²¹⁹⁻²²¹. In both cancer cells and budding yeast, it has been shown that changes in mRNA expression correlate with chromosome copy number^{219, 222, 223}. Nonetheless, aneuploidy not only affects the expression of the genes located on the altered chromosomes, but also the expression of multiple other genes located across the entire genome^{220, 224-226}, likely due to transcriptional regulators located on the aneuploid chromosome(s).

Interestingly, a uniform transcriptional response has been reported in aneuploid model cell lines regardless of the identity of the affected chromosome^{225, 227}. Transcriptome analysis of aneuploidy budding yeast revealed an enrichment for RNA processing and energy metabolisms related transcripts, which resembles the environmental stress response observed in yeast. This response was observed in yeast grown under stress conditions and reflects changes in cell-cycle distribution, emphasizing the negative effect of aneuploidy on cell-cycle progression²²⁵. Likewise, aneuploid mammalian cells show a common expression pattern in response to aneuploidy defined by the upregulation of genes linked to the endoplasmatic reticulum, Golgi apparatus and lysosomes; as well as the downregulation of genes linked to DNA replication, transcription and ribosomes²²⁷.

Furthermore, as protein abundance also scales with changes in gene copy number, aneuploid cells exhibit imbalanced levels of proteins. Indeed, some proteins such as kinases and subunits of macromolecular protein complexes are subjected to dosage compensation

in aneuploid human cell lines in order to restore the protein stoichiometry²²⁸. Nevertheless, the gain or loss of an extra chromosome impairs protein folding and activates protein degradation pathways, thus causing **proteotoxic stress**, which has a negative effect on cellular fitness as it impairs cell proliferation and growth^{229, 230}. This proteotoxic stress is characterized by limited levels of chaperones, such as Hsp90, increased activity of the proteasome²³¹, and insufficient autophagy in aneuploid cells²²⁸.

Despite that aneuploidy has a deleterious effect on cellular fitness under physiological conditions, there are circumstances in which aneuploidy can confer a selective advantage, such as increasing cellular tolerance to stress. Aneuploidy has been shown to provide drug resistance²³² and improve the survival of budding yeast under conditions of limited nutrients or in the presence of genomic damage^{222, 233}. Another recent study reported that aneuploid cancer cells displayed increased fitness under several stress conditions related to the tumor microenvironment, including serum starvation, hypoxia, and chemotherapy treatment²³⁴. Therefore, aneuploidy can serve as an effective mechanism for generating phenotypic variation and adaptation under a strong selective pressure.

5. Tumor heterogeneity

In 1976, Nowell proposed a model of clonal evolution for cancer, in which linear tumor progression is driven by a sequential selection of mutant subpopulations derived from a common progenitor²³⁵. In the years since, the comprehensive genomic analyses of cancer revealed the existence of two important features: **intertumor heterogeneity** and **intratumor heterogeneity (Figure 9)**. The former refers to the fact that each tumor contains multiple genomic aberrations, and that only a few of them are commonly shared between patients with the same tumor subtype. Intratumor heterogeneity, however, is defined as the variation of genomic abnormalities observed within an individual tumor caused by the presence of different subclones. The coexistence of these genetically distinct subclones, we now know, is driven by genomic instability, including CIN^{236, 237}. In fact, intratumor heterogeneity is associated with worse clinical outcome²³⁸ and drug resistance²³⁹.

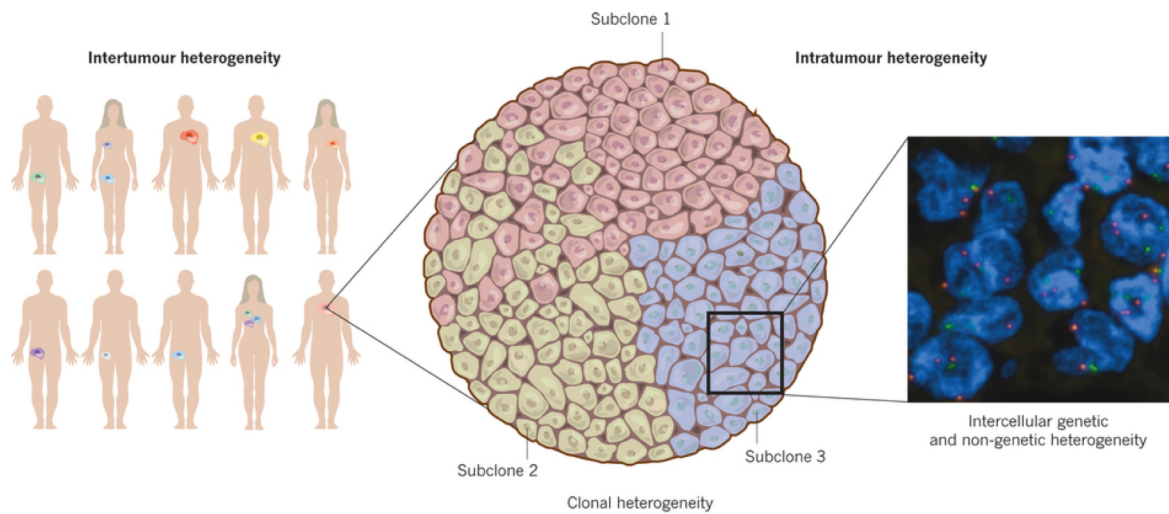


Figure 9. Intertumor and intratumor heterogeneity. *Extracted from Burrell, R.A. et al., Nature 2013.*

The subclonal architecture of a tumor can be reconstructed performing tumor deep-sequencing and single cell analysis, such as multiplex fluorescence in situ hybridization (mFISH), thus enabling the accurate evaluation of intratumor heterogeneity^{240, 241}. Interestingly, an increasing amount of evidence suggests that tumor progression follows a branched rather than a linear pattern of evolution. Accordingly, distinct subclones that are even regionally separated evolve in parallel within the same tumor, producing the extensive subclonal diversity (**Figure 10**). Such branched evolution was first demonstrated in hematological tumors²⁴¹ and later in renal, lung, breast, and more recently in colorectal cancer^{236, 242-244}.

The branched evolution of tumors provides a conceptual framework to study cancer evolution, predict the development of drug resistance and plan effective chemotherapy, as well as identify biomarkers²⁴⁵. Cancer evolution is also referred to as a spatio-temporal process, meaning changes over space and time may occur. Therefore, early somatic events that drive the tumor evolution are present in early clonal progenitors and maintained throughout the process evolution. Conversely, those somatic events occurring after the branched separation are only present in some subclones that might be spatially separated and thus difficult to identify using the current tumor screening strategies (**Figure 9 and 10**). Moreover, the selection for certain subclonal populations within a tumor can occur under

different microenvironment contexts, such as during disease progression or drug treatment^{246, 247}. Consequently, the identification of both high- and low-frequency subclones is essential to understand tumor progression and outcome²⁴⁸.

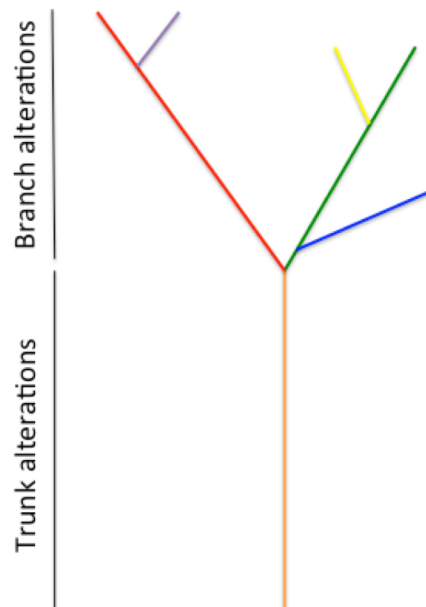


Figure 10. Branched evolution of tumors. Trunk alterations are present in all the subclones forming the tumor; however, branch alterations are characteristic of each subclone and thereby are more difficult to detect.

The paradigm of monoclonal evolution during CRC development was challenged a long ago, when polyclonal adenomas were described²⁴⁹. Later, the analysis of carcinoma in adenoma samples revealed polyclonal evolution and established *APC* mutations as the main trunk mutation for CRC. Moreover, the timing of genetic events was described as more variable as previously stated, indicating that the alteration of *APC* and *TP53* genes by either mutation or loss of heterozygosity (LOH) were the ones preceding cancer development²⁵⁰. Multiregional sequencing of CRC also showed extensive intratumor heterogeneity and branched evolution. Interestingly, trunk alterations leading to progression were composed by not only mutations in the known driver genes *APC* and *KRAS* but also by DNA hypermethylation of CpG islands and CNAs of the following chromosomes: 7p, 13p, 10p, 20p and 20q²⁴⁴. In fact, the gain of the chromosome 20q has been previously described as an

early event in colorectal carcinogenesis^{251, 252}. Therefore, all these data reinforce the branched evolution in CRC and the central role CNAs and aneuploidy during cancer progression.

6. The role of replication in cancer

6.1. Replication in normal conditions

DNA replication ensures the faithful duplication of the genetic material in order for it to be transferred between dividing daughter cells. Importantly, replication is temporarily and spatially controlled within each cell, as occurs with many other metabolic processes, such as transcription, mRNA processing, as well as chromatin remodeling. Replication begins at replication origins, whose activation is controlled by the availability and the subsequent binding of the replicative helicase mini-chromosome maintenance complex 2-7 (MCM2-7) and other factors such as CDC45, the GINS and claspin-TIMELESS-TIPIN complexes. The binding of MCM2-7 to the replication origins occurs during G1 phase and licenses the origin to start replicating the DNA in the following S phase, which is in turn triggered by the cyclin-dependent kinase 2 (CDK2) in what is known as origin firing²⁵³. Nonetheless, many of the licensed origins are not activated during normal replication, but instead, are maintained as back-up origins in the event replication problems arise²⁵⁴.

6.2. Replication stress

Problems that occur during DNA replication result in **replication stress (RS)**, one of the main sources of genomic instability. RS can be caused by numerous endogenous and exogenous sources and results in the stalling of the replication forks and the accumulation of single strand DNA (ssDNA), which is rapidly coated by the protein **RPA**. The presence of RPA bounded to ssDNA leads to the activation of the kinase **ATR** (Ataxia-Telangectasia mutated and Rad3 related), which phosphorylates **CHK1** kinase in the so-called RS response (**Figure 11**)^{255, 256}. Two of the main functions of this response are to stabilize stalled replication forks via the phosphorylation of SMARCAL1 and the claspin-TIMELESS-TIPIN complex^{257, 258} and to arrest cell-cycle and repress any late origins through mitotic entry and origin firing inhibition²⁵⁹. Moreover, encountering a problem during replication may result in an excess of origin firing, excessive consumption of nucleotides (NTPs), and ultimately in

fork collapse, which produces the accumulation of ssDNA and DSBs at the replication forks. Consequently, the accumulation of DSBs activates the DNA damage response (DDR), which results in the recruitment and activation of **ATM** kinase and the subsequent phosphorylation of **CHK2** kinase (**Figure 11**)²⁶⁰. Furthermore, DNA damage during replication also results in the phosphorylation of **H2AX** (γ H2AX) via ATM and ATR, as well as the formation of **53BP1** foci, both which are commonly used as RS markers^{261, 262}. After encountering a problem cells use different pathways to restore replication and resolve DNA damage, including homologous recombination, the Fanconi anemia pathway or post-replicative repair. Accordingly, the proper function of these pathways is essential for maintaining genomic stability²⁵⁸.

6.3. Replication stress, genomic instability and cancer

Growing evidence suggests that RS is a major driver in cancer development and progression. Indeed, regions particularly sensitive to RS such as common fragile sites (CFSs) and early-replicating fragile sites (ERFS), are prone to deletions and LOH in human precancerous lesions and cancers; therefore, suggesting that RS is prevalent in human cancers and a driver of genomic instability^{258, 263-266}. Moreover, some cancers appear to rely on the RS pathway to keep their rates of proliferation and survival high²⁶⁷.

Oncogene activation is a major source of RS during transformation because their activation alters DNA replication, which in turn increases RS, DNA breaks and triggers DDR activation²⁶⁸. Different mechanisms by which oncogenes induce RS have been proposed. One is the deregulation of origin licensing/firing, therefore causing the inappropriate use of replication origins. For example and similarly to what happens with mutations in the MCM components, the overexpression of the oncogene *CCNE1* reduces the number of licensed replication origins causing RS due to the lack of back-up origins to cope with stalled forks²⁶⁹⁻²⁷¹. Conversely, overexpression of other oncogenes, such as *MYC* or *RAS*, increases the number of licensed origins and, consequently, reduces the cellular pool of NTPs, thus causing RS²⁷²⁻²⁷⁴. In fact, oncogenes produce a deregulated entry to cell cycle that is not accompanied by the activation of other pathways needed for the proper progression through cell cycle; for example, the components of the Ribonucleotide Reductase (RNR) RRM1 and RRM2, implicated in the generation and regulation of NTPs. Consequently, the forks

progress slowly and stall, which can be rescued by the supplementation of nucleosides²⁷⁵,²⁷⁶. Interestingly, however, it is the overexpression of RNR components that is frequently observed in human cancers, maybe as an effort by the cell to deal with RS²⁷⁶.

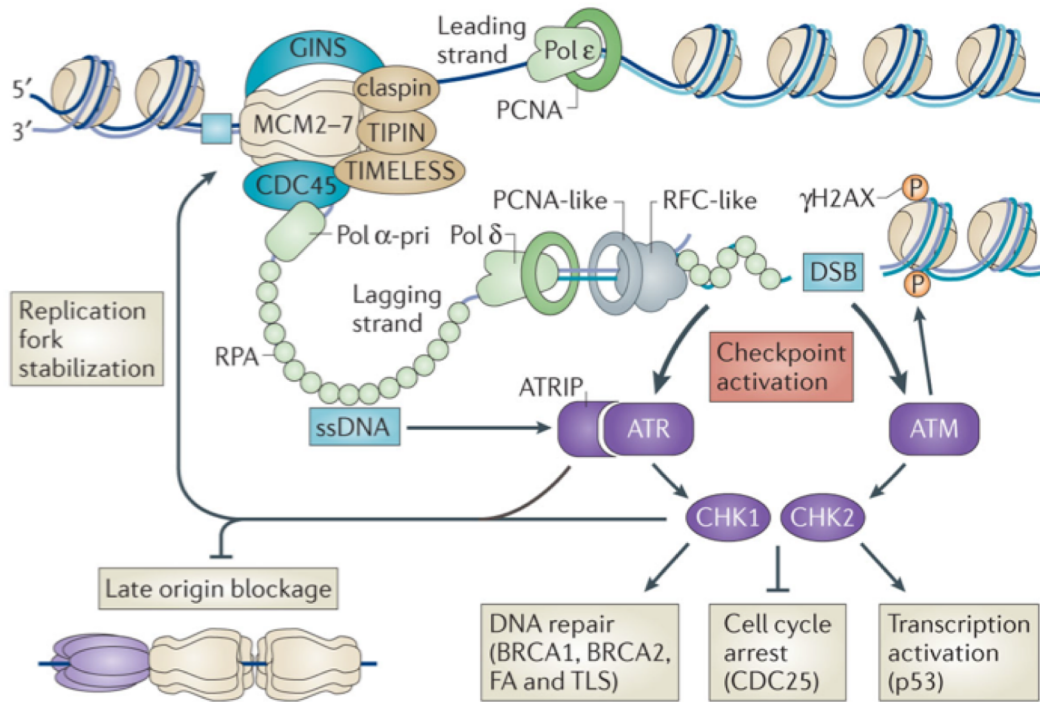


Figure 11. Checkpoints activated during RS. Replication is catalyzed by the replisome multi-subunit complex, which is composed by MCM2-7, CDC45 and the GINS and claspins-TIMELESS-TIPIN complexes. The DNA polymerases are anchored to DNA by the protein PCNA. When a replication fork stalls, it generates the formation of ssDNA and DSBs, which in turn activate the ATR and ATM kinases, respectively. These kinases then phosphorylate different substrates such as histone H2AX and the effector kinases CHK1 and CHK2. These kinases coordinate replication fork stabilization, late origin blockage, DNA repair, cell cycle arrest and transcription activation. *Reviewed in Gaillard, H. et al., Nature 2015.*

The third mechanism implicated in the RS caused by oncogenes is the interference between replication and transcription, which causes the formation of RNA-DNA hybrids known as R-loops²⁷⁷. Highly transcribed genes with GC rich promoters²⁷⁸, as well as large genes and ERFS are more prone to produce R-loops²⁶³. RNA helicases are the main factors cooperating to resolve R-loops; however, the mutation of any of these factors promotes genome instability²⁷⁹.

Furthermore, RS can also lead to chromosomal aberrations observed in malignant transformation. Indeed, as stated above, several deletions observed in cancer occur in regions difficult to replicate such as CFSs, ERFs or large genes^{263-265, 267}. The presence of unreplicated DNA in any of these regions has the potential to interfere in mitosis by forming common and ultrafine anaphase bridges, resulting in DNA damage and genomic instability^{280, 281}. Moreover, RS reportedly can cause CIN. In fact, RS has been found to be a differential characteristic of CIN+ cancer cells, along with chromosome segregation defects, such as the aforementioned anaphase bridges²⁸². In addition, RS has recently been described to cause cells to enter mitosis with chromosomes that have not been completely replicated, which also causes mitotic extra centrosomes and global unbalanced chromosome segregation²⁸³.

Another chromosomal aberration caused by RS is tandem segmental genomic duplications. When there is a collapsed fork, the homology is limited to only one side of the break inducing a specific repair pathway, the break-induced replication (BIR), to produce homologous recombination. Nevertheless, the BIR system can also generate amplifications with microhomology at the repair junctions, which resembles to sections of the common chromosomal alterations observed in human cancers²⁸⁴.

Interestingly, although RS induces genomic instability, mutations in ATR or CHK1 are not commonly observed in human tumors. In contrast, the expression of these RS regulators is upregulated in human cancers^{285, 286}. This can, in part, be explained by the need of tumors to maintain a proficient RS-response to deal with the high levels of RS and ensure their survival²⁵⁵. Consequently, targeting ATR or CHK1 can be toxic for those tumors with high levels of RS^{267, 287-289}. Therefore, taking advantage of the RS to target replication may offer new and promising ways to treat cancer.

HYPOTHESIS

This thesis has been elaborated under the following hypothesis:

1. ACF exhibit genetic and epigenetic alterations that correlate with CRC risk; therefore, these early lesions are suitable biomarkers for CRC development.
2. LINE-1 hypomethylation represents a field defect for CRC.
3. Chromosome alterations can trace the adenoma-to-carcinoma transition during colorectal carcinogenesis.
4. Genome duplication events influence cell behavior and confer increased intercellular heterogeneity.

OBJECTIVES

Main objective:

The general aim of this thesis is to study the colorectal carcinogenesis from two perspectives that cover important steps of the disease progression. The first is the evaluation of ACF, the earliest lesion detected in the colon, as a predictive biomarker for CRC development, as well as the role of LINE-1 methylation as a field defect in CRC. On the other hand, the second main objective is to identify molecular events that drive the transition from adenoma to carcinoma, including the role of genome doubling events.

Specific objectives:

1. Determine whether ACF are predictive biomarkers for CRC by analyzing morphological and molecular features described for colorectal carcinogenesis in a wide cohort of ACF samples from different risk groups, including healthy individuals, adenoma and CRC patients.
2. Evaluate whether LINE-1 methylation represents a field defect for CRC using the same cohort of healthy individuals, adenoma and CRC patients.
3. Analyze progressed adenomas in order to assess chromosomal heterogeneity and define the specific populations and CNAs involved in the adenoma-to-carcinoma transition.
4. Characterize tetraploidy *in vitro* as it is considered an intermediate stage in the progression of many cancers, including CRC, and a way to tolerate CIN.

MATERIALS AND METHODS

1. Patients

1.1. ACF and methylation studies (Objective 1 and 2)

A random selection of patients referred to our hospital for a diagnosis colonoscopy between 2008 and 2010 was prospectively included in the ACF studies. According to the endoscopic findings, patients were divided into two groups: 1) Control Group: healthy individuals with a normal colonoscopy and without a personal history of adenomas or CRC (n = 100) and 2) Case Group: patients with a personal history of CRC or current CC (n = 50) and patients with ≥ 1 current colonic adenomas (n = 50). For the methylation assessment, the case group was divided into two subgroups: adenoma and CC. Patient exclusion criteria were as follows: poor bowel preparation according to Boston scale (less than 2 points in any colonic segment and/or a total score of less than 6), incomplete colonoscopy, past or current rectal cancer, chemotherapy in the last six months, personal history of adenomas but were absent/not detected at time of colonoscopy, known hereditary CRC syndrome and inflammatory bowel disease. Patients with rectal cancer were excluded from the analysis to ensure that biopsies were taken at least 10 cm away from the colorectal neoplasia in all patients. In order to detect ACF, after the colonoscopy was completed, a different examiner who was blind to each patient's study group performed a rectoscopy.

The projects were granted approval from institutional review board of the Hospital Clínic of Barcelona, and written informed consent was obtained from all participants.

1.2. CRC evolution study (Objective 3)

Twenty-three formalin-fixed paraffin-embedded (FFPE) specimens of progressed adenomas and six non-advanced adenomas were retrieved from the archives of the Hospital Clínic of Barcelona. Progressed adenomas were defined as adenomas with a focus of invasive adenocarcinoma, while non-advanced adenomas were defined as lesions with low-grade dysplasia, tubular architecture and a size of < 10 mm. The adenocarcinoma regions of progressed adenomas were further evaluated and classified according to their histological grade, pT stage, millimeters of submucosa infiltration (in the case of pT1), budding and MSS status. Histological grade is considered as the pathological grade of colonic wall infiltration,

pT stage as the pathological grade of mucosa infiltration, and budding as the presence of tumor cells beyond the invasive front of the carcinoma²⁹⁰.

Moreover, all patients included in the study provided written consent for their participation and the project was granted approval by the Hospital Clinic of Barcelona.

1.3. Tetraploidy study (Objective 4)

Colorectal primary tumor FPPE samples were provided by the Hospital Clinic of Barcelona. All patients signed the corresponding informed consent form and the sample collection was approved by the local Ethics Committee. Sections of 4- μ m from FFPE blocks were obtained, and an experienced pathologist marked the corresponding bulk of the tumor and the invasive front based on hematoxylin-eosin (H&E) stained consecutive sections for each sample.

2. Endoscopic assessment of aberrant crypt foci

An expert endoscopist, Dr. Maria López-Céron, performed a systematic exam of the distal 10 cm of the rectum with a high-definition colonoscope (Olympus H180, Olympus, Tokyo, Japan) in all patients. ACF were defined as crypts with a larger diameter than the normal mucosa, a thicker epithelial lining and a dilated crypt lumen. ACF that raised > 2mm were considered polyps. The number of ACF per patient was categorized as less than 5, 5 to 15, and more than 15. Each rectum segment was examined clock-wise, proximal-to-distal to record ACF features and location; first, with Narrow Band Imaging (NBI) and then, with methylene blue 0.5% chromoendoscopy (CE), which was considered the gold standard for ACF detection. The size of the crypts assessed by CE was classified as small (less than 20 crypts per ACF), medium (20-40 crypts per ACF) or large (more than 40 crypts per ACF). The shape of the crypt lumens, as visualized by CE, was characterized as semicircular-oval, asteroid-like or irregular. Finally, ACF vascular pattern intensity (VPI) was described as weak, normal or strong according to the appearance of the surrounding mucosa as visualized by NBI.

3. Pathological assessment of aberrant crypt foci

Normal descending and rectal colon mucosa, and the largest five ACF were biopsied for each patient. ACF were immediately embedded in tissue freezing medium (OCT) and stored at -80°C, while the normal mucosa was conserved in PBS at -80°C. For each sample, three consecutive 5-7 µm sections were stained with H&E and analyzed by light microscopy. All samples were evaluated twice over a period of 6 months by two different pathologists (A and B) who were blinded to the endoscopic classification and each other's diagnoses. Discordant diagnoses were reviewed to reach a consensus diagnosis and histological findings were classified as: inadequate sample, normal mucosa, hyperplastic ACF, serrated ACF and dysplastic ACF based on the WHO classification³⁹. The "serrated morphology" category was incorporated in our final diagnosis in order to evaluate the possible role of ACF in the serrated pathway of colorectal carcinogenesis.

4. Preparation of cytopins from FFPE specimens

First, a 4-µm-thick section from archived paraffin blocks was cut and then stained with H&E so a pathologist assessed the histomorphologic features of each lesion to identify the non-advanced adenomas and the regions containing adenoma (AD) and adenocarcinoma (ADK) in progressed adenomas. The second and third sections were 6-µm-thick (for potential FISH or immunohistochemical studies on tissue sections), and the fourth and fifth sections were 50-µm-thick. In the case of progressed adenoma samples, a sixth 4-µm-thick section was again stained with H&E to ensure the location of the adenoma and adenocarcinoma along the paraffin block (**Figure 12**).

Using the two H&E-stained sections for guidance, the pathologist outlined the AD and ADK areas of progressed adenomas in the unstained 50-µm-thick sections as well as the proper region in non-advanced adenomas. Tissue punches were used to remove each area of interest and were placed in a tube with xylene for 20 minutes. The tube was then centrifuged for 3 minutes at 800xg, and the xylene was removed. The xylene wash was repeated an additional two times. The tissue was then rehydrated in an ethanol series of 100% (x2), 90%, 70%, and 50%.

After 5 minutes, the sample was centrifuged at 4°C for 15 minutes at 16,000 $\times g$. The ethanol was removed, and 1 mL of sterile water was added. After 20 minutes at room temperature, the sample was centrifuged at 4°C for 15 minutes at 16,000 $\times g$. After removing the water, 500 mL of 0.1% protease (type XXIV; P-8038; Sigma-Aldrich, Missouri, USA) in 1x PBS was added. The tube was then placed in a thermomixer at 45°C, shaking for 40 to 60 minutes.

The disintegration of the tissue sample was monitored under microscopic control every 15 to 20 minutes. The degree of disintegration was determined by placing 10 μ L of the protease digest on a slide and staining it with 10 μ L of DAPI-sulforhodamine solution (SR101; S-7635; Sigma-Aldrich, Missouri, USA). Optimal disintegration was determined by the quantity of nuclei and by the lack of cytoplasm and the intensity of the DAPI stain. Once the disintegration was optimal, the reaction was stopped with 500 μ L of 1x PBS, and the sample was centrifuged for 5 minutes at 600 $\times g$. The concentration of nuclei was adjusted so that 80 μ L of the cell suspension centrifuged in a Shandon Cytospin 3 centrifuge (ThermoScientific, California, USA) for 5 minutes at 200 $\times g$ resulted in a medium-dense, monolayer cytopsin. The slides were dehydrated in 70% and 90% ethanol for 5 minutes and in 100% ethanol for 10 minutes, air-dried, and stored at 4°C.

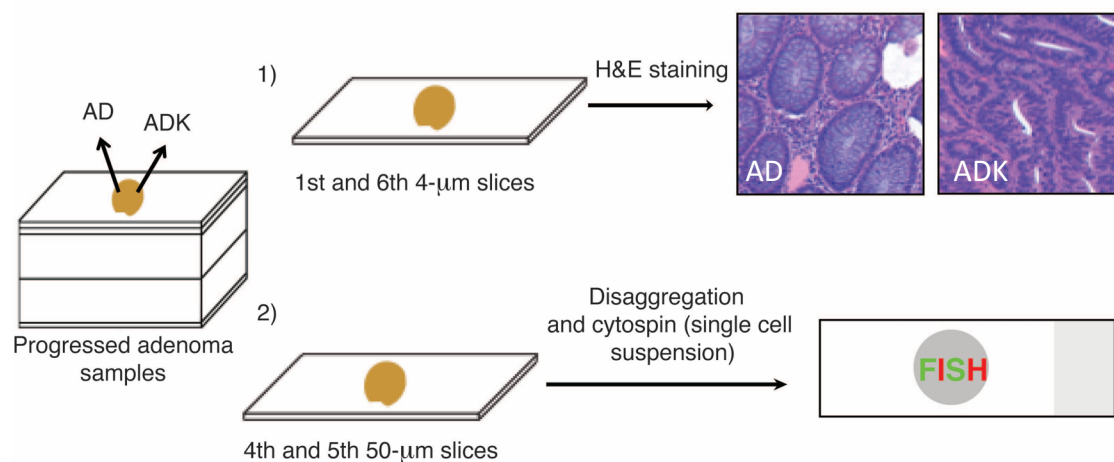


Figure 12. Diagram showing the work-flow for progressed adenoma samples.

5. Cell culture

5.1. Cell lines

DLD-1 and RKO cell lines were obtained from American Type Culture Collection and cultured in RPMI-1640 and DMEM-F12 medium, respectively, supplemented with antibiotics and 10% FBS at 37°C in 5% CO₂. The identity of the cell line was confirmed by the presence of unique chromosomal abnormalities as recorded in the spectral karyotyping (SKY)/array comparative genomic hybridization (aCGH) database (www.ncbi.nlm.nih.gov/sky/).

5.2. Tetraploid cell lines generation

To generate diploid (2N) and near-tetraploid (4N) clones, a suspension of bulk parental DLD-1 cells was single-cell flow-sorted by FACS into 96-well plates based on size differences (i.e., forward scatter). For the RKO cell line, cells were treated with 0.75µM of the actomyosin inhibitor dihydrocytochalasin D (Sigma-Aldrich, Missouri, USA) for 24 hours to generate a tetraploid sub-population before performing the single-cell sorting procedure. The plates were manually scrutinized to ensure that only wells harboring a single cell were used for subsequent culturing. The ploidy of individual clones was determined by karyotypic analysis of metaphase spreads. For the experiments described in this thesis, we used three 2N clones and up to seven 4N clones derived from the DLD-1 cell line. In regard to the RKO cell line, we used two 2N clones and up to four 4N clones.

6. Volume and area quantification

Nuclear volume was measured in 2N and 4N cells synchronized in G1/S with a double thymidine block, as previously described²⁹¹. Cells were harvested and dropped on slides following methanol/acetic acid fixation. Slides were stained with DAPI and systematically analyzed using the automatic acquisition and analysis Bioview 3.6.0.16 software (Bioview, Israel) on an Olympus BX63 (Olympus, Tokyo, Japan). Five thousand nuclei were scanned for each sample, repeated in triplicate, with overlapping cells being discarded. The Bioview software automatically calculated the diameter and area. The volume was measured using the formula: $V = 4/3 * A * c$ (A: Area, c: radius).

7. Nucleic acid extraction and quantification

DNA from ACF and normal mucosa was extracted using the AllPrep DNA/RNA Mini kit (Qiagen, Hilden, Germany) following the manufacturer's recommendations. DNA was then quantified using NanoDrop® Spectrophotometer ND-1000 (Thermo Fisher Scientific, Massachusetts, USA).

DNA from cultured cells was extracted using a standard phenol/chloroform extraction method. Briefly, the cells were first resuspended in 3 mL of DNA buffer and then 125 µL of proteinase K (Sigma-Aldrich, Missouri, USA) at 10 µg/µL were added before incubating the tubes at 37°C overnight for digestion. The contents of the tubes were later mixed with 3 mL of phenol:chloroform:isoamyl alcohol (25:24:1) for 10 minutes and centrifuged at 1,700xg for 10 minutes. The supernatant was transferred to a fresh tube and the phenol:chloroform:isoamyl step was repeated twice. Finally, the DNA was precipitated by adding 3M sodium acetate (pH 5.2) and cold absolute ethanol, and transferred into a sterile tube. Each DNA sample was brought to a final volume of 300-500 µL in ultrapure water and quantified using NanoDrop® Spectrophotometer ND-100 (Thermo Fisher Scientific, Massachusetts, USA).

DNA buffer (100 mL):

- 2 mL 5M NaCl
- 1 mL Tris-HCl 1M (pH 8)
- 5 mL EDTA 0.5M (pH 8)
- 5 mL SDS 10%
- 87 mL dH₂O
-

On the other hand, RNA from cultured cells was extracted using the RNeasy® Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's recommendations and quantified using NanoDrop® Spectrophotometer ND-100 (Thermo Fisher Scientific, Massachusetts, USA). Additionally, RNA quality was evaluated on a Bioanalyzer using the Agilent RNA 6000 Nano kit (Agilent Technologies, California, USA).

8. Mutational analysis

For feasibility reasons the molecular characterization was only performed for the first three histologically confirmed ACF samples from each patient.

KRAS mutational analysis: A 167 bp fragment of the KRAS gene spanning codon 12 and 13 was amplified by COLD-PCR amplification using the following primers: F, 5'-GCCTGCTGAAAATGACTGAA-3' and R, 5'-AGAATGGTCCTGCACCAGTAA-3'. Then, the Sanger sequencing method was used to analyze the PCR fragment.

APC mutational analysis: APC mutations were analyzed from two amplified fragments (A and B) that spanned the majority of the APC gene mutations in terms of frequency (82.6%). Primer sequences for the two fragments were: (a) A-F (5'-CAGTGAGAATACGTCCACACCT-3') and B-F (5'-TTTGAGAGTCGTTTCGATTGC-3'); and (b) A-R (5'-CATTCCACTGCATGGTTCAC-3') and B-R (5'-TGATGACTTTGTTGGCATGG-3'), generating fragment lengths of 598 and 579 bp, respectively, which were then sequenced by Sanger sequencing.

BRAF mutational analysis: BRAF V600E mutation genotyping was performed by Real-Time PCR using primers and probes designed, synthesized, and validated by the ThermoFisher Scientific custom Genotyping Assay Service (TaqMan™ technology). Both *wild-type* and mutant BRAF (V600E) were assessed in a single reaction.

Microsatellite instability (MSI) analysis: The MSI Analysis System, which consists of five nearly monomorphic mononucleotide markers (*BAT-25*, *BAT-26*, *NR-21*, *NR-24*, and *MONO-27*) for MSI determination and two polymorphic pentanucleotide markers (*Penta C* and *Penta D*) for sample identification was used. MSI analysis was performed according to the manufacturer's guidelines (Promega Corp., Wisconsin, USA). We classified the samples as MSI when ≥ 2 mononucleotide were MSI, while we considered samples as microsatellite stable (MSS) when only a single mononucleotide locus was found to be MSI or no instability was observed at any of the loci tested.

9. Bisulfite pyrosequencing and methylation analysis

Bisulfite treatment was performed using the EpiTect® Bisulfite kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Methylation analysis of *MGMT* and LINE-1 promoters was investigated using pyrosequencing-based methylation analysis¹¹³. PCR was carried out in a 25 µL PCR mixture containing 12.5 µL of GoTaq® Coloreless Master Mix (Promega, Madison, Wisconsin, USA) (Reaction Buffer pH 8.5, 400µM dATP, 400µM dGTP, 400µM dCTP, 400µM dTTP and 3mM MgCl₂), 1 µL unit of Taq polymerase, and 100 ng of bisulfite-treated genomic DNA. PCR cycling conditions were 95°C for 15 minutes; 45 cycles of 94°C for 30 seconds, 55°C for 45 seconds, and 72°C for 45 seconds; and finally, 10 minutes at 72°C and 4°C forever. The biotinylated PCR product was purified and made single-stranded to act as a template in a pyrosequencing reaction as recommended by the manufacturer using Pyrosequencing Vacuum Tool (Qiagen, Hilden, Germany). In brief, the PCR product was bound to a Streptavidin Sepharose HP and the Sepharose beads containing the immobilized PCR product were purified, washed, denatured using 0.2M NaOH solution and washed again. Then, 10µM of the pyrosequencing primer was annealed to the purified single-stranded PCR product and pyrosequencing was performed using the PSQ 96MA Pyrosequencing System (Qiagen, Hilden, Germany). As a positive control CpGenome Universal Methylated DNA was used (Millipore, Darmstadt, Germany). Methylation levels of *MGMT* and LINE-1 elements were calculated as the mean percentage of the four CpG sites analyzed. The different DNA primers used are detailed in **Table 5**.

Table 5. DNA primers used for DNA methylation analysis.

Primers	<i>MGMT</i>	LINE-1
Forward	5'-GGTAAATTAAGGTATAGAGTTTT-3'	5'-TTTTGAGTTAGGTGTGGGATATA-3'
Reverse-biotinylated	5'-AAACAATCTACRCATCCT-3'	5'-AAAATCAAAAAATTCCTTTC-3'
Pyrosequencing	5'-GGAAGTTGGGAAGG-3'	5'-AGTTAGGTGTGGGATATAGT-3'

10. Microarrays analysis

10.1. Array comparative genomic hybridization

Three µg of genomic DNA from six 2N and seventeen 4N clones and from commercially available pooled sex-matched control DNA were digested for two hours with AluI and RsaI (Promega, Wisconsin, USA) and purified using a QIAprep® Spin Miniprep Kit (Qiagen, Hilden, Germany). Sample and reference DNA were labeled for two hours with Cy3-dUTP and Cy5-dUTP (Promega, Wisconsin, USA), respectively, in a random priming reaction using Bioprime™ Array CGH Genomic Labeling Module (Life Technologies, California, USA). Unincorporated nucleotides were removed using Microcon® YM-30 columns (Merck-Millipore, Darmstadt, Germany). Cy3- and Cy5-labeled samples were combined in equal amounts according to the incorporation of labeled nucleotides as measured using a Nanodrop®. Oligonucleotide-based microarrays (Human Genome Microarray 4X180K, Agilent Technologies) were hybridized for 40 hours at 65°C. Microarrays were washed using conditions recommended by the manufacturer, and were scanned using a laser scanner (G2565BA, Agilent Technologies, California, USA). Agilent Feature Extraction software (v10.7.3.1) (Agilent Technologies, California, USA) was applied for image processing. Segmentation was done by the R package DNACopy (v1.44.0) using default options²⁹², and was visualized using Circos plots²⁹³. Cut-off for defining gains and losses was set to 0.22.

10.2. Gene expression

High quality total RNA, with a RIN value above eight, from eight 2N and sixteen 4N clones was reversed-transcribed using oligo dT-promoter primers. Amplification and labeling were then performed using a T7 RNA polymerase and Cy3, respectively, from the Low RNA Input Fluorescent Linear Amplification Kit (Agilent Technologies, California, USA). Purified cRNA was hybridized to 4x44K oligonucleotide-based Whole Human Genome Microarrays (Agilent Technologies, California, USA) for 17 hours at 65°C. The microarrays were then washed and imaged with an Agilent G2565CA Microarray Scanner System (Agilent Technologies, California, USA). Signal data were collected with the dedicated software Agilent Feature Extraction (v10.7.3.1). Control probes and probes with “glsPosAndSignif” value of zero in at least three samples were disregarded and labeled as “null”. Background adjustment and normalization were done using normexp and quantile

normalization from the Limma R package²⁹⁴. Median values were calculated for multiple probes with the same Agilent Probe ID. Analysis for differential expression on a probe basis was done using Limma moderated t-statistics and correction for multiple testing (False Discovery Rate, FDR) was performed using the Benjamini-Hochberg method.

For the current study, Gene Set Enrichment Analysis (GSEA) was used with gene sets extracted from Molecular Signature Database (MSigDB v.4-0) C2 (manually curated gene sets from online pathway databases, publications in PubMed and knowledge of domain experts). Only gene sets with more than 15 genes and less than 500 genes were included in the analysis. Analysis was based on a Fisher's exact t-test and a weighted scoring scheme with 150,000 permutations²⁹⁵.

11. Real-Time PCR

According to the results of the gene expression microarray, quantification of *ESPL1* expression was performed using the TaqMan® Assay Technology (ThermoFisher Scientific, California, USA). cDNA was generated from 1 µg of total RNA using oligo dT primers (ThermoFisher Scientific, California, USA) together with the High capacity cDNA RT kit (ThermoFisher Scientific, California, USA) according to the manufacturer's instructions. Real Time PCR amplifications were performed in triplicate on an ABI 7900 Sequence Detection System using Universal PCR Master Mix without AmpErase® UNG (ThermoFisher Scientific, California, USA) and TaqMan® made-to-order probes. The thermal cycling conditions comprised an initial denaturation step at 95°C for 10 minutes followed by 45 cycles at 95°C for 15 seconds and 60°C for 1 minute.

ESPL1 expression was first normalized to that of a control transcript, *GAPDH* or *YWAHZ*, with $\Delta Ct = Ct_{ESPL1} - Ct_{GAPDH \text{ or } YWAHZ}$. Diploid clones were used as calibrator sample, thereby *ESPL1* expression was expressed as $2^{-\Delta\Delta Ct}$ where $-\Delta\Delta Ct = -(\Delta Ct_{\text{sample}} - \Delta Ct_{\text{calibrator}})$.

12. Protein analysis

12.1. Western blot

Cells were trypsinized and incubated in RIPA buffer (50mM Tris-HCl, 1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS, 150mM NaCl, 2mM EDTA, and 50mM NaF, with protease inhibitors). The lysates were boiled for 5 minutes at 95°C with NuPAGE™ LDS Sample Buffer (4X) (ThermoFisher Scientific, California, USA). Protein samples were resolved by 4-12% SDS-PAGE and electroblotted onto a PVDF membrane. The membrane was blocked by soaking in TBS, 0.1% Tween 20, and either 5% milk or 3% BSA for 1 hour, incubated with primary antibody diluted in blocking solution overnight at 4°C, washed three times with TBST (TBS, 0.1% Tween 20), incubated with secondary antibodies for 1 hour at room temperature, and washed three times with TBST. For the detection of signals, SuperSignal™ West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific, California, USA) was used according to the manufacturer's recommendations. Membranes were imaged on an ImageQuant™ LAS400 device (GE Healthcare Life Sciences, Little Chalfont, UK) and blots were quantified using ImageJ software.

RIPA buffer:

- 25 mL Tris-HCl 1M (pH 7.4)
- 5 mL NP-40
- 2.5 g Na-deoxycholate
- 0.5 g SDS
- 15 mL NaCl 5M
- 2 mL EDTA 0.5M
- 1.05 g NaF
- 453 mL dH₂O

The primary antibodies and dilutions used were as follows: rat anti-RPA32 (Cell Signaling), 1:2,000; rabbit anti-phospho-MCM2 (Bethyl Laboratories, Texas, USA), 1:1,000; rabbit anti-CHK1 (Santa Cruz Biotechnology, Texas, USA), 1:500; rabbit anti-phospho-CHK1 (Cell Signaling, Massachusetts, USA), 1:100; rabbit anti-CHK2, 1:500; and rabbit anti-phospho-CHK2 (Thr68) (Novus Biologicals, Colorado, USA), 1:500. Blots were detected using

goat anti-rabbit (1:2,000) and goat anti-rat (1:2,000) secondary antibodies conjugated to horseradish peroxidase (ThermoFisher Scientific, California, USA).

12.2. Immunostaining

Cells were grown on sterilized 22x22 mm coverslips inside 6-well plates. When required, cells were treated with 0.2 μ M of the DNA polymerase inhibitor aphidicolin (Sigma-Aldrich, Missouri, USA) for 24 hours. Cells were then fixed and permeabilized either using 4% paraformaldehyde for 10 minutes followed by 10 minutes in 0.5% Triton X-100 or in a single step with ice-cold methanol. Incubation with selected primary antibody was done based on manufacturer's recommendations. When required, cells were treated with 0.2 μ M of the DNA polymerase inhibitor aphidicolin (Sigma-Aldrich, Missouri, USA) for 24 hours.

The primary antibodies and dilutions used in this study included: rabbit anti- γ -tubulin (Sigma-Aldrich, Missouri, USA), 1:1,000; rabbit anti-53BP1 (Novus Biologicals, Colorado, USA), 1:1,000; mouse anti-phospho-H2AX (S139) (Millipore, Massachusetts, USA), 1:500; rabbit anti-phospho-Histone H3 (Millipore, Massachusetts, USA), 1:1,000; mouse anti- γ -tubulin (DM1A, Sigma-Aldrich, Missouri, USA), 1:500; human anti-centromere antigen (Antibodies Inc., California, USA), 1:100; and mouse anti-cyclin D1, 1:50 (Abcam, Cambridge, UK). The secondary antibodies used were as follows: Rhodamine Red-X goat anti-human (Jackson ImmunoResearch, Philadelphia, USA), 1:1,000; Alexa Fluor488 goat anti-rabbit 1:500; AlexaFluor594 goat anti-mouse 1:1,000; Alexa Fluor 488 goat anti-mouse (ThermoFisher Scientific, California, USA), 1:1,000.

12.3. Microscopy analysis

Immunostained cells were imaged with a swept field confocal system (Prairie Technologies, Texas, USA) equipped with a multiband pass filter on a Nikon Eclipse TE2000-U inverted microscope (Nikon Instruments Incorporated, Tokyo, Japan) equipped with a 100x/1.4 NA Plan-Apochromatic objective and an automated ProScan stage (Prior Scientific, Cambridge, UK). Digital image stacks were acquired with a HQ2 CCD camera (Photometrics, Arizona, USA) using 0.5 μ m Z-steps. Exposure time, Z-axis position, laser line power, and confocal system were all controlled by NIS Elements AR software (v4.0) (Nikon Instruments

Incorporated, Tokyo, Japan). All data were quantitatively analyzed with the software ImageJ²⁹⁶.

13. Cytogenetic procedures

13.1. Chromosome harvesting and cytokinesis-block micronucleus assay

Metaphase chromosomes were harvested after one hour incubation with colcemid (0.1 µg/mL) (Life Technologies, California, USA). Cells were then treated with hypotonic solution (0.075M KCl) and fixed (methyl alcohol / glacial acetic acid, 3:1). Similarly, interphase cells were prepared from cell cultures to which colcemid had not been added. Cell suspensions were dropped on microscope slides and allowed to air-dry.

Binucleated cells were generated as previously described²⁹⁷. Briefly, proliferating cells were incubated with 6 µg/mL of cytochalasin B (Sigma-Aldrich, Missouri, USA) for 24 hours. The cells were then suspended in 0.075M KCl at 4°C and centrifuged at 300xg for 8 minutes. After centrifugation, a mixture of fixative (methanol:acetic acid, 3:1) and 37% formaldehyde was added to the cells. Finally, fixed binucleated cells were dropped on microscope slides and air-dried.

13.2. Fluorescence *in situ* hybridization and spectral karyotyping

Four color fluorescence *in situ* hybridization (FISH) was performed in cultured cells using contigs of overlapping bacterial artificial chromosome (BAC) clones that corresponded to four genes: *TERC* (3q26), *EGFR* (7p12), *CCND1* (11q13.3), and *CDX2* (13q12). Alternatively, a sequential bicolor FISH was performed in cytopsin samples using BAC clones that corresponded to six genes: *EGFR* (7p12), *MYC* (8q24.21), *CDX2* (13q12), *SMAD7* (18q21), *ZINF20* (20q13.2), and *APC* (5q21). All genes were used as surrogate markers for corresponding chromosome copy number. Nick-translation was used to label each contig with Dy505 (Dyomics, Jena, Germany), Biotin-dUTP (Roche Applied Science, Penzberg, Germany), Dy415 (Dyomics, Jena, Germany), Spectrum Orange-dUTP, Spectrum Red-dUTP or Spectrum Green-dUTP (Abbott Molecular, Illinois, USA). Yeast artificial chromosome (YAC) clones were used to generate centromere fluorescent probes for chromosomes 4 and 6 labeled with FITC and Cy3, respectively. Before hybridization, the samples were pretreated

with either 0.05% pepsin or 0.1M EDTA and 0.05% pepsin (Sigma-Aldrich, Missouri, USA). The slides were then dehydrated in an ethanol series (70%, 90%, and 100%) for 5 minutes each and allowed air-dry. The probes were denatured for 5 minutes and preannealed for 1 hour before being hybridized onto the samples. Coverslips were then added and sealed with rubber cement. The slides were placed into a humid hybridization chamber and incubated at 37°C overnight. After hybridization, slides were washed in 0.4x SSC / 0.3% NP-40 for 1 minute at 74°C and in 2x SSC / 0.1% NP-40 for 2 minutes at room temperature, both with gentle agitation. The slides were then dehydrated in 70%, 90%, and 100% ethanol for 2 minutes each, air-dried, and counterstained with Vectashield antifade mounting medium with DAPI (Vector Laboratories, California, USA). The slides were then imaged using the 40X/1.00 NA objective in a Leica DM-RXA fluorescence microscope (Leica, Wetzlar, Germany) and images of interphase FISH were captured on an automated slide scanning system Metafer (MetaSystems, Heidelberg, Germany) using an Olympus microscope with 40X/1.30 or 100X/1.25 NA objectives. For the tetraploidy study, five thousand cells were automatically captured and scored by three independent investigators; whereas, 200 cells were captured and scored *per* sample and lesion for the CRC evolution study.

Spectral karyotyping (SKY) was performed as previously described²⁹⁸. The protocol basis of SKY is the same as for the FISH procedure, but in this case the probes consisted of a cocktail of chromosome-specific painting probes prepared by amplification of flow-sorted chromosomes using degenerate oligonucleotide-primer (DOP)-PCR. SKY images were acquired with a 63X/1.40 NA objective using a custom-designed triple-pass filter using the SpectraCube SD200 (Applied Spectral Imaging, Edinger-Neckarhausen, Germany) connected to an epifluorescence microscope (DMRXA, Leica Microsystems, Wetzlar, Germany). For each clone, at least 15 metaphases spreads were analyzed with the SkyView 2.1 software (Applied Spectral Imaging, Edinger-Neckarhausen, Germany) and karyotypes defined using standard rules governing cytogenetic nomenclature²⁹⁹.

13.3. FISH data analysis

One cytospin *per* case was hybridized and evaluated with the above mentioned three different probe sets using single-cell FISH methods, as described previously²⁴⁶. All 23 progressed adenomas and six non-advanced adenomas were analyzed with a final count of

99 to 201 interphase nuclei for each sample. According to the data obtained, gain and loss patterns were assigned for each sample, sorted according to their frequency and displayed in color charts giving an overview of the clonal populations and the overall heterogeneity observed for every case. Moreover, the pattern of the exact copy number count for each of the six chromosomes analyzed was also registered for every sample.

In order to explore whether the adenocarcinomas are more heterogeneous than the adenomas we estimated tumor heterogeneity by calculating the instability index, defined as 100 times the number of cell count patterns divided by the number of cells.

13.4. Modeling of tumor progression and analysis of node depth

For each of the 23 progressed adenomas, the accumulation of copy number changes was modeled using FISHtress software³⁰⁰, which infers phylogenetic trees describing progression among the observed cell types as distinguished by their probe copy numbers. A tree is inferred to heuristically seek to minimize the total number of copy number changes across the tree. In this analysis, the phylogenetic trees were generated for the non-advanced adenomas and the AD and ADK regions of progressed adenomas separately. Moreover, a consensus tree was also generated from the combination of AD and ADK, reproducing the evolution of the whole progressed adenoma lesion. FISHtrees created a tree model in which the normal state (2, 2, 2, 2, 2, 2) is at the root of the inferred tree and each edge moving away from the root to a new node corresponds to the minimal number of changes in the copy number of genes required to connect the closest nodes. For each node, we also stored the number of cells observed to match the six components signal count pattern for that node. An example of a phylogenetic tree inferred with FISHtrees software is shown in **Figure 13**.

The number of steps away from the root is called the depth node, which is predictive of progression potential³⁰⁰. Therefore, to test whether the adenocarcinoma regions exhibited more progression than the adenoma region and non-advanced adenomas, we compared the average depths for all the sample categories.

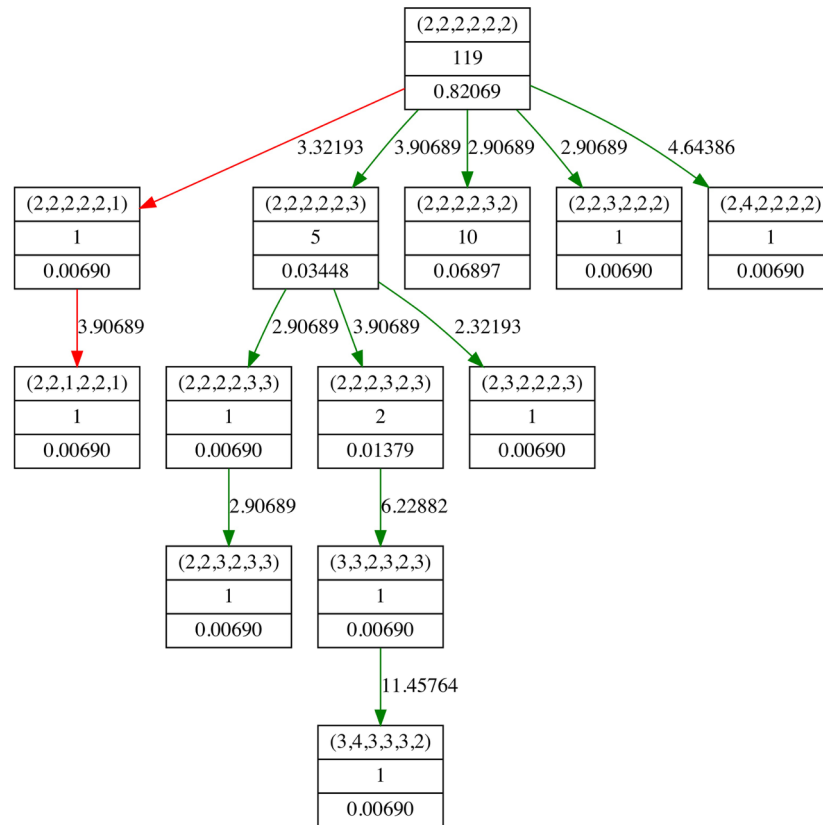


Figure 13. Phylogenetic tree generated by software FISHTrees. Every node is a clone with a certain signal count pattern, which is indicated in the first row of the node in the following order: 8q, 7p, 18q, 13q, 20q and 5q. The second and third rows of each node indicate the number of cells forming the clone and its frequency in the total population of the sample. The numbers next to the arrows connecting nodes are the value of the likelihood of the connection, considering the number of copy number changes between nodes.

By analyzing the consensus tree of all samples, the common ancestor for each lesion was determined, which was defined as the clone present in both the AD and ADK regions that leads to the appearance of the ADK population. Moreover, the drivers of the AD-to-ADK transition were inferred considering the CNAs that occur from the common ancestor to the first adenocarcinoma clones.

Finally, we defined fitness as a measure of the allelic expansion of a determined copy number status in the ADK respect to the AD. To calculate the fitness for each cell, each of the six loci analyzed were classified into five allelic options, monosomic, diploid, trisomic, tetrasomic, or polysomic. Then, for each locus the allelic frequency was calculated for the AD and the ADK of the same patient. The (logarithmic) ratio between the allelic frequency

between the ADK and the AD represents the fitness value associated to that allele. Then, the sum of individual locus represents the total fitness of the genotype or clone. The fitness value associated to a diploid allele was fixed to 0.

As an example with two loci, there are 10 cells in the AD carrying the genotype (8q,7p) = (2,3) and 10 cells carrying the genotype (2,2), while in the ADK there are 19 cells carrying the genotype (2,3) and only 1 carrying the genotype (2,2). Thus, in total there are two genotypes observed in the population, the genotypes (2,3) and (2,2). The genotype (2,2) has a fixed fitness of 0. The genotype (2,3) has a fitness of $0 + X(3)$. $X(3)$ is calculated in the following way:

$$X(3) = \frac{\frac{\text{number of AD cells carrying (3)}}{\text{total number of AD cells}}}{\frac{\text{number of ADK cells carrying (3)}}{\text{total number of ADK cells}}}$$

Then, 10 out of 20 cells have a value of 3 in the adenoma (allele frequency of 0.5) versus 19 out of 20 cells in the carcinoma (allele frequency of 0.95). The ratio of 0.95 over 0.5 is 1.9 and the log is 0.28. Therefore, the final fitness is $0 + 0.28 = 0.28$.

14. Flow cytometry

Cells were treated with a low concentration of aphidicolin (0.05 μ M) (Sigma-Aldrich, Missouri, USA) for 30 minutes to facilitate measurement of small differences, and with nocodazole (250 ng/mL) (Sigma-Aldrich, Missouri, USA) to arrest cells in mitosis. Previously BrdU-labeled cells were trypsinized, washed in ice-cold 1x PBS, and fixed with ice-cold 1x PBS diluted 1:10 in ethanol (70%) overnight at -20°C. Afterwards, cells were washed and treated with 2N HCl diluted in 1x PBS containing 0.1% (v/v) Triton X-100 for 15 minutes at room temperature. A Na₂B₄O₇·10H₂O buffer with a pH of 8.5 was used twice to neutralize the HCl solution. After that, cells were washed and blocked with 3% BSA (Sigma-Aldrich, Missouri, USA) in 1x PBS containing 0.05% (v/v) Tween-20 (Sigma-Aldrich, Missouri, USA) for one hour at room temperature prior to incubation with rat anti-BrdU (Abcam, Cambridge, UK) primary antibody (1:250 dilution) for one hour at room temperature. After washing, cells were incubated with secondary Alexa Fluor488-anti-rat conjugated antibody (ThermoFisher Scientific, Massachusetts, USA) for 45 minutes at room temperature, washed

with 1xPBS containing 0.05% (v/v) Tween (Sigma-Aldrich, Missouri, USA) and resuspended in 1x PBS containing 10 µg/mL propidium iodide and 0.1 mg/mL RNase A (Sigma-Aldrich, Missouri, USA), prior to the analysis by FACS (FACSCalibur, BD Biosciences, California, USA).

15. DNA ploidy measurement

Nuclear DNA ploidy status was assessed by image cytometry using Feulgen-stained cytopins of dissociated cells from all progressed adenoma samples. The staining procedure, internal standardization, and cell selection criteria were based on published methods³⁰¹. An average of 1,000 cells per sample was interactively selected in the ICM cell gallery and quantitatively measured for their DNA content. All DNA values were expressed in relation to the corresponding staining controls (lymphocytes), which were given the value 2c, denoting normal diploid DNA content. The DNA profiles were classified as previously described³⁰¹. Samples were assessed as diploid when <5% of cells showed DNA values >4.5c. Polyploidy was defined as >5% of cells presenting with DNA values >4.5c and the presence of cells with DNA values >6c.

16. Viability, growth and colony formation assays

CellTiter-Blue Cell Viability Assay (Promega, Wisconsin, USA) was used to determine the viability of individual DLD-1 clones. A total of 1,000 cells were seeded in 96-well plates and incubated for five days. CellTiter-Blue was added to the media for one hour and fluorescence signal was measured as *per* manufacturer's instructions using a SpectraMax M2 plate-reader (Molecular Devices, California, USA). To obtain growth curves, 70,000 cells of each clone were seeded, in triplicate, in 6-well plates and the growth evolution determined by Neubauer chamber counting every 24 hours over a five-day period. To assess for colony formation in both treated and untreated conditions, 200 cells of each clone were seeded in a 6-well plate. When required, cells were exposed to 10µM of the ATR inhibitor VE821 (Selleckchem, Texas, USA) for 24 hours before seeding them. Cells were then cultured for 14 days, with a change of media every two days, and finally stained with crystal violet (Sigma-Aldrich, Missouri, USA). Colonies were imaged on a Zeiss Axio Observer.A1 microscope (Carl Zeiss, Oberkochen, Germany) equipped with 5X/0.12 NA and 10X/0.25 NA

Plan-Apochromatic objectives. Digital images were captured using AxioVision software (Carl Zeiss, Oberkochen, Germany) and quantified using ImageJ software.

17. Migration and invasion assays

For the wound healing (scratch) assay, confluent cells in a 24-well plate were treated with 10 µg/mL of mytomicin (Sigma-Aldrich, Missouri, USA) for 1 hour at 37°C. A pipet tip was used to create a straight-line scratch. Migration was monitored by acquiring images every 24 hours until the cells displaying fastest migrating characteristics achieved a complete closure of the scratch.

For invasion, a Matrigel mixture (Corning, New York, USA) was mixed with the appropriate culture media to a concentration of 200 µg/mL and poured over the transwell membrane. A total of 1×10^6 cells were plated onto the Matrigel and 500 µl of extra media were added to the bottom well, then the chambers were incubated for 24 hours at 37°C. After the incubation, non-invasive cells were scraped off the chamber side of the transwell membrane and invasive cells were fixed with 100% methanol, washed twice with 1x PBS, and stained with 1% Giemsa solution. The samples were viewed on a Nikon Eclipse TS100 (Nikon Instruments Incorporated, Tokyo, Japan) and the total number of invasive cells was quantified in six random fields of view.

18. Live cell imaging

For phase-contrast live-cell imaging, DLD-1 2N and 4N cells were grown on sterile glass-bottom 35 mm Petri dishes. Petri dishes at 60–70% confluency were filled with L-15 medium supplemented with 4.5 g/l glucose and placed on the microscope stage within a stage top incubator (Tokai Hit, Shizuoka-ken, Japan) for temperature (36°C) and humidity control. Images were acquired on a Nikon Eclipse Ti inverted microscope (Nikon Instruments Incorporated, Tokyo, Japan) equipped with phase-contrast trans-illumination, transmitted light shutter, ProScan automated stage (Prior Scientific, Cambridge, UK), and a HQ2 CCD camera (Photometrics, Arizona, USA). 10-12 different fields of view were imaged every 6 minutes for 24 hours using a 20X / 1.4 NA Plan-Apochromatic phase-contrast objective and Nikon Perfect Focus (Nikon Instruments Incorporated, Tokyo, Japan) for focus control.

19. Statistical analysis

Chromoendoscopy and the consensus histological diagnosis were considered the gold standard for ACF detection and ACF confirmation diagnosis, respectively. The diagnostic yield of endoscopy was defined by the number of histologically diagnosed ACF divided by the number of ACF detected by chromoendoscopy.

SPSS statistical software (IBM Corp. 2012. IBM SPSS Statistics, Version 20.0. Armonk, NY) was used to analyze ACF data. Results for continuous variables were summarized using mean and standard deviation (SD) or median and interquartile range (IQR) for skewed data. Frequencies (%) were used to summarize categorical variables and 95% confidence intervals (95%CI) were calculated when relevant. Student's *t* or Mann-Whitney tests were used to compare the distribution of continuous variables by their outcome. Pearson chi-squared or Fisher's exact tests were used to test for any association between categorical variables and outcome. All analyses were exploratory and two-tailed tests with a significance level of 5% were used throughout. Comparisons of age- and gender-adjusted data were carried out using binary logistic regression. Paired analyses were performed when comparing molecular changes in normal mucosa and ACF. Inter-explorer and intra-explorer concordance for histological diagnosis was calculated by Weighted *k*-statistics and defined as: fair, 0.21-0.40; moderate, 0.41-0.60; good, 0.61-0.80 and very good, 0.81-1.00.

Distribution of LINE-1 methylation levels was assessed with the Shapiro-Wilk test showing that data was not normally distributed ($P = 0.017$). Accordingly, Wilcoxon signed-rank test was used for comparing average methylation levels of paired ACF and normal mucosa. When comparing global methylation between two independent groups, and more than two groups, Mann-Whitney U test and Kruskal-Wallis test were used, respectively. LINE-1 methylation levels are expressed as median and interquartile range (IQR).

For the CRC evolution and tetraploidy studies, Fisher's exact *t*-test for paired or unpaired samples, and proportions test (`prop.test`) were used to assess statistical significance using R packages, when appropriate, unless another method has been indicated. When more than three groups were compared, one-way ANOVA was performed.

RESULTS

- 1. Chapter I: Rectal aberrant crypt foci in humans are not surrogate markers for colorectal cancer risk**

As stated previously, the fact that colorectal carcinogenesis is a long stepwise process through a known precursor lesion provides a perfect scenario for prevention. The identification of reliable biomarkers that could help identify individuals at risk of CRC and measure the effect of chemopreventive treatments is of outmost importance. Despite ACF being promising biomarkers for CRC risk, discrepancies exist in the endoscopic, morphologic and molecular findings, which renders their clinical utility uncertain. Consequently, there is a need for well-standardized translational studies in order to define the role of ACF in colorectal carcinogenesis.

1.1. Patient baseline characteristics

Two hundred patients (56% female, 62.9 ± 13.8 years old) were included in our study's cohort. The control group consisted of 100 healthy individuals and had more young females than the case group (66% vs. 46%; 57.62 ± 15.58 years, $P = 0.05$ vs. 68.15 ± 9.26 years, $P < 0.001$, respectively). On the other hand, the case group included 100 patients with either adenomas or colon cancer (CC), as individuals with rectal cancer were excluded in order to ensure that the biopsies were taken at least 10 cm away from the neoplasm in all patients. Among the adenoma patients, 29 (58%) had advanced adenoma, while 22 CC patients (44%) had a previous CC and 28 (56%) had a current CC. Colonoscopy indications are described in **Table 6** and Inclusion and exclusion criteria are shown in **Figure 14**.

Table 6. Recommendations for colonoscopy.

Indications for colonoscopy (N = 308 patients)	
Abdominal pain:	27 patients
Anemia:	30 patients
Change in bowel habit:	37 patients
Rectal bleeding:	37 patients
Abnormal imaging or laboratory findings:	10 patients
Polyp surveillance:	26 patients
Colorectal cancer screening:	
- Positive immunologic fecal blood test (population screening program):	36 patients
- Family history:	46

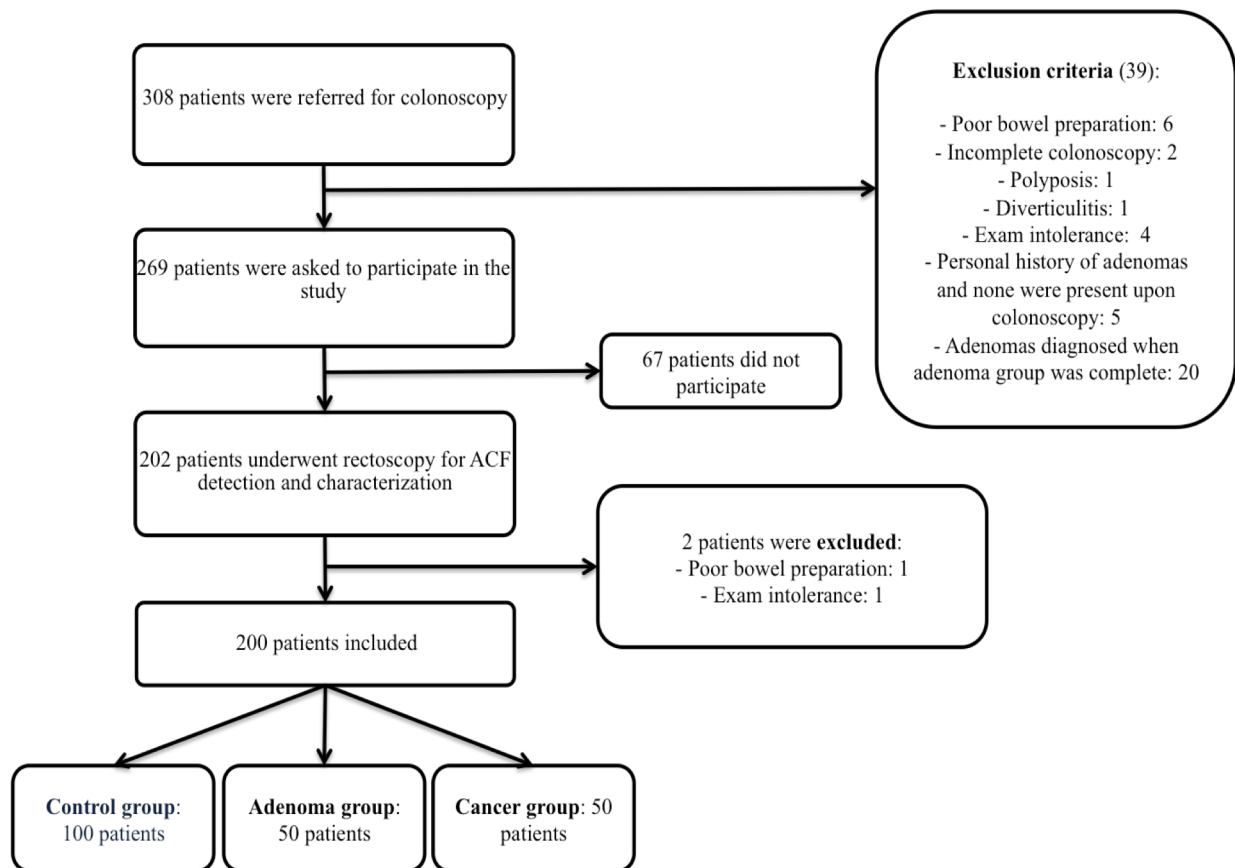


Figure 14. Patients flow-chart.

1.2. Endoscopic features of ACF as a surrogate marker for CRC risk

Methylene blue 0.5% high-definition chromoendoscopy (CE) was considered the gold standard for ACF detection and identified the presence of at least one ACF in 176/200 (88%) patients. Forty-three (21.5%) individuals had less than five ACF, 94 (47.0%) had 5-15, and 63 (31.5%) had more than 15. A total of 1,103 ACF were characterized by CE, whereas Narrow Band Imaging (NBI) was used to describe 768. Crypts size was assessed by CE, which determined that 305 (27.2%) ACF were small, 366 (33.2%) were medium, and 432 (39.2%) were large. Moreover, the shape of the crypt's lumens was semicircular-oval in 884 (80.1%) of the ACF, asteroid-like in 111 (10.1%) and irregular in 108 (9.8%). ACF vascular pattern intensity (VPI) was visualized by NBI and classified as weak in 525 (68.4%) ACF, normal in 220 (28.6%) ACF, and strong in only 23 (3%) ACF.

As it is shown in **Table 7 and 8**, only the presence of large ACF was related to the CRC risk group. In fact, the number of large ACF progressively increased toward CRC risk (control, 6.06%; adenoma, 20%; CC, 28%). Conversely, neither the number of ACF per subject, the lumen morphology nor VPI was predictive of CRC risk.

Table 7. Analysis of individuals and ACF features for CRC risk adjusted per patient's age and gender.

	Control N = 100	Case N = 100	P value	Exp (B)
Female (%)	66 (66.0)	46 (46.0)	0.005	0.32
Age (years $\pm$ standard deviation)	57.62 $\pm$ 15.58	68.15 $\pm$ 9.26	0.0001	1.08
Num. of ACF per patient	23/48/29	20/46/34	n.s	
Categories: <5/5-15/>15 (%)	(23.0/48.0/29.0)	(20.0/46.0/34.0)		
Patients with:				
• At least 1 large ACF (%)	62 (62.0)	78 (78.0)	0.046	2.03
• More than 1 large ACF (%)	41 (41.0)	66 (66.0)	0.002	2.70
• More than 4 large ACF (%)	6 (6.0)	24 (24.0)	0.003	3.47
• One irregular ACF (%)	26 (26.0)	35 (35.0)	n.s	
• At least one dysplastic ACF (%)	32 (32.0)	40 (40.0)	n.s	
• More than one dysplastic ACF (%)	10 (10.0)	11 (11.0)	n.s	
• At least one serrated ACF (%)	19 (19.0)	26 (26.0)	n.s	

Table 8. Analysis of ACF features predictive of CRC risk adjusted per patient's age and gender.

	Control N = 100	Case N = 100	P value	Exp (B)
Large Size (Large ACF/ ACF evaluated by chromoendoscopy)	32.1% (159/496)	45.0% (273/607)	0.001	1.61
Irregular shape (Irregular ACF/ ACF evaluated by chromoendoscopy)	7.4% (35/470)	11.9% (65/547)	n.s	
Large or irregular (Large or irregular ACF/ ACF evaluated by chromoendoscopy)	36.8% (173/470)	54.1% (296/547)	0.0001	1.76
Intensive vascular pattern (Hypervascular ACF/ ACF evaluated by NBI)	2.2% (8/369)	3.8% (15/399)	n.s	
ACF histologically confirmed (ACF histologically confirmed/ACF biopsied)	75.6% (235/311)	84.8% (318/375)	0.018	1.69
Dysplastic ACF (Dysplastic ACF/histologically confirmed ACF)	19.6% (46/235)	17.3% (55/318)	n.s	
Serrated ACF (Serrated ACF/histologically confirmed ACF)	10.2% (24/235)	10.1% (32/318)	n.s	
Hyperplastic ACF (Hyperplastic ACF/histologically confirmed ACF)	70.2% (165/235)	72.6% (231/318)	n.s	

1.3. ACF histology as a surrogate marker for CRC risk

In order to reach a histological consensus diagnosis, two expert pathologists (A and B) evaluated all ACF samples two times over a period of six months, and all discordant diagnosis were then discussed. Both pathologists were blinded to the endoscopic classification and each other's diagnosis. Accordingly, 686 ACF samples were classified as follows: inadequate samples (5.5%), normal mucosa (13.8%), hyperplastic ACF (57.8%), serrated ACF (8.2%) or dysplastic ACF (14.7%). The diagnostic yield from the endoscopy was 80.6% and the prevalence of dysplasia was 18.3% (**Table 9**).

Table 9. ACF samples description.

Number of ACF described with chromoendoscopy	1,103
Number of ACF described with NBI	768
Number of ACF biopsied	686
Non-valid biopsies:	133
- Inadequate	38
- Normal mucosa	95
Histology confirmed ACF:	553
- Hyperplastic (%)	396 (71.6)
- Serrated (%)	56 (10.1)
- Dysplastic (%)	101 (18.3)
Chromoendoscopy diagnosis yield (%):	80.6
Histologically confirmed ACF/ACF detected with chromoendoscopy	553/686

Pathologist A showed a weighed K for intra-observer concordance of 0.59, and pathologist B 0.71 ($P \leq 0.001$). However, the best weighed K for inter-observer concordance was only of 0.25 ($P < 0.001$). Interestingly, the histology of ACF was not related to gender, sex or CRC risk group (**Table 10**). Moreover, although dysplastic and serrated ACF exhibited an irregular shape more frequently than hyperplastic ACF, their size and VPI were not associated with ACF histology (**Table 10**).

Table 10. Analysis of ACF features related to histology.

	Hyperplastic	Serrated	Dysplastic	P value
Large Size (Large ACF/histologically confirmed ACF)	59.6% (236/396)	50.5% (33/56)	58.9% (51/101)	n.s
Irregular shape (Irregular ACF/histologically confirmed ACF)	10.1% (39/385)	20.4% (11/54)	20.6% (20/97)	0.006
Asteroid-like or irregular shape (Asteroid-like or irregular ACF/histologically confirmed ACF)	20.8% (80/385)	25.9% (14/54)	35% (34/97)	0.012
Intensive vascular pattern (Hypervascular ACF/histologically confirmed ACF)	3.9% (12/306)	5.1% (2/39)	5.6% (4/71)	n.s
Control group (ACF from control group/histologically confirmed ACF)	41.7% (165/396)	42.9% (24/56)	45.5% (46/101)	n.s
Case group (ACF from study group/ histologically confirmed ACF)	58.3% (231/396)	57.1% (32/56)	54.5% (55/101)	n.s

1.4. Molecular analysis of ACF

The first three histologically confirmed ACF for each patient were then used for molecular analysis, accounting a total of 294 ACF samples. Of these, 128 were from 67 healthy control individuals (47 Female, (70%); 58.3 ± 15.2 years old) and 166 were from 81 case patients, of which 40 were from adenoma and 41 from CC patients (32 Female, (39%); 67.7 ± 9.9 years old). Histological examination revealed that 197 (67.0%) ACF were hyperplastic, 35 (12.0%) serrated, and 62 (21.0%) dysplastic lesions (**Figure 15**).

1.4.1. APC mutation and microsatellite instability analysis

APC mutations were assessed in 285/294 (96.9%) ACF samples by sequencing two PCR fragments spanning 82.6% of APC mutations in terms of frequency. The APC variant E1317Q was found in 5/285 (1.7%) samples (**Table 11**); however, these ACF belonged to patients who exhibited the same variant in their normal mucosa, indicating it was a germline alteration not responsible for ACF formation. On the other hand, MSI status was evaluated for 276/294 (94.2%) ACF samples; although all samples analyzed were MSS.

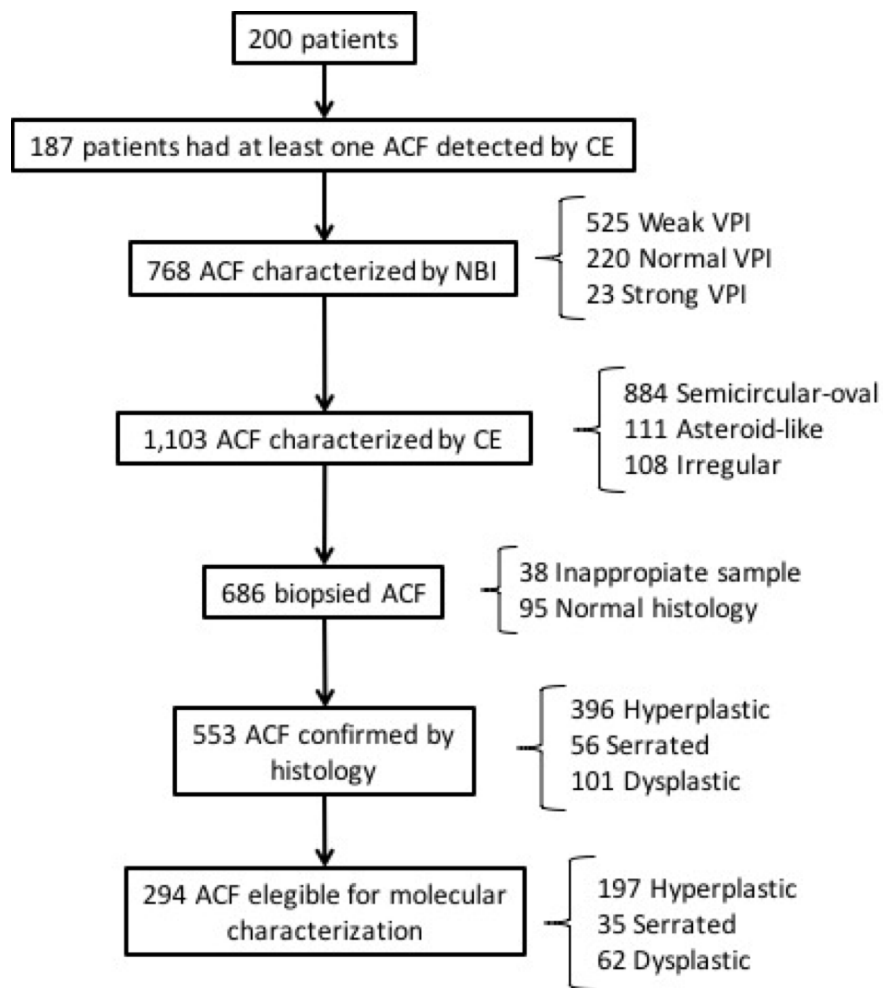


Figure 15. ACF samples flow-chart. ACF, aberrant crypt foci; CE, chromoendoscopy; NBI, narrow band imaging; and VPI, vascular pattern intensity.

Table 11. ACF molecular features related to CRC risk group.

	Control	Case	P value
<i>KRAS</i> mutations	17/124 (13.7%)	30/165 (18.2%)	n.s
<i>BRAF</i> mutations	3/128 (2.3%)	3/162 (1.9%)	n.s
<i>APC</i> mutations	0/127 (0%)	5/158 (3.2%)	n.s
<i>MGMT</i> hypermethylation	6/37 (16.2%)	6/71 (8.5%)	n.s

1.4.2. *BRAF* and *KRAS* mutations

BRAF and *KRAS* mutations were examined for 290/294 (98.6%) and 289/294 (98.3%) ACF samples, respectively. The *BRAF* V600E mutation was assessed by Real Time PCR of allelic discrimination and was found in only 6/290 (2.1%) of the ACF samples, whereas *KRAS* was mutated in 48/289 (16.3%) ACF samples as determined by COLD-PCR and Sanger sequencing. As previously described, none of the ACF samples simultaneously exhibited *BRAF* and *KRAS* mutations³⁰². *KRAS* and *BRAF* mutations were somatic, as they were not detected in the corresponding normal mucosa. Furthermore, we did not observe any positive correlation between *BRAF* or *KRAS* mutational status and the case or histology groups (Table 11 and 12).

Table 12. ACF molecular features related to histology.

	Hyperplastic	Serrated	Dysplastic	P value
<i>KRAS</i> mutations	33/195 (16.9%)	2/35 (5.7%)	12/59 (20.3%)	n.s
<i>BRAF</i> mutations	3/193 (1.6%)	1/35 (2.9%)	2/62 (3.2%)	n.s
<i>APC</i> mutations	2/190 (1.1%)	1/35 (2.9%)	2/60 (3.3%)	n.s
<i>MGMT</i> hypermethylation	10/76 (13.2%)	0/9 (0%)	2/23 (8.7%)	n.s

1.4.3. *MGMT* methylation status analysis

As *KRAS* mutations have been previously associated with the serrated pathway of colorectal carcinogenesis³⁰³, the methylation status of the *MGMT* gene was determined in 33 *KRAS* mutated and 75 *KRAS* wild-type ACF samples, as well as in their corresponding normal mucosa. Overall, the level of methylation was significantly higher in the ACF samples compared to their corresponding normal mucosa (5.54 ± 3.70 vs 4.23 ± 2.31 , $P = 0.004$). Nonetheless, there were no differences in methylation levels among the control and case groups or between those ACF samples with mutated or wild-type *KRAS* status (Figure 16A, B and C). We then analyzed the methylation results as a categorical variable using the mean levels in the normal mucosa methylation as the cut-off for ACF hypermethylation (mean $\pm$ 2 standard deviations, which corresponded to 8.66%). By using this cut-off we found significantly more hypermethylated samples in the set of *KRAS* mutated ACF than in the set of ACF displaying wild-type *KRAS* (21% vs 5.71%, $P = 0.02$), thus confirming the association between these two molecular alterations. Nevertheless, we did not detect any difference in

the distribution of the hypermethylated ACF among the control and case groups, or according to ACF histology (Table 11 and 12).

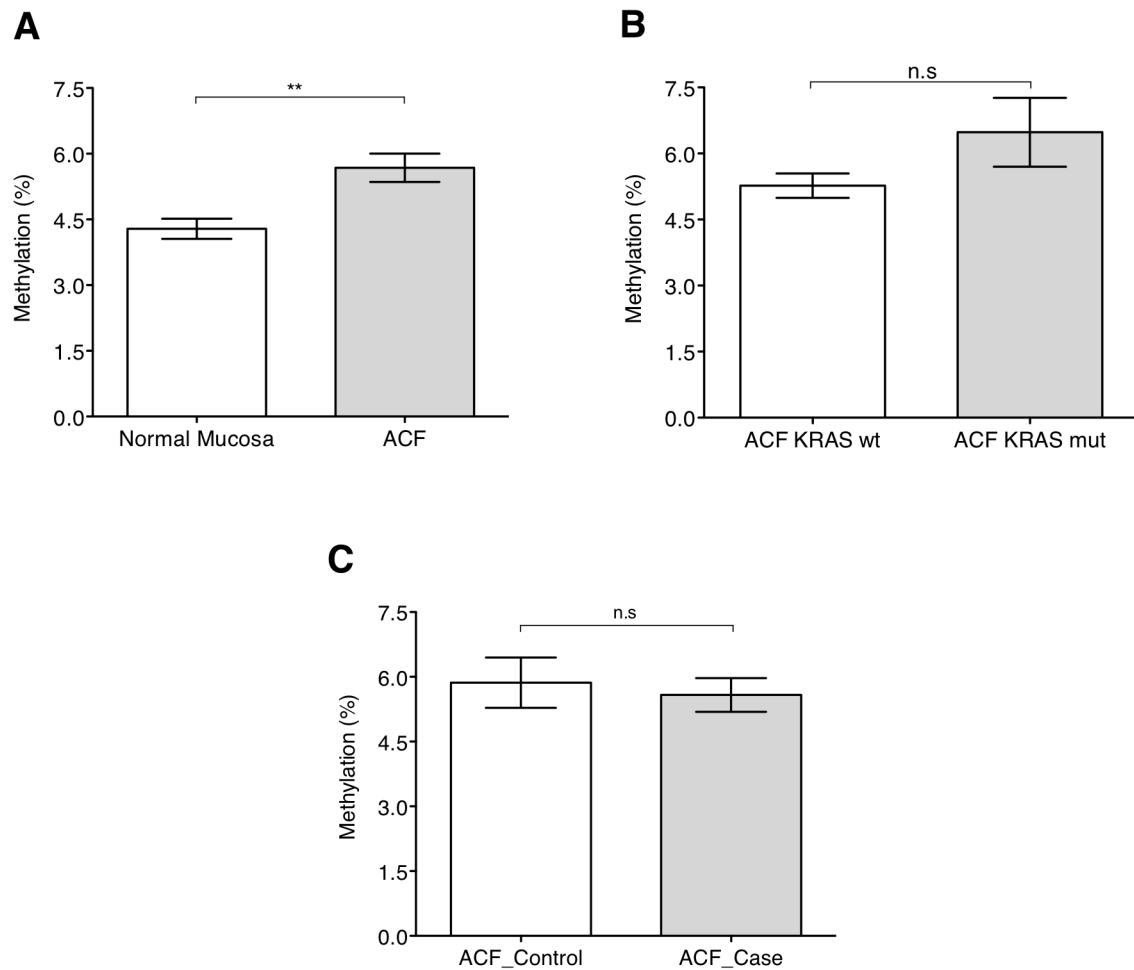


Figure 16. *MGMT* methylation levels in ACF samples. **(A)** *MGMT* methylation levels of normal mucosa vs. ACF samples. **(B)** *MGMT* methylation of ACF with *KRAS* mutated or *wild-type*. **(C)** *MGMT* methylation levels in ACF from control and case patients. Data are reported as mean $\pm$ SEM. ** $P < 0.001$.

- 2. Chapter II: LINE-1 hypomethylation is neither present in rectal aberrant crypt foci nor associated with field defect in sporadic colorectal neoplasia**

Global DNA hypomethylation is an early event in colorectal carcinogenesis, thus its evaluation is also relevant when investigating precancerous lesions. In this sense, long interspersed nuclear element-1 (LINE-1) methylation status is a well-known marker for genome-wide DNA methylation, which has been reported as a field defect for CRC¹¹⁶. However, LINE-1 methylation has never been assessed in ACF samples and more evidence is needed to depict its role as field defect.

2.1. Patients baseline characteristics

For this molecular study, a cohort of rectal ACF and paired normal descending (NDM) and rectal (NRM) mucosa obtained from the first study was used. Specifically, this sample cohort comprised 70 healthy control, 40 adenoma and 42 CC patients (**Figure 17**).

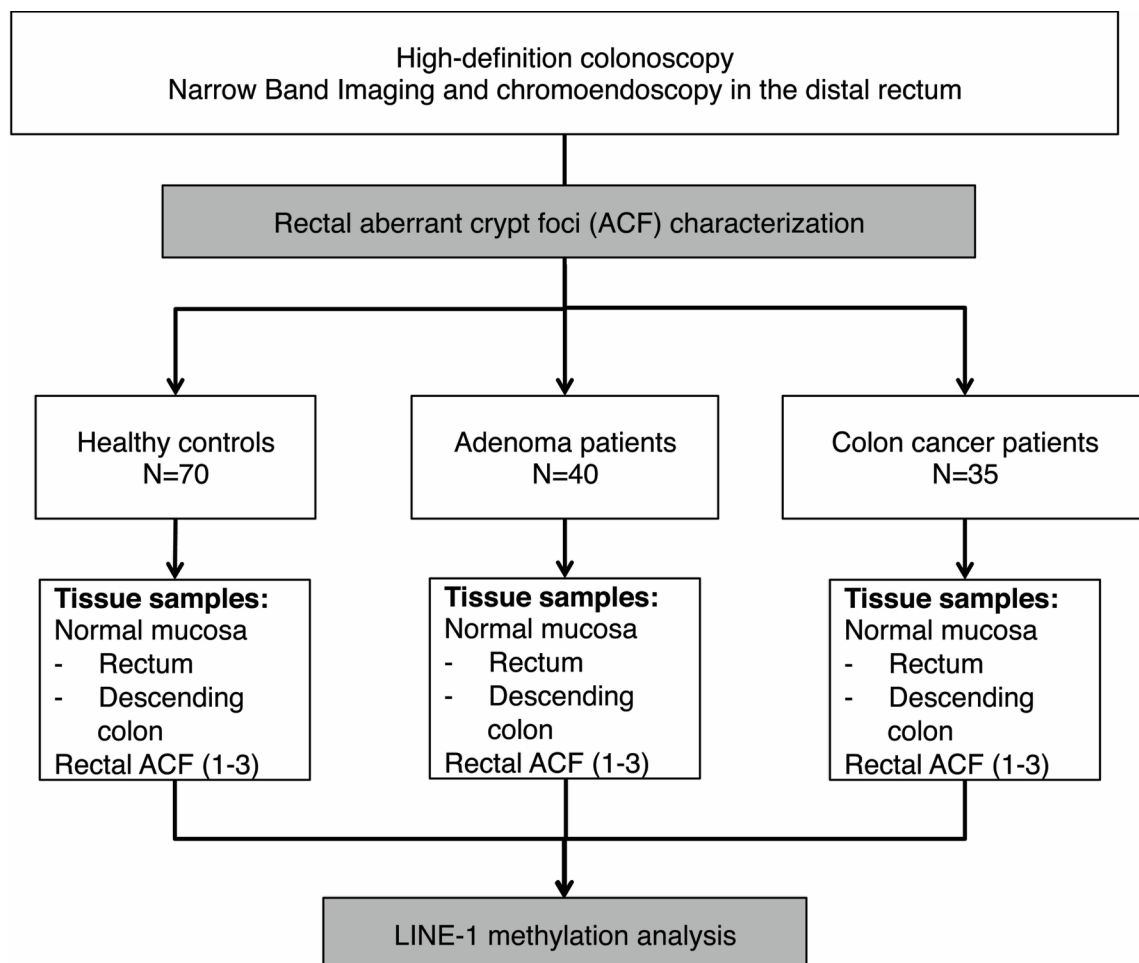


Figure 17. Flow chart summarizing the study design. LINE-1 methylation analysis was performed in DNA extracted from ACF and normal colorectal mucosa samples from three colon cancer risk groups (individuals with a normal colonoscopy, patients with adenomas, and patients with colon cancer).

Baseline characteristics of these patients and their rectal ACF related to CRC risk are detailed in **Table 13**. As it is shown, the adenoma group mainly included patients with advanced colorectal adenomas (mean size: 12.3 ± 9.8 mm) located throughout the colorectum. Moreover, although 16 (40%) adenoma patients had synchronous serrated polyps, the vast majority of them were hyperplastic polyps located in the recto-sigmoid. On the other hand, 12 CC patients had synchronous adenomas, while eight exhibited synchronous serrated polyps, all of them hyperplastic.

Table 13. Baseline characteristics of individuals and ACF in each colon cancer risk group.

	Control n=70	Adenoma n=42	Colon Cancer n=35
Gender, n (%)			
- Female	48 (68.6%)	17 (40.5%)	13 (37.1%)
- Male	22 (31.4%)	25 (59.5%)	22 (62.9%)
Age, mean (years) $\pm$ SD	58.1 $\pm$ 15.3	68.6 $\pm$ 9.9	66.8 $\pm$ 8.6
Colon cancer location, n (%)			
- Sigmoid	-	-	21 (60%)
- Descending	-	-	2 (5.7%)
- Transverse	-	-	1 (2.9%)
- Ascending	-	-	9 (25.7%)
- Cecum	-	-	2 (5.7%)
Colon cancer TNM stage, n (%)			
- I	-	-	10 (28.6%)
- II	-	-	11 (31.4%)
- III	-	-	13 (37.1%)
- IV	-	-	1 (2.9%)
Adenoma features (per patient)¹			
- Tubular, n (%)	-	28 (70%)	11 (31.4%)
- Tubulovillous, n (%)	-	7 (17.5%)	1 (2.9%)
- Villous, n (%)	-	5 (12.5%)	0 (0%)
- High-grade dysplasia, n (%)	-	11 (27.5%)	5 (14.3%)
- Adenoma size, mean (mm) $\pm$ SD	-	12.3 $\pm$ 9.8	8.4 $\pm$ 5
- Number of adenomas per patient, mean $\pm$ SD (range)	-	2.4 $\pm$ 2 (1-9)	1.2 $\pm$ 1.2 (1-4)
- Advanced adenoma ² , n (%)	-	22 (55%)	7 (20%)
Adenoma location (per patient)³, n (%)			
- Rectum	-	12 (30%)	2 (5.7%)
- Sigmoid	-	10 (25%)	3 (8.6%)
- Descending	-	5 (12.5%)	3 (8.6%)
- Transverse	-	5 (12.5%)	2 (5.7%)
- Ascending	-	5 (12.5%)	2 (5.7%)
- Cecum	-	3 (7.5%)	0 (0%)

Table 13. Baseline characteristics of individuals and ACF in each colon cancer risk group (Continued).

	Control N = 70	Adenoma N = 42	Colon Cancer N = 35
Serrated polyps features (per patient)			
- Presence of synchronous serrated polyps ⁴ , n (%)		16 (40%)	8 (22.9%)
- Recto-sigmoid hyperplastic polyps, n (%)		12 (30%)	3 (8.6%)
- Proximal ⁵ serrated polyps, n (%)		4 (10%)	5 (14.3%)
- Serrated polyp ≥ 10 mm, n (%)		2 (5%)	2 (5.7%)
- Number of serrated polyps per patient, mean $\pm$ SD (range)		1.6 $\pm$ 0.8 (1-3)	3.1 $\pm$ 3.4 (1-5)
- Hyperplastic polyp, n (%)		15 (37.5%)	8 (22.9%)
- Sessile serrated adenoma, n (%)		1 (%)	0 (0%)
- Traditional serrated adenoma, n (%)		0 (%)	0 (0%)
Number of ACF for molecular study	117	76	77
Dysplastic ACF (%)	24 (17.4%)	23 (24%)	17 (25%)
Hyperplastic ACF (%)	114(82.6%)	73(76%)	51(75%)

SD: standard deviation; ACF: aberrant crypt foci.

¹ In the adenoma group, referred to 40 individuals with full pathological information of adenomas available.

² Advanced adenoma: ≥ 10 mm in size, or presence of high-grade dysplasia or villous features.

³ In case of more than one adenoma, referred to location of the most advanced adenoma.

⁴ Includes hyperplastic polyps, sessile serrated adenomas and traditional serrated adenomas.

⁵ Proximal to the sigmoid colon.

2.2. Rectal ACF do not display LINE-1 hypomethylation

To further test the role of rectal ACF as a potential precursor lesion of CRC, we analyzed the LINE-1 methylation levels in ACF and paired NRM and NDM from all patients. The median LINE-1 methylation levels of up to three ACF for each patient were considered for statistics using the Wilcoxon signed-rank test, as the data was not normally distributed (see Materials and Methods). As it is shown in **Figure 18A**, our data indicated that LINE-1 methylation levels in ACF were higher than in the NRM [80.02% (77.54-82.22) vs. 76.13% (73.74-78.06); $P < 0.0001$]. In fact, only 20 subjects exhibited lower LINE-1 methylation in their rectal ACF compared to their NRM, and LINE-1 was not hypomethylated in any of the three risk groups (**Figure 18B**).

Likewise, comparable results were obtained when rectal ACF LINE-1 methylation levels were compared to those of the NDM [80.02% (77.54-82.22) vs. 78.42% (76.51-80.33); $P = 0.0001$] (**Figure 19A and B**). We also stratified these analyses by sex and found similar results for both males and females (**Table 14**). Interestingly, there were significant differences in LINE-1 methylation between dysplastic and hyperplastic ACF [81.46% (78.75-82.97) vs. 79.54 (76.88-81.79); $P = 0.031$].

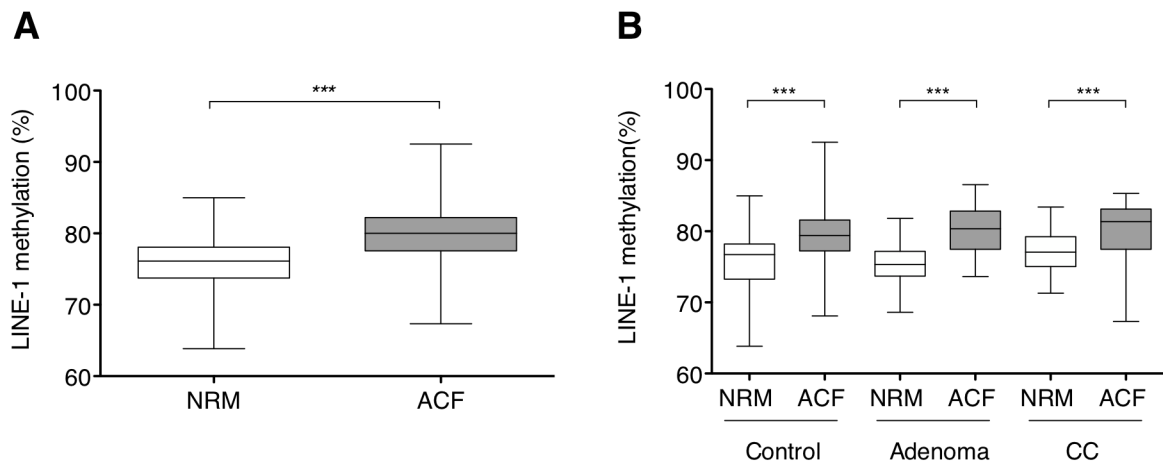


Figure 18. LINE-1 methylation status in aberrant crypt foci (ACF) and normal rectal mucosa (NRM). **(A)** LINE-1 methylation levels in ACF compared to NRM samples. **(B)** ACF LINE-1 methylation levels compared to NRM samples according to risk group. Box-and-whisker plot indicating the median methylation level expressed as a percentage (horizontal line), 25th and 75th percentile (box), and maximum and minimum levels (whiskers). *** $P < 0.0001$.

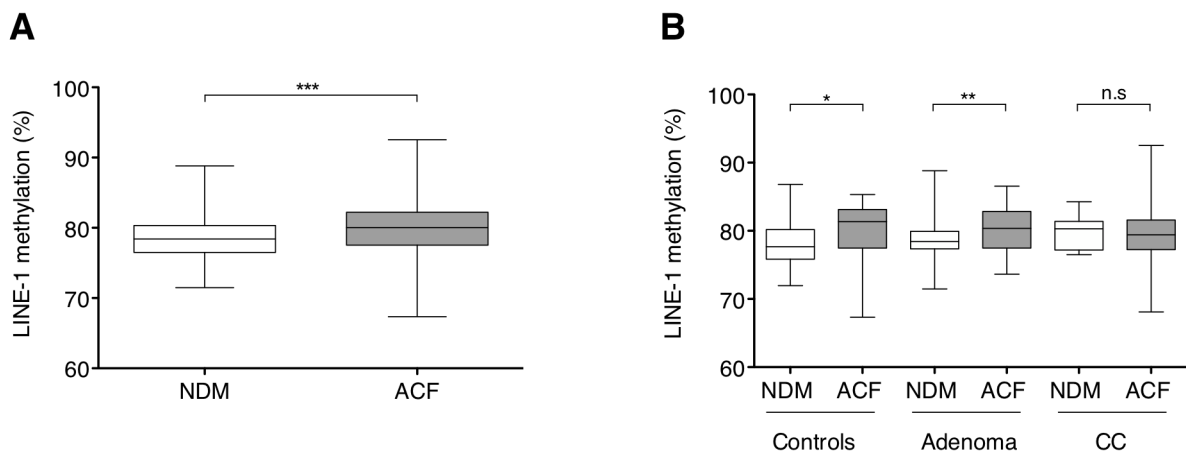


Figure 19. LINE-1 methylation status in aberrant crypt foci (ACF) and normal descending mucosa (NDM). **(A)** LINE-1 methylation levels in ACF compared to NDM. **(B)** ACF LINE-1 methylation levels compared to NDM samples according to risk group. Box-and-whisker plot indicating the median methylation level expressed as a percentage (horizontal line), 25th and 75th percentile (box), and maximum and minimum levels (whiskers). * $P < 0.05$; ** $P < 0.001$; *** $P < 0.0001$.

Table 14. LINE-1 methylation analysis stratified per gender.

	Gender	N	Median (IQR)	P value
LINE-1 methylation in normal rectal mucosa versus ACF*				
- Overall	Male	53	76.76 (74.25-78.55) vs 80.32 (77.51-82.39)	$P < 0.0001$
	Female	64	75.69 (73.33-77.95) vs 79.72 (77.08-82.00)	$P < 0.0001$
- Healthy controls	Male	19	76.95 (73.50-80.02) vs 80.19 (77.00-81.55)	$P = 0.021$
	Female	43	76.11 (72.74-78.00) vs 79.33 (77.59-81.78)	$P < 0.0001$
- Adenoma	Male	18	74.87 (73.10-77.21) vs 79.93 (77.37-82.36)	$P = 0.0005$
	Female	13	74.40 (73.72-77.41) vs 80.79 (77.32-83.91)	$P = 0.011$
- Colon cancer	Male	16	77.60 (76.02-79.80) vs 81.73 (77.51-84.51)	$P = 0.003$
	Female	8	75.63 (73.27-79.91) vs 80.07 (74.33-82.34)	$P = 0.156$
LINE-1 methylation in normal descending mucosa versus ACF*				
- Overall	Male	49	78.44 (76.42-80.66) vs 80.32 (77.51-82.39)	$P = 0.103$
	Female	54	78.07 (76.34-80.28) vs 79.72 (77.08-82.01)	$P = 0.032$
- Healthy controls	Male	11	78.90 (77.13-81.23) vs 80.19 (77.00-81.55)	$P = 0.320$
	Female	31	77.64 (75.33-79.66) vs 79.33 (77.59-81.78)	$P = 0.016$
- Adenoma	Male	22	78.44 (76.52-80.72) vs 79.93 (77.37-82.36)	$P = 0.614$
	Female	14	79.13 (77.49-80.69) vs 80.79 (77.32-83.91)	$P = 0.454$
- Colon cancer	Male	15	78.40 (75.87-79.41) vs 81.73 (77.51-84.51)	$P = 0.092$
	Female	8	78.19 (76.72-80.15) vs 80.07 (74.33-82.34)	$P = 0.844$
LINE-1 methylation according to ACF histology				
- Hyperplastic ACF vs. dysplastic ACF	Male	47 vs. 15	79.66 (77.06-82.01) vs 81.54 (79.15-83.07)	$P = 0.094$
	Female	54 vs. 18	79.33 (76.30-81.60) vs 81.02 (78.70-82.52)	$P = 0.145$
LINE-1 methylation in normal rectal mucosa				
- Healthy controls vs. Adenoma vs. CC	Male	19 vs. 18 vs. 16	76.95 (73.50-80.02) vs 74.87 (73.10-77.20) vs 77.60 (76.02-79.80)	$P = 0.051$
	Female	43 vs. 13 vs. 8	76.11 (72.74-78.00) vs 75.40 (73.72-77.41) vs 75.63 (73.27-79.91)	$P = 0.840$
LINE-1 methylation in normal descending mucosa				
- Healthy controls vs. Adenoma vs. CC	Male	11 vs. 23 vs. 15	78.90 (77.13-81.23) vs 78.44 (76.52-80.72) vs 78.40 (75.87-79.41)	$P = 0.692$
	Female	31 vs. 15 vs. 8	77.64 (75.33-79.66) vs 79.13 (77.49-80.69) vs 78.19 (76.72-80.15)	$P = 0.201$
LINE-1 methylation in normal mucosa from different colonic segments				
- Normal rectal mucosa vs. normal descending mucosa	Male	53 vs. 49	76.76 (74.25-78.55) vs 78.44 (76.42-80.66)	$P = 0.004$
	Female	64 vs. 54	75.69 (73.33-77.95) vs 78.07 (76.34-80.28)	$P = 0.0002$
LINE-1 methylation in ACF among CC risk groups				
- ACF Healthy controls vs. adenoma vs. CC	Male	22 vs. 22 vs. 18	80.19 (77.00-81.55) vs 79.93 (77.37-82.36) vs 81.73 (77.51-84.51)	$P = 0.362$
	Female	46 vs. 46 vs. 10	79.33 (77.59-81.78) vs 80.79 (77.32-83.91) vs 80.07 (74.33-82.34)	$P = 0.667$
LINE-1 methylation in ACF stratified by gender				
- Healthy controls (Male vs. Female)		22 vs. 46	80.19 (77.00-81.55) vs 79.33 (77.59-81.78)	$P = 0.969$
- Adenoma (Male vs. Female)		22 vs. 16	79.93 (77.37-82.36) vs 80.79 (77.32-83.01)	$P = 0.756$
- Colon cancer (Male vs. Female)		18 vs. 10	81.73 (77.51-84.51) vs 80.07 (74.33-82.34)	$P = 0.240$

IQR: interquartile range; ACF, aberrant crypt foci; CC, colon cancer; NRM, normal rectal mucosa; and NDM, normal descending mucosa. *Paired-samples analysis.

2.3. LINE-1 methylation levels in normal colorectal mucosa are similar in all CRC risk groups

In order to evaluate the LINE-1 field defect, we compared the methylation status of LINE-1 in NRM and NDM of individuals from the three CRC risk groups (controls vs. adenoma vs. CC). Since the control group was significantly younger than both the adenoma and CC groups (**Table 13**), we excluded an age-effect in the LINE-1 methylation levels (**Figure 20**).

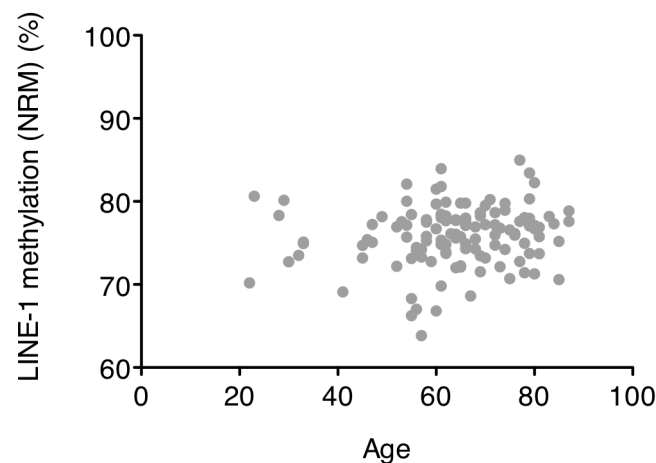


Figure 20. Evaluation of the age-effect for LINE-1 methylation data. Representation of the correlation between LINE-1 methylation levels in normal rectal mucosa (NRM) and the age of patients.

Next, as it is shown in **Figure 21**, we found the methylation of LINE-1 in NRM to be independent of the patient group [controls: 76.73% (73.23-78.22); adenoma: 75.34% (73.7-77.2); CC: 77.06% (75.04-79.23); $P = 0.107$] (**Table 15**). Moreover, we obtained similar results when we considered the NDM [controls: 77.67% (75.84-80.18); adenoma: 78.45% (77.37-79.92); CC: 80.30% (77.19-81.39); $P = 0.147$] (**Table 15**), as well as when we stratified according to sex (**Table 14**).

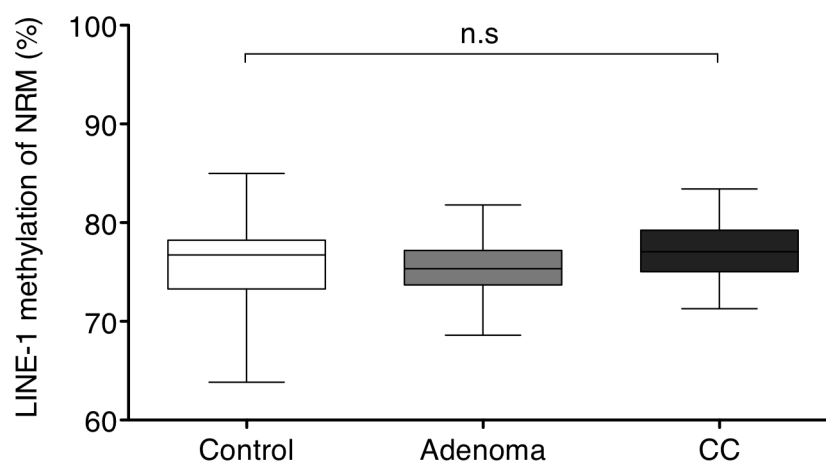


Figure 21. Evaluation of LINE-1 methylation in normal colorectal mucosa as a potential field defect. LINE-1 methylation levels in normal rectal mucosa (NRM) according to colon cancer (CC) risk group. Box-and-whisker plot indicating the median methylation level expressed as a percentage (horizontal line), 25th and 75th percentile (box), and maximum and minimum levels (whiskers).

Table 15. LINE-1 methylation levels in normal mucosa and ACF for each of the 3 groups of subjects.

	Controls N=70	Adenoma N=42	Colon Cancer N=35	<i>P</i> value
Normal rectal mucosa, n/N	62/70	31/42	24/35	
Median	76.73%	75.34%	77.06%	0.107
(IQR)	(73.23-78.22)	(73.7-77.2)	(75.04-79.23)	
Normal descending colon mucosa, n/N	42/70	38/42	21/35	
Median	77.67%	78.45%	80.30%	0.147
(IQR)	(75.84-80.18)	(77.37-79.92)	(77.19-81.39)	
Aberrant crypt foci, n/N	68/70	38/42	24/35	
Median	79.41%	80.32%	81.37%	0.409
(IQR)	(77.24-81.58)	(77.45-82.82)	(77.48-83.14)	
Dysplastic aberrant crypt foci, n/N	15/70	11/42	9/35	
Median	80.58%	80.99%	82.74%	0.320
(IQR)	(78.54-82.37)	(79.15-83.07)	(79.83-83.43)	
Hyperplastic aberrant crypt foci, n/N	53/70	27/42	24/35	
Median	79.32%	79.59%	80.02%	0.812
(IQR)	(76.94-81.54)	(76.80-82.22)	(76.50-82.55)	

IQR, interquartile range; n, number of samples analyzed in each group; and N, total number of samples included in each group.

The potential field defect for LINE-1 methylation was also analyzed taking into consideration the distance of the lesion to the normal mucosa biopsies. However, the levels of LINE-1 methylation in normal mucosa did not differ according to their proximity to the adenoma or CC (**Table 16**). Thus, altogether these results suggest that LINE-1 hypomethylation does not represent a field defect for CRC.

Table 16. LINE-1 methylation levels in normal mucosa stratifying by location of the neoplastic lesion.

LINE-1 methylation	Patients with proximal ¹ adenoma (N=10)	Patients with distal ¹ adenoma (N=30)	P value
Normal rectal mucosa (NRM)	73.87% (72.47-75.90)	76.06% (74.78-77.33)	0.087
Normal descending mucosa (NDM)	80.72% (76.62-82.68)	78.38% (76.55-80.01)	0.177
	Patients with proximal ¹ colon cancer (N=11)	Patients with distal ¹ colon cancer (N=24)	P value
Normal rectal mucosa (NRM)	77.59% (72.78-80.12)	77.40% (75.70-79.79)	0.794
Normal descending mucosa (NDM)	77.14% (75.59-80.06)	78.59% (76.90-79.41)	0.458

¹ Referred to the splenic flexure.

2.4. LINE-1 methylation levels in normal mucosa differ between colonic segments

Finally, when we compared the LINE-1 methylation status in the normal mucosa from different segments of the colorectum, we found a higher degree of hypomethylation in the rectum compared to the descending colon [76.13% (73.75-78.06) vs. 78.42% (76.51-80.33); $P < 0.0001$] (**Figure 22**), which was also independent of gender (**Table 14**). Therefore, this finding suggests different genome-wide methylation profiles among colorectum segments.

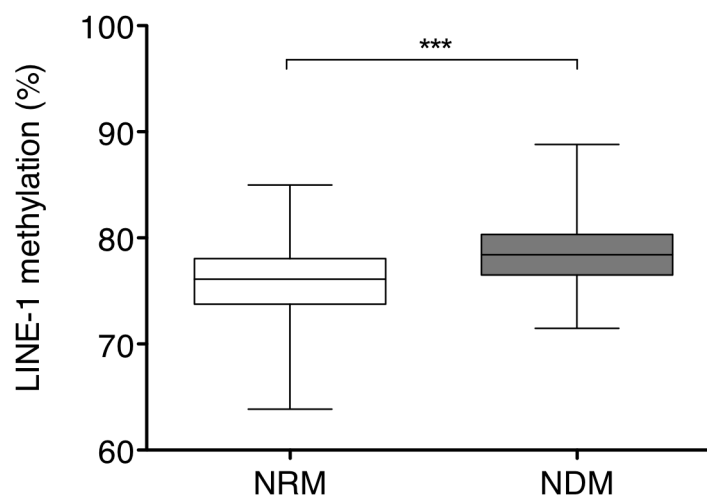


Figure 22. LINE-1 methylation levels in normal rectal mucosa (NRM) compared to normal descending mucosa (NDM). Box-and-whisker plot indicating the median methylation level expressed as a percentage (horizontal line), 25th and 75th percentile (box), and maximum and minimum levels (whiskers). *** $P < 0.0001$

3. Chapter III: The role of copy number alterations in the adenoma-to-adenocarcinoma transition of colorectal cancer

Together with preventive strategies, another decisive step in the investigation of colorectal carcinogenesis is to understand the transition from adenoma to adenocarcinoma. In this respect, the onset of intratumor heterogeneity and the increasing main role of chromosome abnormalities during malignant transformation have challenged the classical idea of CRC development. Indeed, while CNAs are important players in CRC progression, their role in the adenoma-to-adenocarcinoma transition has yet to be fully elucidated and more comprehensive studies are necessary in order to define the landscape of this transition. Accordingly, we used a single-cell based strategy to evaluate the copy number changes involved in this malignant transformation using adenomas with a focus of invasive adenocarcinoma (i.e. malignant or progressed adenomas).

3.1. Sample baseline characteristics

Twenty-three pedunculated adenoma samples with a focus of invasive carcinoma, termed progressed adenomas, were used to assess single-cell CNAs implicated in the adenoma-to-adenocarcinoma transformation. These lesions represent models for colorectal carcinogenesis, as they ensure that the origin of the adenocarcinoma is maintained in the same sample and enable to trace the path of the malignant transformation. In addition, six non-advanced adenomas, defined as adenomas with low-grade dysplasia, tubular architecture and a size of <10mm, were also included in the study. The pertinent clinical parameters of the adenocarcinoma region of progressed adenomas are summarized in **Table 17**.

The majority of adenocarcinomas exhibited a low histological grade (69.6%), with the rest considered to be of high histological grade. 78.3% of the samples were classified as pT1. In them, the average submucosa infiltration was 3 mm. Lastly, we determined the presence of cells beyond the invasive front of the adenocarcinomas (i.e. budding), which was classified as high in 13% of our samples.

Table 17. Clinical sample data. All information refers to the adenocarcinoma region of the progressed adenoma samples.

Case	Histological grade	Submucosa Invasion (mm)	pT	Budding	MSS/MSI
1	High	-	pT2	Low	MSS
2	Low	3	pT1	High	MSS
3	Low	5	pT1	Low	MSS
4	Low		pTis	0	MSS
5	Low	5	pT1	Low	MSS
6	Low	2	pT1	Low	MSS
7	High	-	pT3	Low	MSS
8	Low	3	pT1	Low	MSS
9	Low	6	pT1	Low	MSS
10	Low	1	pT1	Low	MSS
11	Low	1	pT1	Low	MSS
12	High	2.5	pT1	High	MSS
13	High	1.5	pT1	Low	MSS
14	Low	1.3	pT1	Low	MSS
15	High	3	pT1	High	MSS
16	High	-	pT2	Low	MSS
17	Low	2	pT1	0	MSS
18	Low	2.5	pT1	Low	MSS
19	Low	-	pTis	0	MSS
20	Low	2	pT1	Low	MSS
21	High	3	pT1	Low	MSI
22	Low	2	pT1	Low	MSS
23	Low	1	pT1	Low	MSS

Histological grade, level of cell differentiation; pT, pathological grade of colonic wall infiltration; Budding, presence of tumor cells beyond the invasive front of the carcinoma; MSS; microsatellite stable; and MSI, microsatellite instable.

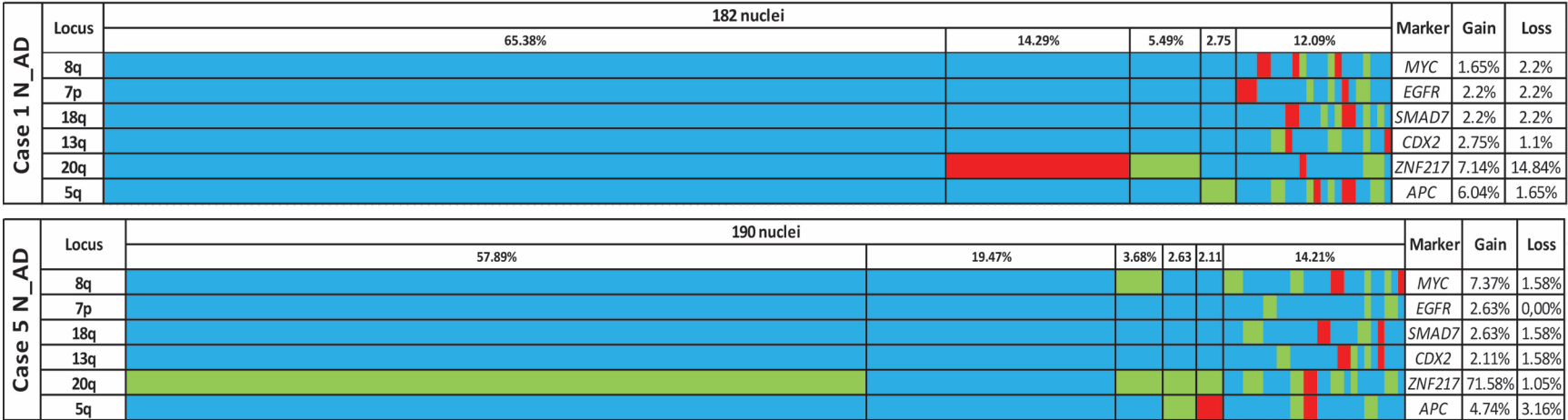
3.2. Chromosomal instability and clonal patterns

Three panels of FISH probes were sequentially hybridized to individual nuclei of the non-advanced adenoma samples (N_AD), as well as the adenoma (AD) and adenocarcinoma (ADK) regions of progressed adenomas. Repeated hybridization and relocation of the cells enabled us to enumerate the clonal aberration patterns on a cell-by-cell basis for all six gene probes, which served as surrogate markers for chromosome arm copy number. Based on our observations two groups of FISH signal patterns were established: single pattern clones and imbalanced clones. Single pattern clones were composed of cells that had the exact same number of signals for all the loci analyzed, such as 2,2,2,3,4,2. On the other hand, cells for which the different FISH signals matched each other in the direction of the change (gain,

loss, or unchanged) were considered part of the same imbalanced clone. For example, cells with 2,2,2,3,4,2 and 2,2,2,4,5,2 would be part of the same imbalanced clone because they both exhibit a gain in the 4th and 5th loci analyzed. Next, in order to visualize and compare the FISH results for the different samples, we organized cells based on their imbalanced clone classification. Finally, we arranged the different imbalanced clones for each case according to their incidence and the frequency of gains and losses for each locus was calculated as a percentage of the overall cell population (**Figure 23**).

The results for two N_AD and three progressed adenoma cases are shown in **Figure 23**. Several chromosome CNAs were observed in both the N_AD and progressed adenoma samples, showing chromosome heterogeneity (**Figure 23**). Interestingly, although the N_AD cases generally exhibited a more diploid chromosome content compared with the AD and ADK regions of progressed adenomas, the presence of several chromosome imbalances indicated that chromosome alterations were already present in early precursors of CRC.

A



B

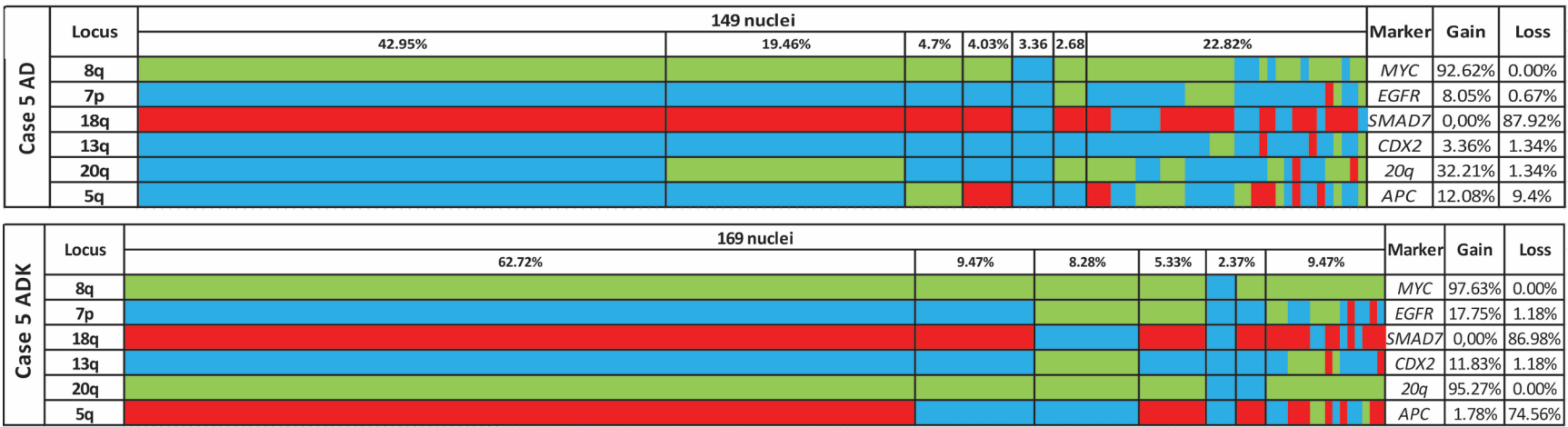


Figure 23. Summary of the imbalance clones in two non-advanced adenomas (N_AD) (A) and three progressed adenomas (AD and ADK regions) (B to D). Green, gains; red, losses; blue, unchanged. The organization of the graphs is the same for all cases and is explained for Case 1 N_AD (A top panel) in detail from left to right. The locus column shows the chromosome arms analysed by FISH. Each vertical line separates the most common imbalance clones. The row above the imbalance clone displays the percentages at which the clones were found. For example, the most frequent clone in Case 1 N_AD comprised 66.38% of the cells, with diploid content for all loci. The second most recurrent clone comprised 14.29% of the cells, with loss of chromosome 20q. The third most frequent clone comprised 5.49% of the cells, with gain of chromosome 20q. The Marker column shows the gene/probe name. The Gain column shows the percentage of gain and loss for each locus.

C

Case 14 AD	Locus	140 nuclei						Marker	Gain	Loss
		61.43%		15,00%		7.14%	2.86			
	8q					7.14%	2.86			

D

	Locus	121 nuclei									Marker	Gain	Loss
		21.49%	8.26%	7.44%	6.61%	5.79%	4.96%	4.13	3.31%	34.71%			
Case 15 AD	8q										MYC	86.78%	1.65%
	7p										EGFR	89.26%	1.65%
	18q										SMAD7	1.65%	75.21%
	13q										CDX2	2.48%	21.49%
	20q										ZNF217	49.59%	9.92%
	5q										APC	5.79%	27.27%

	Locus	147 nuclei						Marker	Gain	Loss
		25.85%	22.45%	5.44%	4.08%	2.72%	25.85%			
Case 15 ADK	8q							MYC	93.88%	0.00%
	7p							EGFR	93.2%	0.00%
	18q							SMAD7	16.33%	6.8%
	13q							CDX2	82.99%	1.36%
	20q							ZNF217	74.15%	7.48%
	5q							APC	36.05%	12.93%

Figure 23. Summary of the imbalance clones in two non-advanced adenomas (N_AD) (A) and three progressed adenomas (AD and ADK regions) (B to D) (Continued).

In order to quantify the levels of chromosome heterogeneity for the different sample types, we first calculated the frequency of the main imbalanced clone for each case. As a result, we observed the highest frequency of the main clone for N_AD, while the incidence of the main clone amongst the AD and ADK regions of progressed adenomas did not differ (**Figure 24**).

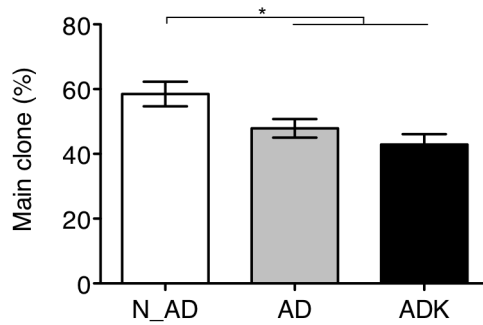


Figure 24. Percentage of the main imbalanced clone in the three sample categories analyzed. Data are reported as mean $\pm$ SEM. N_AD, non-advanced adenomas; AD, adenoma region of progressed adenomas; and ADK, adenocarcinoma region of progressed adenomas.

Additionally, we assessed the instability index, which is a measurement of the number of different cell types (i.e. cells with different FISH signals combinations) within a given cell population, and the overall incidence of chromosome imbalances in our samples. Accordingly, we detected a significant stepwise increase in both the instability index and the incidence of chromosome imbalances from the N_AD to AD and from AD to ADK (**Figure 25A and B**). Thus, indicating a significant increase in both CIN and chromosome heterogeneity during CRC development.

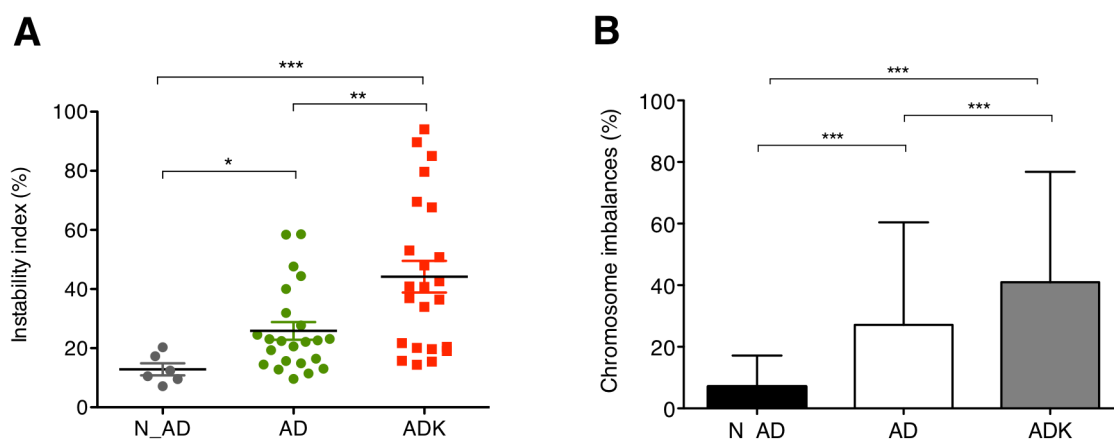


Figure 25. (A) Representation of the instability index for the three sample categories analyzed. Instability index calculated as follows: Number of patterns $\times$ 100 / number of cells analyzed. **(B)** Graph showing the percentage of overall chromosome imbalances observed in the three types of samples analyzed by FISH. N_AD, non-advanced adenoma; AD, adenoma region of progressed adenomas; and ADK, adenocarcinoma region of progressed adenomas. Data are reported as mean $\pm$ SEM. * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0001$.

When evaluating chromosome gains and losses individually, the AD and ADK showed the expected patterns for CRC: chromosomes 7p, 8q, 13q and 20q were mainly gained, while chromosomes 18q and 5q were mainly lost. Moreover, the gain of 20q and the loss of 18q were the most recurrent alterations for both the AD and ADK regions of progressed adenomas (39.6% vs. 63.2% for 20q gain; 14.9% and 20.4% for 18q loss, respectively). In the N_AD samples, however, the patterns of gains and losses were lower and more similar for all chromosomes, indicating there was no selection for specific imbalances yet. Overall, the AD region exhibited greater number of imbalances than N_AD for all the chromosomes analyzed (**Figure 26**). Interestingly, the frequency of chromosome 7q, 8q, 13q and 20q imbalances were also significantly higher in the ADK in comparison to the AD (**Figure 26**). Therefore, these results indicated the levels of aneuploidy sequentially increased during CRC pathogenesis and can distinguish the different lesions.

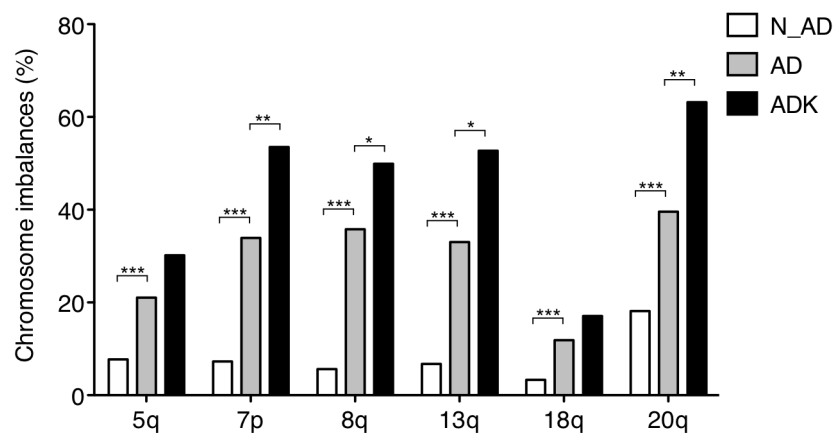


Figure 26. Graph showing the mean frequency of specific chromosome imbalances observed in non-advanced adenomas (N_AD) and in the adenoma (AD) and adenocarcinoma (ADK) regions of progressed adenomas. * $P < 0.05$, ** $P < 0.005$.

3.3. Cells from ADK exhibit a greater potential for progression than cells from AD

As we found a sequential increase in CIN and aneuploidy from the N_AD to the different regions of progressed adenomas (AD and ADK), we used the FIShtree software along with the single pattern data to construct tree-based models of clonal evolution. This software infers phylogenetic trees describing the most probable evolution of a clone based

on the observed signal counts. The model begins by assuming an initially diploid root cell within each lesion to heuristically minimize total copy number changes across the tree (see Materials and Methods).

Interestingly, we observed branched evolution in all the samples evaluated, from N_AD to AD and ADK. Next, we sought to evaluate whether there were statistically significant differences between the inferred trees derived from the different sample types by calculating the node depth, defined as the number of tree levels and considered a measure of progression potential. This analysis revealed the greatest node depth in trees from the ADK samples (**Figure 27**), thus indicating that cells from the ADK had a higher tendency to acquire genomic imbalances and progress into advanced disease than either the N_AD or AD cells. In order to assess whether any clinical characteristics of the ADK were related to their node depth, we performed the corresponding correlation analysis but were unable to find any positive correlations.

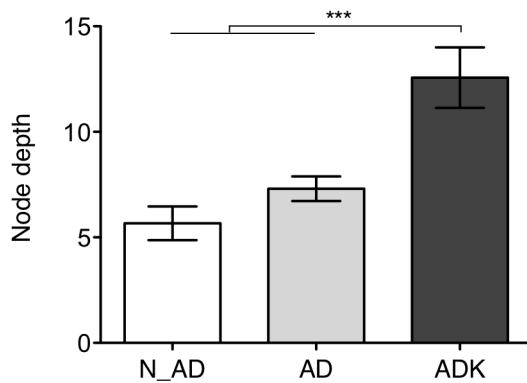


Figure 27. Graph representing the node depth of the three sets of samples analyzed. N_AD, non-advanced adenoma; AD, adenoma region of progressed adenomas; and ADK, adenocarcinoma region of progressed adenomas. Data are reported as mean $\pm$ SEM. *** $P < 0.0001$.

3.4. Consensus evolutionary trees reveal common ancestor clones and drivers of malignant transformation

Considering that the AD and ADK regions were part of the same lesion within progressed adenomas, we assumed that for every case, the ADK came from the evolution of its corresponding AD. Therefore, by using FIShtrees software we inferred consensus evolutionary trees for the progressed adenoma samples, thus enabling us to reconstruct the evolution of each AD into ADK (**Figure 28**).

As is shown in **Figure 28**, the resulting consensus trees maintained the branched evolution. The clones of the AD were closer to the root of the tree, while the clones from the ADK arose from the AD population and were located further away from the root. The frequency of each clone's signal pattern was also included in the tree in order to show the distribution of the AD and ADK populations. Interestingly, we observed that low-frequency clones led to the heterogeneity and represented the majority of the population for each case (**Figure 29**). In fact, none of the cases showed a strong selection for a specific single pattern clone. A summary of the results obtained is detailed in **Table 18**.

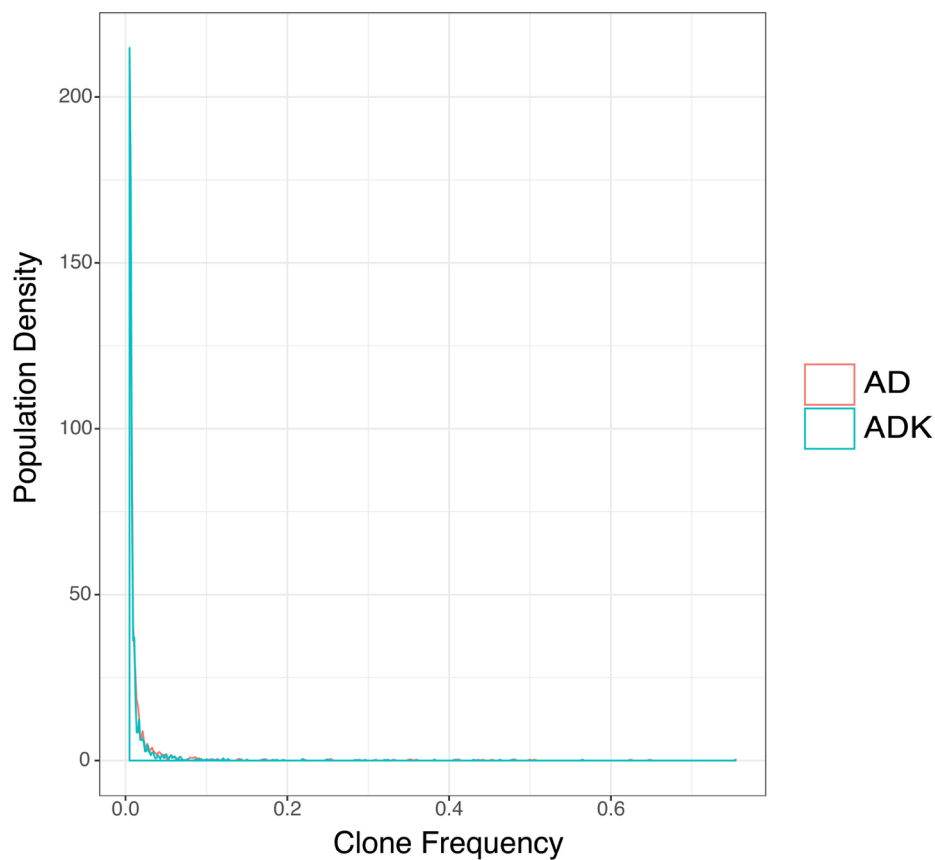


Figure 29. Representation of the different clones in the overall cell population according to their frequency. AD, adenoma region of progressed adenomas; and ADK, adenocarcinoma region of progressed adenomas.

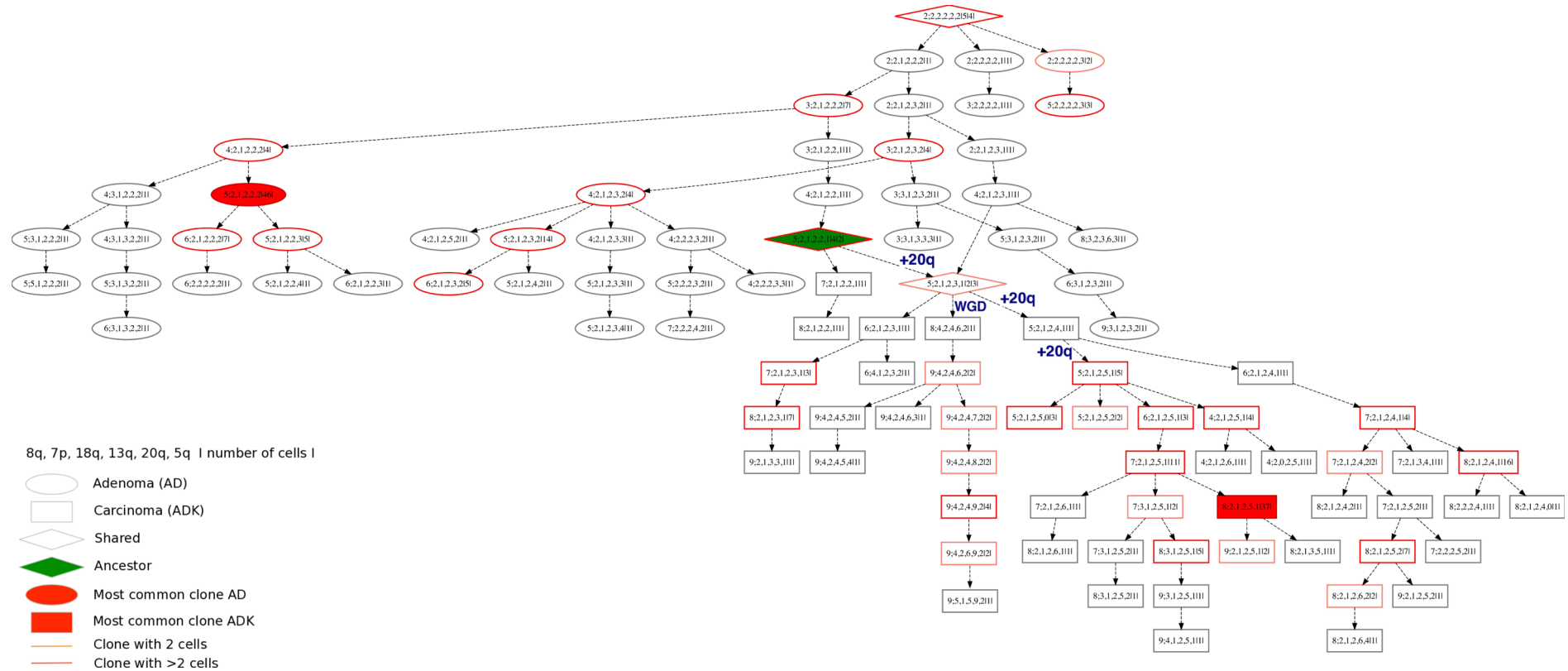
A

Figure 28. Consensus phylogenetic trees for three progressed adenoma cases. Circles represent clones from AD, squares represent clones from ADK and diamonds are shared clones between AD and ADK. Green indicates common ancestor and red the most recurrent clone. **(A)** Consensus tree for case 5 of progressed adenoma samples. The gain of chromosome 20q is the driver of the AD-to-ADK transition. Moreover, a WGD leads the transition to a secondary branch of the evolution. AD, adenoma region of progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and WGD, whole genome duplication.

B

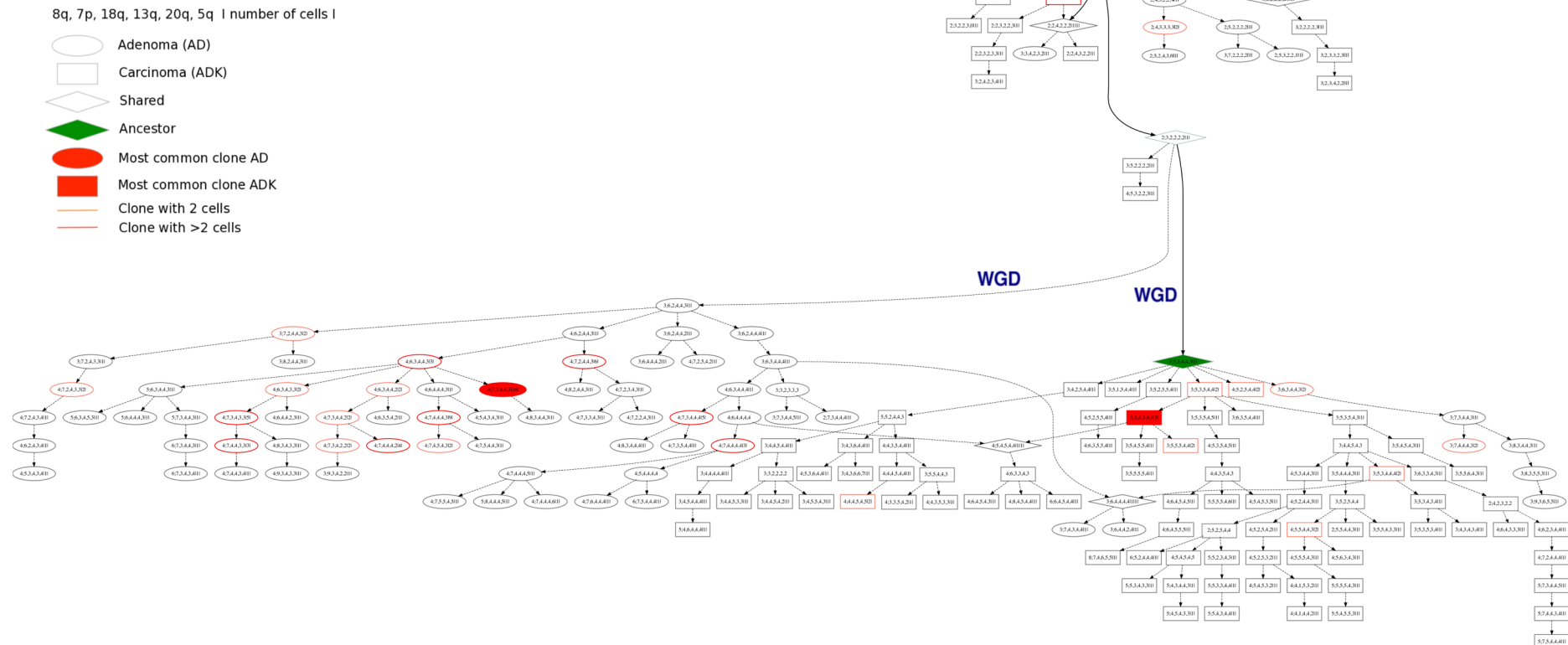


Figure 28. Consensus phylogenetic trees for three progressed adenoma cases. Circles represent clones from AD, squares represent clones from ADK and diamonds are shared clones between AD and ADK. Green indicates common ancestor and red the most recurrent clone. **(B)** Consensus tree for case 14 of progressed adenoma samples. A WGD event leads to the AD-to-ADK transition and also the progression of the AD population. AD, adenoma region of progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and WGD, whole genome duplication.

C

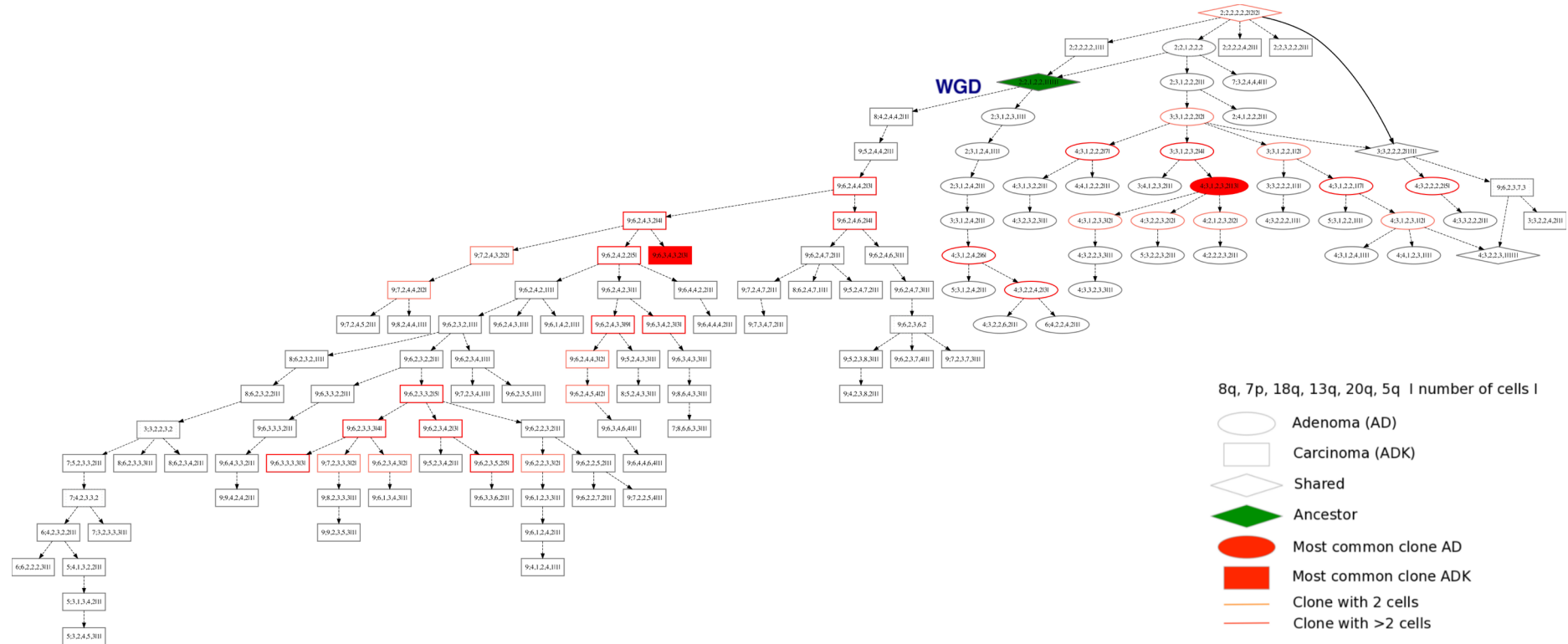


Figure 28. Consensus phylogenetic trees for three progressed adenoma cases. Circles represent clones from AD, squares represent clones from ADK and diamonds are shared clones between AD and ADK. Green indicates common ancestor and red the most recurrent clone. (C) Consensus tree for case 15 of progressed adenoma samples. A WGD event is the driver of the AD-to-ADK transition and it is not present in the AD population. AD, adenoma region of progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and WGD, whole genome duplication.

Table 18. Summary of the results obtained from the inferred phylogenetic trees of AD and ADK progressed adenoma samples.

Case	Lesion	Depth Node	Instability Index	Main clone*	Common ancestor*	Driver(s) AD-to-ADK	Driver(s) progression	WGD events**	Ploidy range***
1	AD	6	15.66	8q (2), 7p (2), 18q (2), 13q (3), 20q (2), 5q (2)	- 8q (2), 7p (2), 18q (2), 13q (2), 20q (3), 5q (2) - 8q (2), 7p (2), 18q (2), 13q (4), 20q (2), 5q (2)	Gain 13q and 20q	Gain 13q	None	1.8c - 5c
	ADK	13	36.42	8q (2), 7p (3), 18q (2), 13q (4), 20q (4), 5q (2)					1.8c - 7c
2	AD	11	40	8q (2), 7p (3), 18q (2), 13q (2), 20q (3), 5q (2)	8q (2), 7p (2), 18q (2), 13q (2), 20q (3), 5q (2)	Gain 13q	Gain 5q and 20q	ADK progression	1.8c - 4.4c
	ADK	16	50.84	8q (3), 7p (4), 18q (2), 13q (5), 20q (4), 5q (4)					2.1c - 16.1c
3	AD	4	11.41	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (3)	ND	Gain 7p, 8q and 20q	ADK progression	1.9c - 5.4c
	ADK	14	40.91	8q (3), 7p (4), 18q (2), 13q (5), 20q (4), 5q (4)					1.9c - 9.2c
4	AD	4	9.64	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	8q (2), 7p (2), 18q (2), 13q (2), 20q (3), 5q (2)	Gain 7p	Gain 7p	None	1.9c - 5.4c
	ADK	7	15.43	8q (2), 7p (3), 18q (2), 13q (2), 20q (3), 5q (2)					1.9c - 4.8c
5	AD	8	31.94	8q (5), 7p (2), 18q (1), 13q (2), 20q (2), 5q (2)	8q (5), 7p (2), 18q (1), 13q (2), 20q (2), 5q (1)	Gain 20q	Gain 20q	ADK progression	ND
	ADK	11	36.97	8q (8), 7p (2), 18q (1), 13q (2), 20q (5), 5q (1)					ND
6	AD	6	13.07	8q (3), 7p (3), 18q (2), 13q (2), 20q (2), 5q (2)	8q (3), 7p (3), 18q (2), 13q (2), 20q (2), 5q (2)	ND	ND	None	1.9c - 5.8c
	ADK	5	14.38	8q (3), 7p (3), 18q (2), 13q (2), 20q (2), 5q (2)					1.9c - 4.9c
7	AD	6	12.80	8q (2), 7p (2), 18q (1), 13q (2), 20q (3), 5q (2)	8q (2), 7p (2), 18q (1), 13q (2), 20q (3), 5q (2)	ND	Gain 8q	None	1.9c - 4.9c
	ADK	6	20.14	8q (2), 7p (2), 18q (1), 13q (2), 20q (3), 5q (2)					1.9c - 4.4c
8	AD	8	14.87	8q (3), 7p (3), 18q (2), 13q (3), 20q (4), 5q (3)	8q (3), 7p (3), 18q (2), 13q (2), 20q (4), 5q (3)	ND	ND	None	1.8c - 5.2c
	ADK	7	20.44	8q (3), 7p (3), 18q (2), 13q (2), 20q (4), 5q (3)					1.8c - 4.5c

AD, adenoma region progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and ND, not determined.

* Main clone and common ancestor are shown as: Chromosome (number of copies), **WGD events confirmed by the detection of cells that have undergone genome doubling by ploidy cytometry, ***Ploidy range obtained by DNA ploidy cytometry.

Table 18. Summary of the results obtained from the inferred phylogenetic trees of AD and ADK from progressed adenoma samples (Continued).

Case	Lesion	Depth Node	Instability Index	Main clone*	Common ancestor*	Driver(s) AD-to-ADK	Driver(s) progression	WGD events**	Ploidy range***
9	AD	8	27.73	8q (2), 7p (2), 18q (2), 13q (3), 20q (2), 5q (1)	8q (2), 7p (2), 18q (1), 13q (3), 20q (3), 5q (1)	Gain 13q and 20q	Gain 20q	None	1.8c - 5.7c
	ADK	16	53.06	8q (2), 7p (2), 18q (1), 13q (4), 20q (3), 5q (1)					1.8c - 5.4c
10	AD	6	19.33	8q (2), 7p (2), 18q (2), 13q (3), 20q (3), 5q (2)	8q (2), 7p (2), 18q (1), 13q (3), 20q (3), 5q (2)	Loss 5q and gain 20q	Gain 7p and 20q	None	1.8c - 3.3c
	ADK	13	40.71	8q (2), 7p (2), 18q (1), 13q (3), 20q (5), 5q (1)					1.8c - 4.6c
11	AD	5	14.44	8q (2), 7p (2), 18q (2), 13q (2), 20q (3), 5q (2)	8q (2), 7p (2), 18q (2), 13q (3), 20q (3), 5q (2)	Gain 7p and 13q	Gain 7p and 13q	None	1.8c - 3.8c
	ADK	6	15.74	8q (2), 7p (2), 18q (2), 13q (3), 20q (3), 5q (2)					1.4c - 4.9c
12	AD	11	58.45	8q (6), 7p (4), 18q (2), 13q (5), 20q (6), 5q (3)	8q (6), 7p (4), 18q (2), 13q (5), 20q (6), 5q (3)	WGD (in AD)	Several	Main transition AD progression	1.9c - 9.8c
	ADK	16	69.57	8q (6), 7p (4), 18q (2), 13q (5), 20q (6), 5q (3)					1.8c - 7.7c
13	AD	9	44.38	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	8q (2), 7p (3), 18q (2), 13q (2), 20q (2), 5q (2)	Gain 13q	Gain 20q	AD and ADK progression	1.9c - 7.9c
	ADK	14	79.71	8q (2), 7p (3), 18q (2), 13q (2), 20q (2), 5q (2)					1.9c - 7.3c
14	AD	15	58.57	8q (4), 7p (7), 18q (3), 13q (4), 20q (4), 5q (3)	8q (3), 7p (5), 18q (3), 13q (4), 20q (4), 5q (3)	WGD (in AD)	Several	Main transition AD progression	1.9c - 12.3c
	ADK	20	89.66	8q (3), 7p (5), 18q (4), 13q (5), 20q (4), 5q (4)					1.8c - 7.1c
15	AD	10	47.67	8q (4), 7p (3), 18q (1), 13q (2), 20q (3), 5q (2)	8q (2), 7p (2), 18q (1), 13q (2), 20q (2), 5q (1)	WGD	Gain 7p and 8q	Main transition	1.8c - 5.2c
	ADK	19	85.07	8q (9), 7p (6), 18q (2), 13q (4), 20q (3), 5q (3)					1.8c - 9.5c
16	AD	5	22.14	8q (2), 7p (2), 18q (2), 13q (2), 20q (3), 5q (2)	8q (3), 7p (3), 18q (2), 13q (2), 20q (2), 5q (2)	Gain 7p	Gain 13q and 20q	None	1.8c - 4.4c
	ADK	20	67.65	8q (4), 7p (5), 18q (3), 13q (6), 20q (4), 5q (4)					1.8c - 5.8c

AD, adenoma region progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and ND, not determined.

* Main clone and common ancestor are shown as: Chromosome (number of copies), **WGD events confirmed by the detection of cells that have undergone genome doubling by ploidy cytometry, ***Ploidy range obtained by DNA ploidy cytometry.

Table 18. Summary of the results obtained from the inferred phylogenetic trees of AD and ADK from progressed adenoma samples (Continued).

Case	Lesion	Depth Node	Instability Index	Main clone*	Common ancestor*	Driver(s) AD-to-ADK	Driver(s) progression	WGD events**	Ploidy range***
17	AD	9	24.55	8q (2), 7p (2), 18q (2), 13q (3), 20q (3), 5q (2)	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (1)	WGD	Gain 20q	Main transition ADK progression	1.8c - 4.8c
	ADK	11	42.61	8q (4), 7p (3), 18q (2), 13q (4), 20q (8), 5q (2)					1.8c - 7.6c
18	AD	9	22.41	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	8q (2), 7p (2), 18q (1), 13q (2), 20q (3), 5q (1)	WGD (in AD)	Several	Main transition AD and ADK progression	1.8c - 9.2c
	ADK	16	47.96	8q (2), 7p (2), 18q (1), 13q (3), 20q (4), 5q (1)					1.9c - 8.8c
19	AD	4	20.59	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	ND	ND	ND	None	1.8c - 5.2c
	ADK	7	21.71	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)					1.8c - 5.2c
20	AD	8	22.65	8q (3), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	8q (4), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	Gain 8q and 20q	Gain 7p, 13q and 20q	ADK progression	1.8c - 6.0c
	ADK	34	94.02	8q (6), 7p (2), 18q (2), 13q (2), 20q (3), 5q (2)					1.8c - 9.0c
21	AD	7	23.08	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	8q (2), 7p (2), 18q (1), 13q (2), 20q (2), 5q (2)	Gain 7p	ND	None	1.8c - 5.3c
	ADK	8	33.98	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)					1.8c - 4.4c
22	AD	5	23.14	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	ND	ND	ND	None	1.8c - 4.2c
	ADK	5	19.71	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)					1.8c - 4.9c
23	AD	4	16.41	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)	ND	ND	ND	None	1.8c - 4.4c
	ADK	5	19.05	8q (2), 7p (2), 18q (2), 13q (2), 20q (2), 5q (2)					1.9c - 4.4c

AD, adenoma region progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and ND, not determined.

* Main clone and common ancestor are shown as: Chromosome (number of copies), **WGD events confirmed by the detection of cells that have undergone genome doubling by ploidy cytometry, ***Ploidy range obtained by DNA ploidy cytometry.

In order to assess the potential for disease progression of the different single pattern clones, we assigned each of them a fitness value. This fitness value is a measure of the allelic expansion of a determined copy number status in the ADK with respect to the AD (see Materials and Methods). We then defined the most common branch as the one with more number of cells and the fittest branch as the one with the clones exhibiting more fitness. Interestingly, we observed that the most common branch was not the fittest one. When considering the fitness for individual chromosome alterations, genomic gains affecting chromosomes 7p, 13q and 20q were the fittest ($P = 0.034$), indicating that they were the main drivers of progression (**Figure 30**). We also observed that the gain of four or more copies of a chromosome enhanced the fitness for all chromosomes except for 18q, for which its loss conferred the greatest fitness ($P = 0.022$ and $P = 0.003$) (**Figure 30**). Thus, these results suggested that the accumulation of recurrently gained chromosomes leads to high levels of progression.

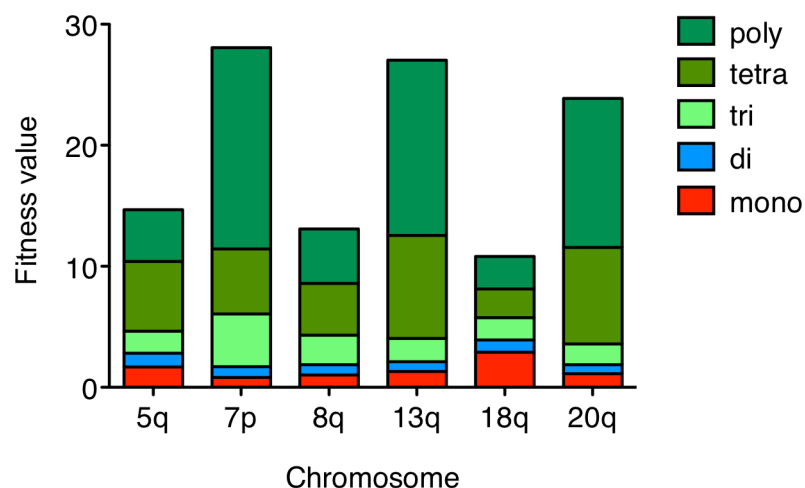


Figure 30. Fitness values for the different chromosomes imbalances observed during FISH analysis. Poly, >4 chromosome copies; tetra, 4 chromosome copies; tri, 3 chromosome copies; di, 2 chromosome copies; and mono, 1 chromosome copy.

Analysis of the consensus trees also allowed us to infer the common ancestor clone for 20 of the 23 progressed adenomas analyzed (**Table 18**). Among the clones shared between the AD and ADK of every case, the common ancestor was considered to be the clone responsible for leading to ADK initiation. In three cases, the common ancestor was not determined because the FISH data resulted in consensus trees with a mixture of populations

and did not allow for the proper separation between AD and ADK. Next, the identified common ancestors were divided into two categories according to their origin: the fittest or the most common branch of the AD population. Consequently, we observed that the ADK population did not always arise from the most common branch of the AD, but emerged from low-frequency clones that displayed greater fitness (45% vs. 55%). Also, in 43.8% of the cases the common ancestor already exhibited a primary imbalance, mainly as a trisomy, in the chromosome driving the ADK progression (**Table 18**).

Furthermore, the identification of the common ancestor led us to assess the drivers of transition, defined as those chromosome imbalances that lead to the transition from the common ancestor to the earliest clones of the ADK. These drivers of transition were identified for 70% of the cases analyzed and although the major ones corresponded to previously identified drivers of progression, for six cases they were composed of different chromosome imbalances (**Figure 31** and **Table 18**).

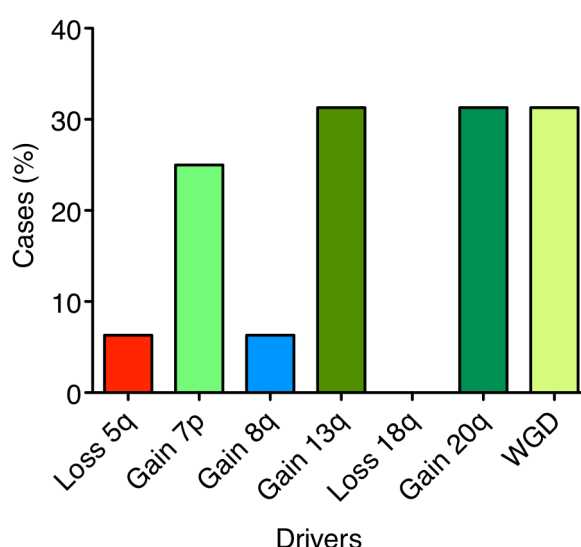


Figure 31. Graph showing the CNAs involved in the AD-to-ADK transition based on their frequency. WGD, whole genome duplication; AD, adenoma region of progressed adenomas; ADK, adenocarcinoma region of progressed adenomas; and CNAs, copy number alterations.

3.5. Whole genome duplication events play a main role in CRC progression

A potential whole genome duplication (WGD) event was identified in five cases as the driver of transition (31.3%) (**Figure 31** and **Table 18**). Indeed, in these five cases several chromosome imbalances showed similar fitness, indicating that different chromosomes

contributed to the progression in a similar way, which is in agreement with the consequences of a genome-doubling event. Moreover, the fact that the fittest imbalances were conferred by the gain of multiple copies of chromosomes, as explained above, further supports the existence of WGD events occurring in some of the cases analyzed.

Interestingly, in three of these cases, the WGD event occurred in the AD just prior to the transition (**Figure 28B**), and WGD events were observed during the evolution of 20% of the AD and 35% of the ADK cases. For example, as is shown in **Figure 28B**, WGD events lead to the transition to ADK and also the evolution of AD.

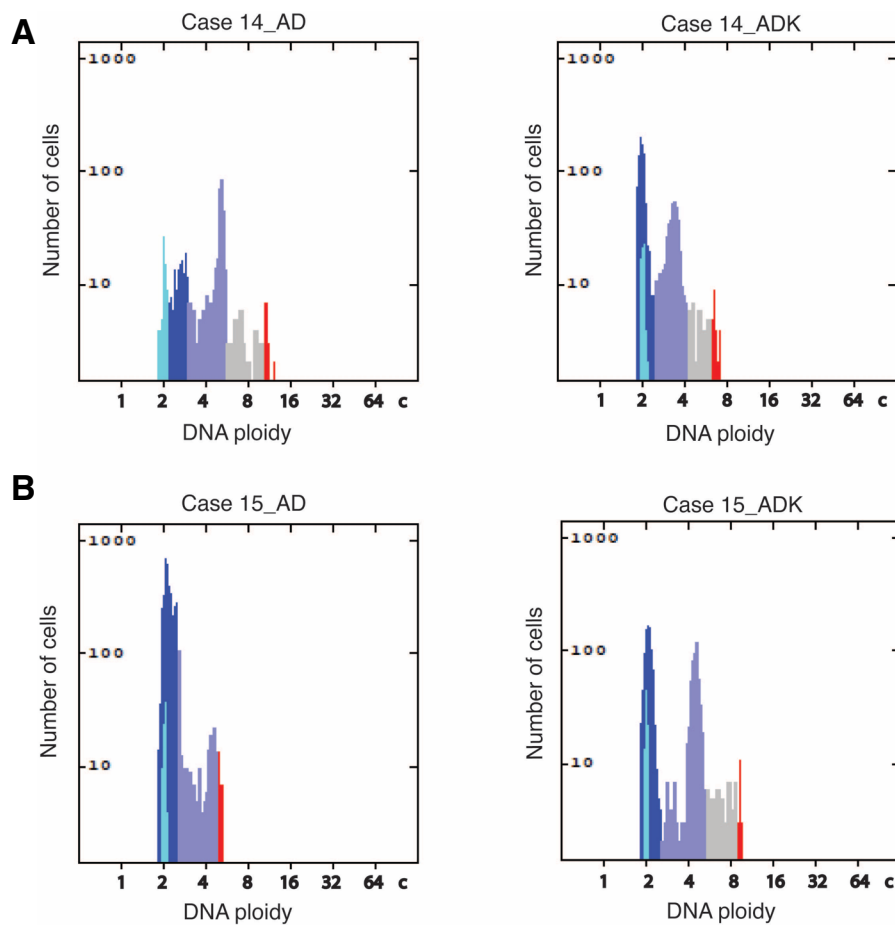


Figure 32. Graphs showing the ploidy of two progressed adenoma cases. The colors indicate cells with different levels of ploidy. The light blue shows the ploidy of the reference cells. **(A)** DNA ploidy measurements for case 14 of progressed adenomas showing polyploid cells in both AD and ADK. **(B)** DNA ploidy measurements for case 15 of progressed adenomas showing diploid cells in the AD and polyploid cells in the ADK.

Therefore, in order to confirm the presence of WGD events suggested by the evolutionary trees, we sought to analyze the nuclear DNA content of the AD and ADK samples by image cytometry. The results obtained from these measurements confirmed the presence of polyploid cells in all suspected cases (**Figure 32**), thus supporting the importance of genome doubling events during the early stages of colorectal carcinogenesis.

4. Chapter IV: Near-tetraploid cancer cells show chromosome instability triggered by replication stress and exhibit enhanced invasiveness

In further support of the previous findings, tetraploidy has been widely proposed to be an intermediate stage during malignant transformation, often preceding the chromosomally unstable populations observed in many cancers. In fact, as the aforementioned results help demonstrate, the presence of tetraploidy during early stages of colorectal carcinogenesis leads to the development of adenocarcinoma. Despite these findings, the physiological consequences of WGD remain to be fully elucidated. Therefore, we sought to investigate tetraploidy *in vitro* in order to unveil the consequences of genome doubling in CRC.

4.1. Generation and characterization of diploid and near-tetraploid isogenic clones

To investigate the consequences of tetraploidization in cancer cells, we isolated diploid and near-tetraploid (hereafter also referred to as 2N and 4N, respectively) isogenic cell lines derived from the diploid and stable CRC cell lines DLD-1 and RKO. Cytogenetic analysis suggested the existence of a 4N subpopulation in the DLD-1 cell line (**Figure 33**), leading us to perform single-cell flow sorting to isolate 2N and 4N clones based on differences in cell size and then confirm the ploidy status (**Figure 34**).

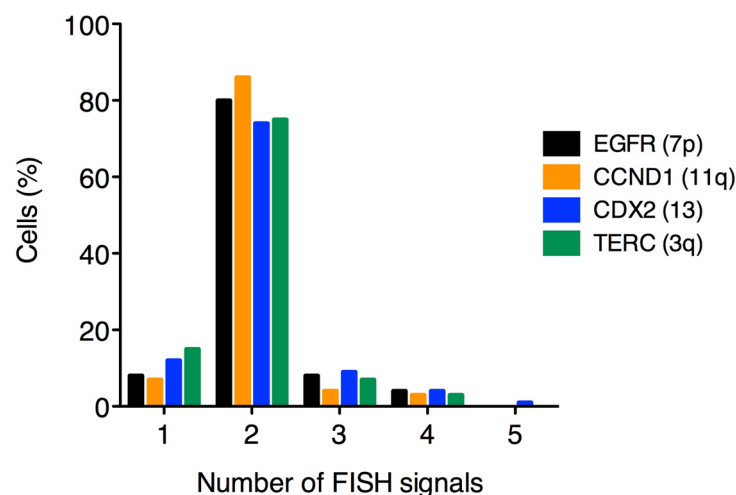


Figure 33. Percentage of FISH signals for four locus-specific probes on 5,000 parental DLD-1 cells deciphering the presence of a 4N cellular subpopulation.

Conversely, 4N clones for the RKO cells were generated by cytokinesis inhibition followed by single-cell isolation and ploidy confirmation (see Materials and Methods for further details).

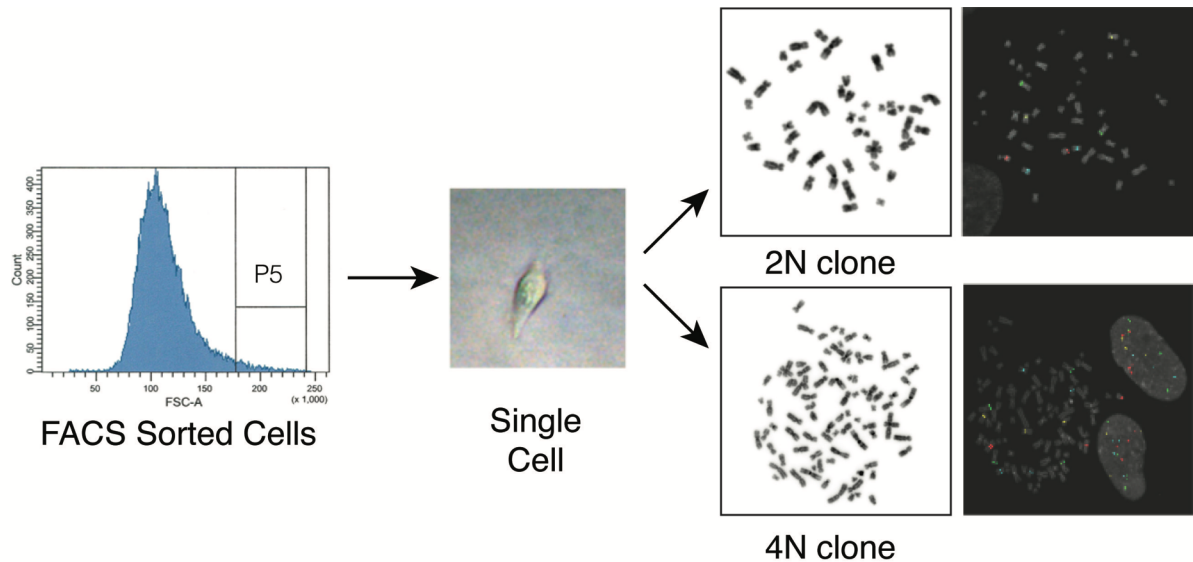


Figure 34. Diagram showing the procedure for single cell sorting and generation of pure 2N and 4N isogenic clones. The sort window P5 selected cells having a high forward scatter, which are larger and have a higher probability of being 4N than cells with a low forward scatter. The ploidy of the generated clones was then confirmed by chromosome counts and FISH experiments.

Once we established the different clones, we conducted proliferation assays in the DLD-1 cells, which showed that 4N cells proliferated slower than their 2N counterparts (**Figure 35**). Additionally, this growth impairment was further confirmed by cell viability and colony formation assays (**Figure 36A and B**).

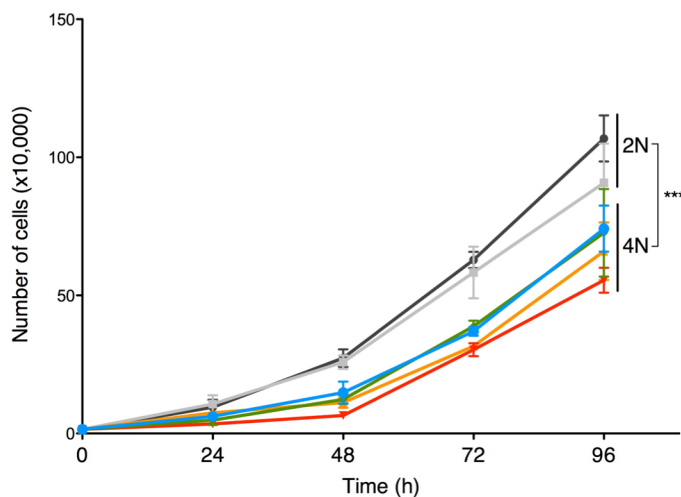


Figure 35. Growth curve for DLD-1 derived 2N and 4N clones. Statistical analysis was performed on the trends using R package geepack for longitudinal analysis. *** $P < 0.0001$

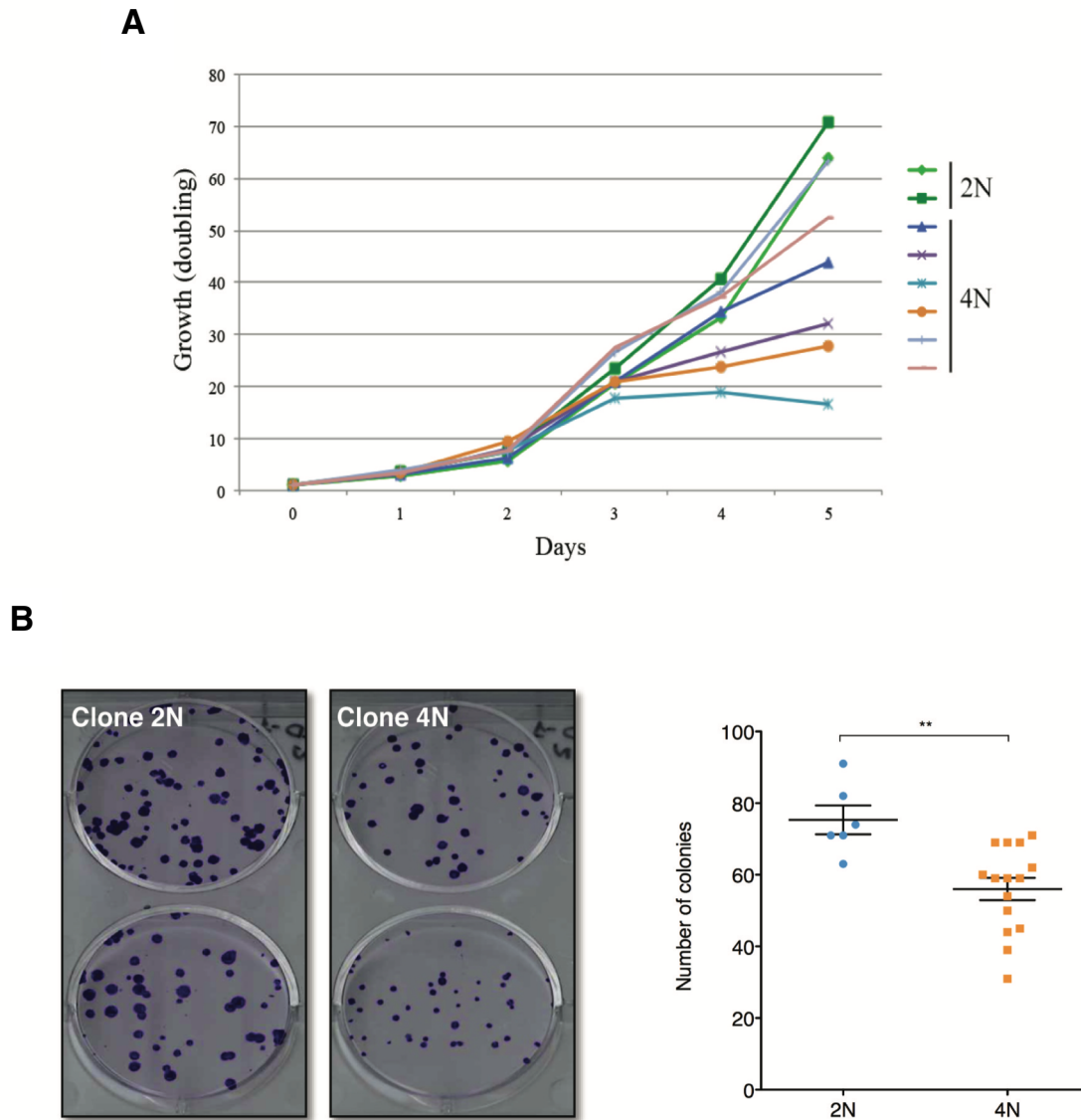


Figure 36. Impairment of cellular growth in DLD-1 derived 4N cells. **(A)** Proliferation assay using CellTiter-Blue in two 2N and six 4N clones. **(B)** Left: Representative colony formation assay in two 2N clones and two 4N clones. Right: Quantification of colony formation assay. Levels represented as number of colonies $\pm$ SEM ($n=3$). $**P < 0.001$.

Next, in order to study the effect of WGD on nucleic acid and protein levels we synchronized 2N and 4N DLD-1 cells in G1/S phase and assessed their total DNA, RNA, and protein content. While 4N clones had an approximately 2-fold increase in DNA and RNA content, the protein quantification showed a 3-fold increase and significant heterogeneity (**Figure 37**).

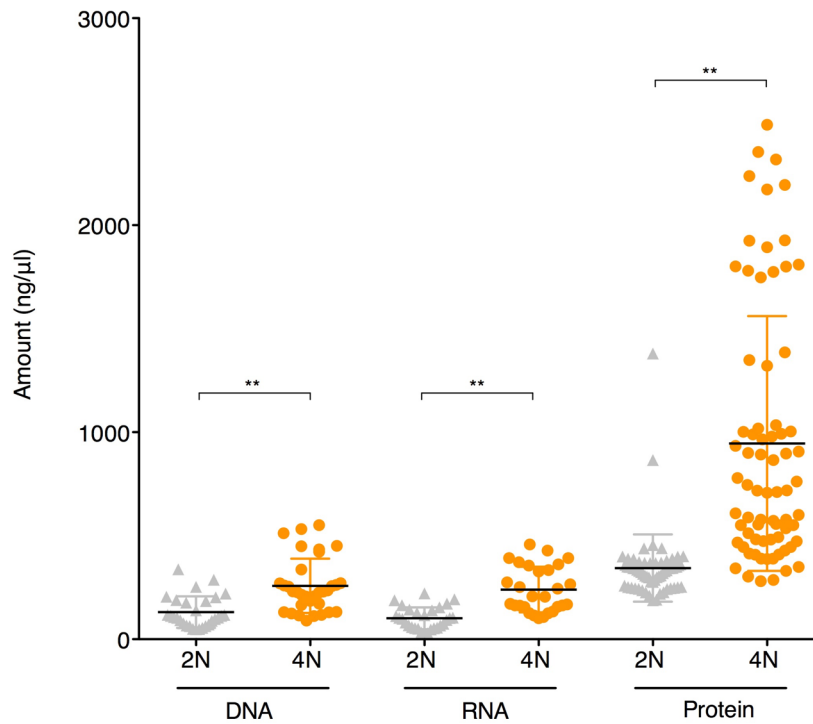


Figure 37. Comparison of total DNA, RNA and protein amounts in seven 2N and seven 4N DLD-1 clones. Each data point represents a measurement for biological replicates (n= 4-9). ** $P < 0.001$.

Finally, we sought to investigate whether 4N cells displayed an increase in nuclear size. Automated image analysis revealed that 4N cells exhibited a 1.7-fold increase in nuclear area (**Figure 38A**) and a 2-fold increase in volume (**Figure 38B**).

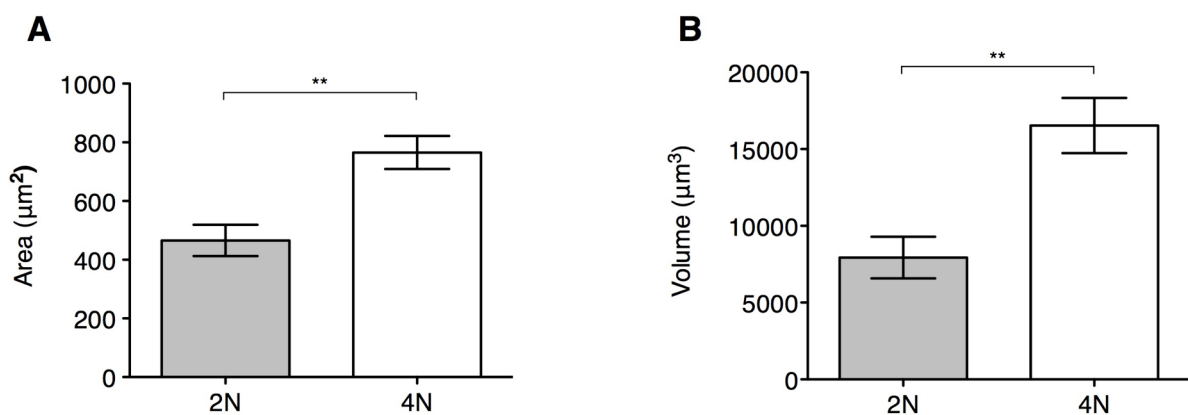


Figure 38. Measurements for the area (μm^2) (**A**) and volume (μm^3) (**B**) for approximately 5,000 DLD-1 cells of each 2N and 4N individual clone are represented. Data are reported as mean $\pm$ SEM. ** $P < 0.001$.

4.2. Transcriptional profiling of 4N cells reveals a deregulation of genes associated with DNA replication and mitosis

We next explored the molecular changes associated with tetraploidy by performing a genome-wide transcriptome profiling of the 2N and 4N DLD-1 clones. With the exception of one clone, we found a distinct difference in gene expression between the two sets of clones, according to the principal component analysis. In fact, hierarchal gene clustering revealed a gene signature that further distinguished the two sets of clones (**Figure 39**).

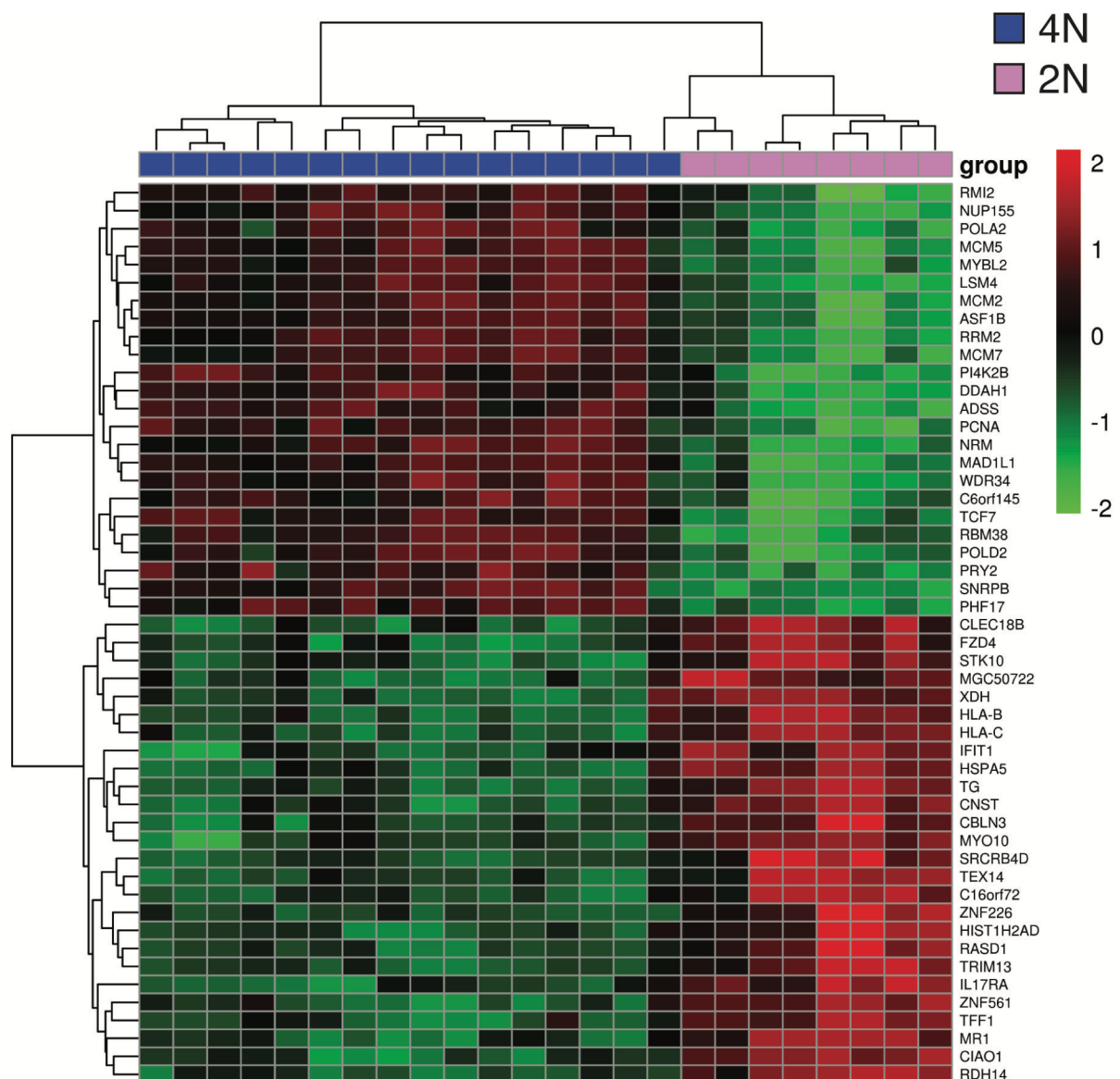


Figure 39. Gene expression profiling. Hierarchical clustering for genes that distinguishes between DLD-1 derived 2N and 4N clones. Top-ranked 50 genes sorted by FDR are shown. FDR, False Discovery Rate.

Overall, 328 genes were differentially regulated in 4N vs. 2N clones (fold change > 1.5, FDR < 0.05) and gene set enrichment analysis showed a significant deregulation of genes associated with the cell cycle, mainly those involved in DNA replication and mitosis (Normalized Enrichment Score > 2.00, $P < 0.001$) (Figure 40A and B).

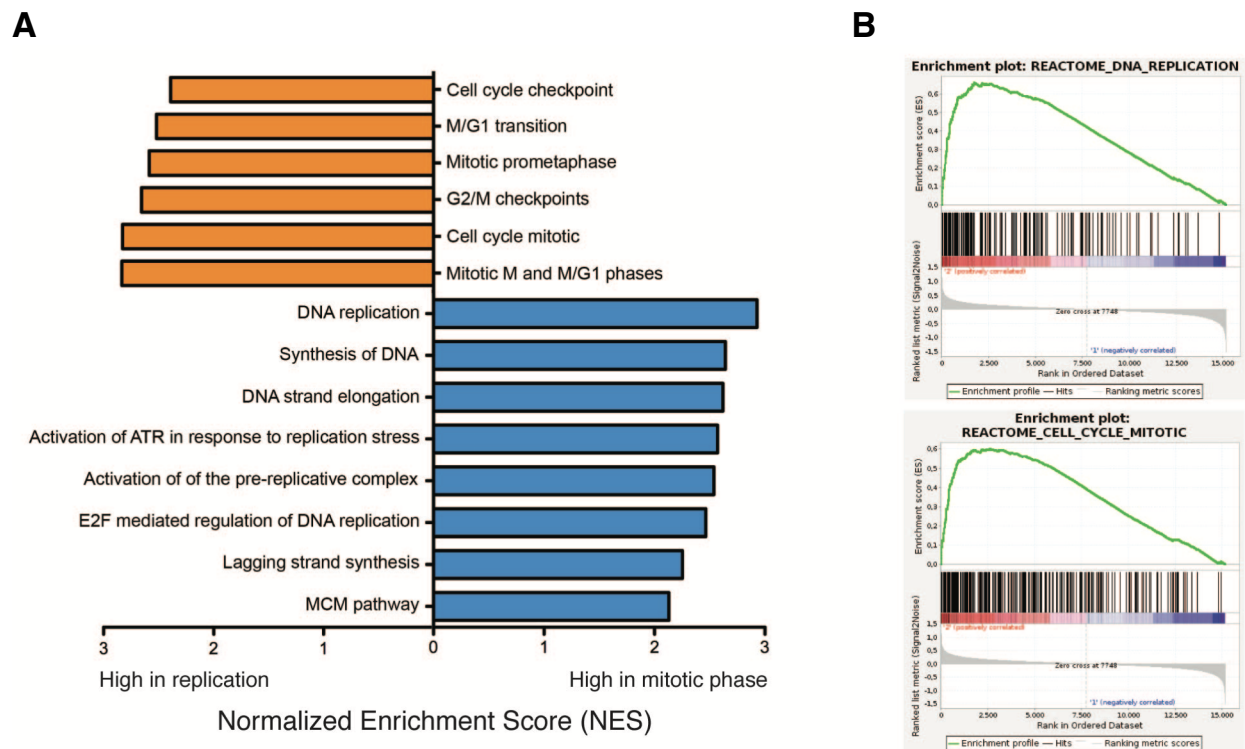


Figure 40. (A) Summary of gene set enrichment analysis showing the most over-represented pathways included in the Molecular Signature Database. (B) Gene set enrichment analysis for DNA replication and mitosis comparing 2N and 4N DLD-1 clones.

Notably, overexpression of the replication-associated genes *RRM2* and *MCM2*, which have previously been shown to be involved in the replication stress response, and the overexpression of the mitosis-related gene *ESPL1* were confirmed at the protein and mRNA level, respectively (Figure 41A and B). Nevertheless, we observed variability in the levels of expression for *RRM2* in 4N clones and only two clones showed a clear increase in the mRNA levels for this gene.

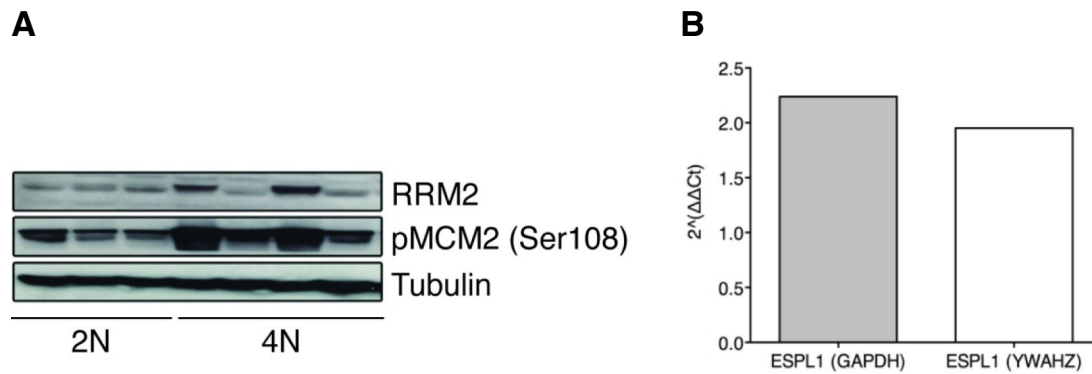


Figure 41. (A) Immunoblot validating the upregulation of *RRM2* and phospho-*MCM2* in DLD-1 4N clones. Tubulin was used as a protein loading control. (B) Graph representing the fold-change of the gene *ESPL1* in the 4N cells vs 2N cells. GAPDH and YWAHZ were used as expression control.

4.3. Near-tetraploid cells display increased replication stress and DNA damage

As the transcriptomic profiling data revealed an overexpression of replication stress-related genes in the 4N clones, we decided to further characterize the levels of replication stress in the 2N compared with 4N clones. First, we assessed the activation of DNA replication and DNA damage checkpoints. Interestingly, while we did not detect any differential activation of the ATM/CHK2 pathway (**Figure 42**), we did observe enhanced phosphorylation of CHK1 and increased levels of RPA in the 4N clones (**Figure 43A and B**). Importantly, we determined the levels of total CHK1 and CHK2 in order to exclude an increase in the protein levels due to the elevated DNA content *per se*.

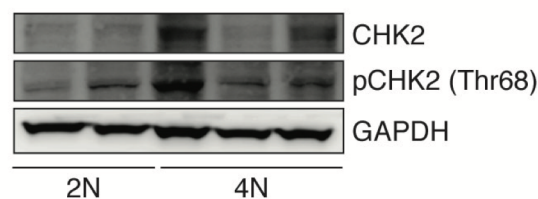


Figure 42. Representative western-blot of the levels of CHK2 and its active form phospho-CHK2 (pCHK2) in DLD-1 clones. Average ratio pCHK2/CHK2 of 0.82 for 2N and 0.74 for 4N cells. GAPDH was used as a protein loading control.

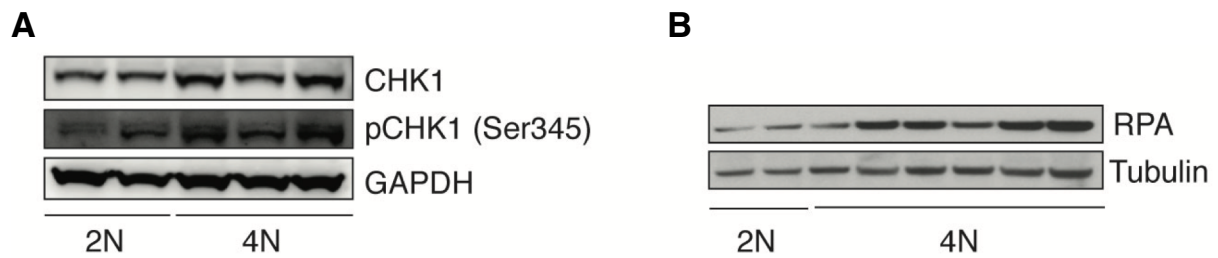


Figure 43. (A) Immunoblot analysis of total CHK1 and phospho-CHK1 (pCHK1) levels in DLD-1 2N and 4N cells. GAPDH was used as a protein loading control. Average ratio pCHK1/CHK1 of 0.63 for 2N and 1.14 for 4N cells. (B) Immunoblot analysis for RPA showing higher levels in 4N DLD-1 cells. Tubulin was used as a protein loading control.

We also determined the number of fragile sites *per* chromosome, for which 4N cells exhibited the highest incidence (**Figure 44**). Additionally, 4N cells showed a greater increase in the fragile sites incidence (2-fold vs. 2.7-fold for 2N and 4N, respectively) and the phospho-CHK1 levels following treatment with the DNA polymerase inhibitor aphidicolin, indicating that they had higher sensitivity to replication stress (**Figure 44 and 45**).

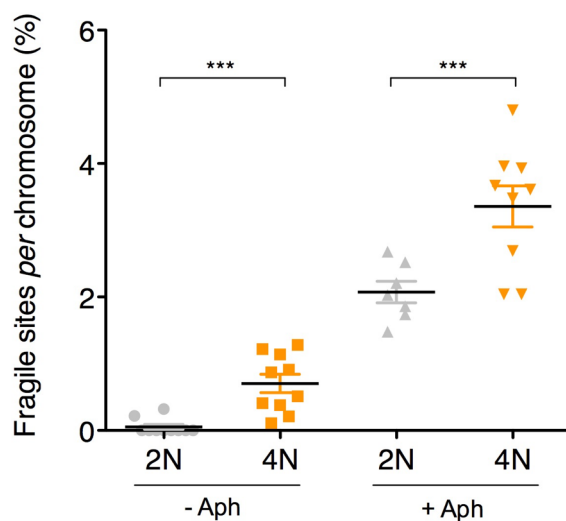


Figure 44. Percentage of random fragile sites *per* chromosome in untreated as well as treated DLD-1 cells with aphidicolin (0.2 μ M) for 24 hours ($n > 50$ metaphases / clone). *** $P < 0.0001$.

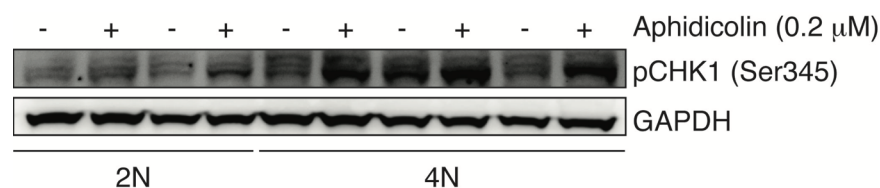


Figure 45. Analysis of phospho-CHK1 (pCHK1) levels after the addition of aphidicolin (0.2 μ M) for 24 hours to both sets of DLD-1 clones. GAPDH was used as a protein loading control.

In order to rule out that replication stress findings in 4N clones were cell line specific, we determined the presence of replication stress response in RKO clones. In support of the findings in DLD-1 cells, we also observed an increase in the phosphorylation of CHK1 and a greater sensitivity to replication stress in the 4N clones derived from the RKO cell line (**Figure 46A and B**).

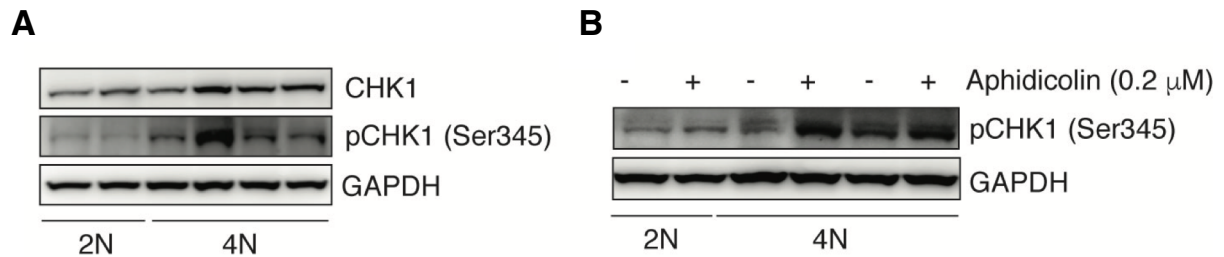


Figure 46. (A) Immunoblot analysis of phospho-CHK1 (pCHK1) levels in RKO 2N and 4N cells. Average ratio pCHK1/CHK1 of 0.23 for 2N and 0.75 for 4N cells. (B) Analysis of phospho-CHK1 levels after the addition of aphidicolin (0.2μM) for 24 hours to both sets of RKO clones. GAPDH was used as a protein loading control.

To further evaluate the levels of endogenous replication stress in 4N clones, sensitivity to the ATR inhibitor VE-821 was analyzed. Whereas 2N cells maintained approximately the same growth levels, 4N clones displayed a significant growth impairment following VE-821 treatment ($P < 0.0001$) (**Figure 47**), supporting the presence of replication stress in these cells.

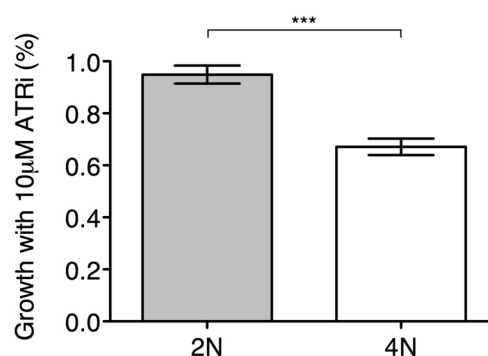


Figure 47. Graph depicting percentage of growth impairment of 4N compared to 2N cells as measured by colony formation assay after treatment with ATRi (10μM) for 24 hours ($n = 3$ / clone). Data are reported as mean $\pm$ SEM. *** $P < 0.0001$.

Next, we investigated whether the growth impairment previously observed in 4N cells resulted from a delay in cell cycle progression due to replication stress by pulse-labeling the cells with BrdU and analyzing them at two-hour intervals. Consequently, we observed a delay in S-phase for the 4N clones, which resulted in a slower progression into G2/M-phase, thus confirming the association between growth impairment and replication stress-dependent S-phase delay (**Figure 46**).

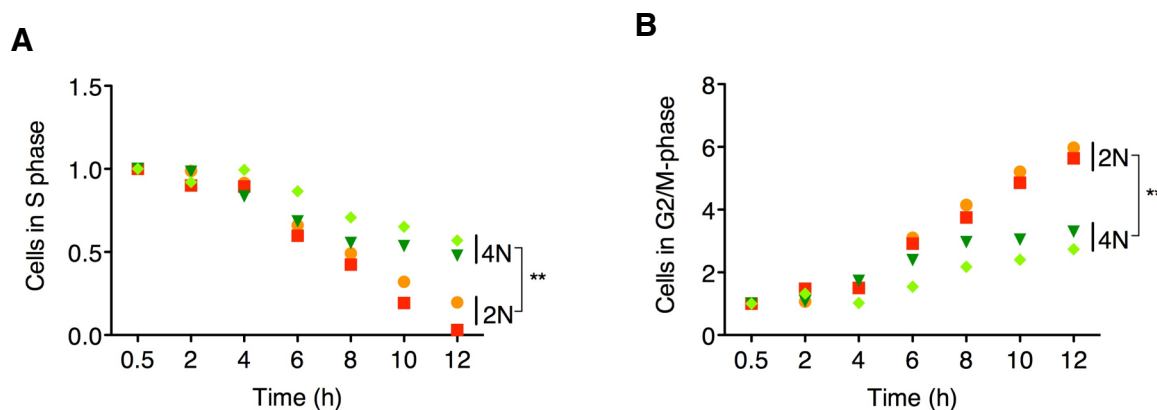


Figure 48. Plots showing time-course experiments to characterize the delay in the cell cycle. DLD-1 cells were pulse-labeled with BrdU and analyzed by FACS every two hours for a 12-hour period. Near-tetraploid cells exhibit a slower progression through S-phase (**A**) and a slower rate of entry into G2/M (**B**). Statistical analysis was performed on the trends using the R package geeppack for longitudinal data analysis. $**P < 0.001$.

Given our finding that mitosis-related genes were also overexpressed in 4N cells (**Figure 40**), we assessed whether a similar delay to that observed in S-phase occurred during cell division. To this end, we evaluated mitotic timing for both sets of clones by performing live-cell imaging. However, we did not observe any significant difference between the 2N and 4N clones (**Figure 49**), suggesting the cell cycle delay was specific to S-phase.

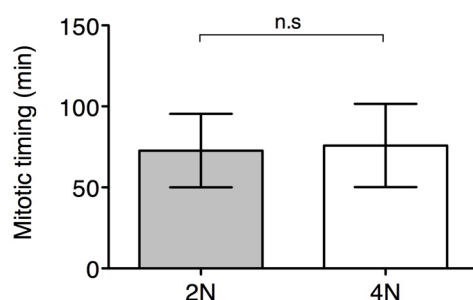


Figure 49. Histogram showing the mitotic timing of 2N and 4N DLD-1 cells quantified by live cell imaging. (n = 50 cells/ clone). Data are reported as mean $\pm$ SEM.

Taken together, our data demonstrated that 4N clones experience replication stress, which appears to be kept at a level compatible with cell survival, likely due to the concerted actions of S-phase checkpoint proteins. Nevertheless, such levels of replication stress could be detrimental to 4N cells upon treatment with ATR inhibitors.

Finally, in order to determine whether replication stress resulted in increased DNA damage, we used immunofluorescence staining for the DNA damage markers γ H2AX and 53BP1. We found that 4N cells displayed an approximately 2.7-fold and 2-fold increase in γ H2AX and 53BP1 foci *per nucleus*, respectively ($P < 0.0001$) (**Figure 50**).

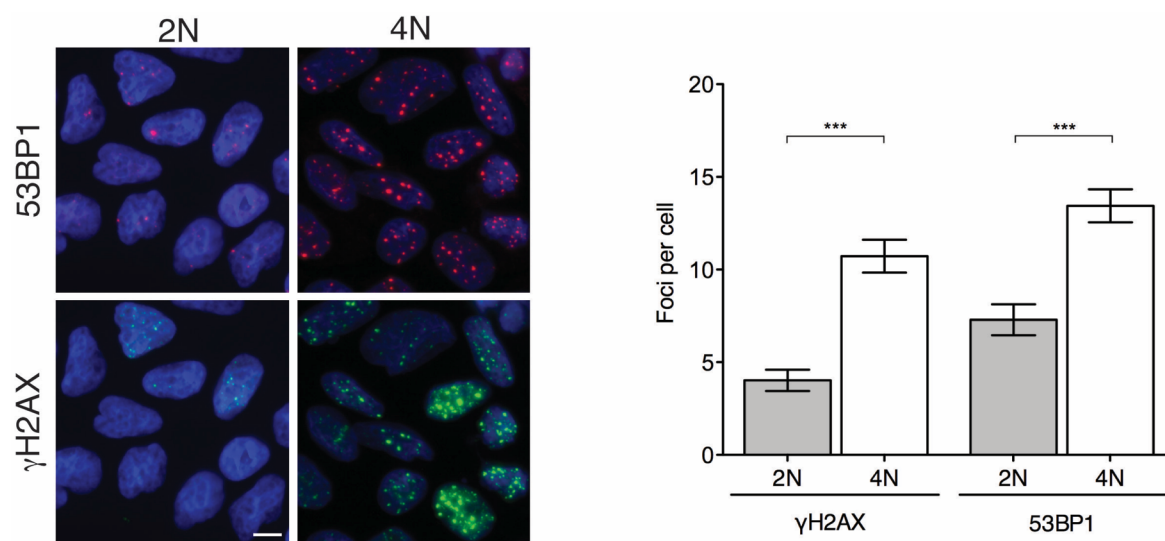


Figure 50. Left: Representative images of immunofluorescence for 53BP1 (red) and γ H2AX (green). DAPI was used for nuclear counterstaining. Right: Histogram depicting the number of foci *per cell* for γ H2AX and 53BP1 for 2N and 4N DLD-1 cells ($n > 240$ cells / clone). Scale bar = 20 μ m. Data are reported as mean $\pm$ SEM. *** $P < 0.0001$.

As these results were indicative of higher levels of DNA damage in the 4N clones, we sought to further delineate when this damage occurred. To do this, we specifically evaluated DNA damage in S- and M-phase by co-staining cells with EdU or phospho-H3 and γ H2AX. The levels of γ H2AX-positive cells were significantly higher in 4N vs. 2N clones during S-phase ($P < 0.001$) (**Figure 51**). In fact, almost all of the 4N cells were positive for DNA damage.

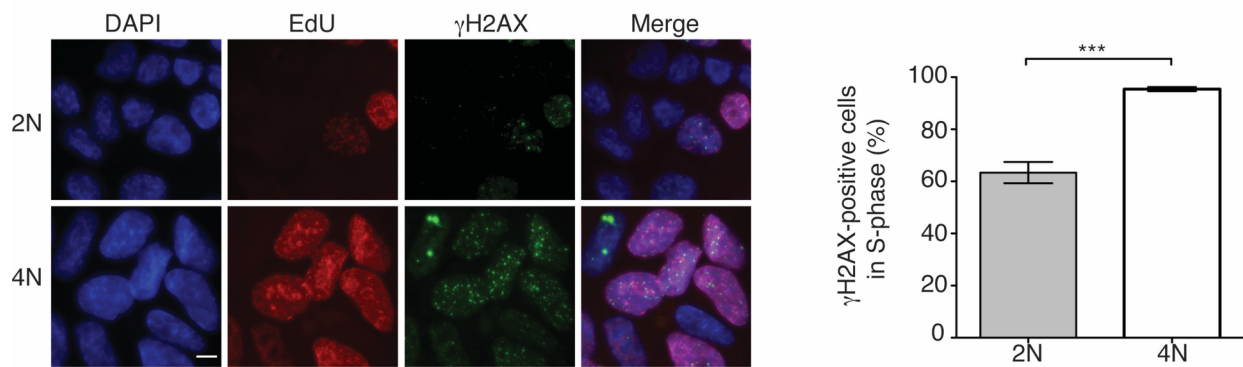


Figure 51. Left: Representative images of DAPI (blue), EdU (red), γ H2AX (green), and merged images in both 2N and 4N DLD-1 cells. Right: Plotted is the percentage of γ H2AX positive cells (> 15 foci) with positive EdU staining ($n > 240$ cells / clone). Scale bar = 20 μ m. Data are reported as mean $\pm$ SEM. *** $P < 0.0001$.

Additionally, when comparing the 4N with 2N cells we identified a 2.4-fold increase in the number of γ H2AX foci during M-phase ($P < 0.0001$) (**Figure 52**), suggesting that not all the DNA damage occurring in S-phase is repaired by the time cells reach M-phase.

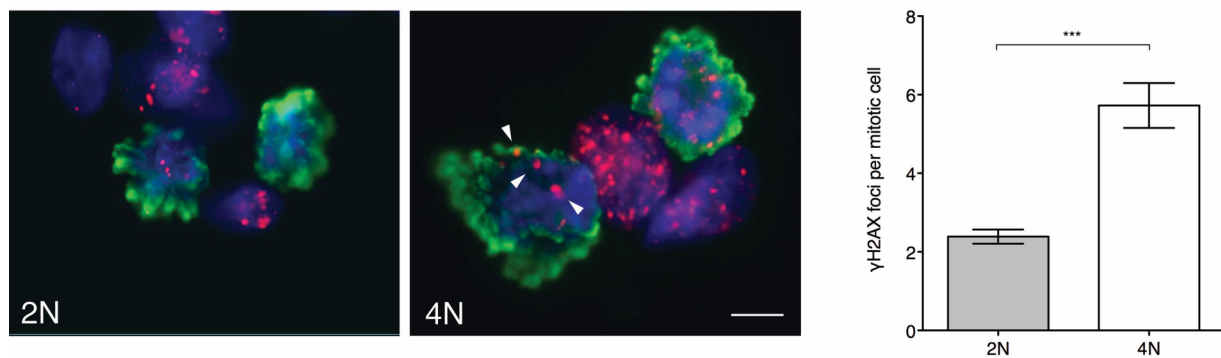


Figure 52. Left: Representative images of 2N and 4N DLD-1 cells in M-phase immunostained with γ H2AX (red), phospho-H3 (green) and DAPI (blue). Arrowheads point at γ H2AX foci in a prometaphase cell. Right: Quantitative analysis of γ H2AX foci *per* mitotic cell for two 2N and two 4N clones. Scale bar = 10 μ m. Data are reported as mean $\pm$ SEM. *** $P < 0.0001$.

4.4. Near-tetraploid cells show karyotypic heterogeneity and genomic instability

Considering that 4N cells overexpressed genes related to M-phase (**Figure 40**), yet progressed through mitosis with normal timing (**Figure 49**), we wondered whether certain aspects of cell division other than timing might also be affected. Accordingly, we examined the effect of ploidy on chromosome segregation and genomic instability by assessing karyotypic heterogeneity. By enumerating chromosome numbers we found, as expected, 2N clones exhibited rather stable chromosome contents (modal number of 45 or 46 chromosomes in DLD-1 and 49 chromosomes in RKO) whereas the 4N clones had a wider variability in chromosome numbers and tended to lose chromosomes in both lines, indicative of a higher degree of heterogeneity (**Figure 53A and B**).

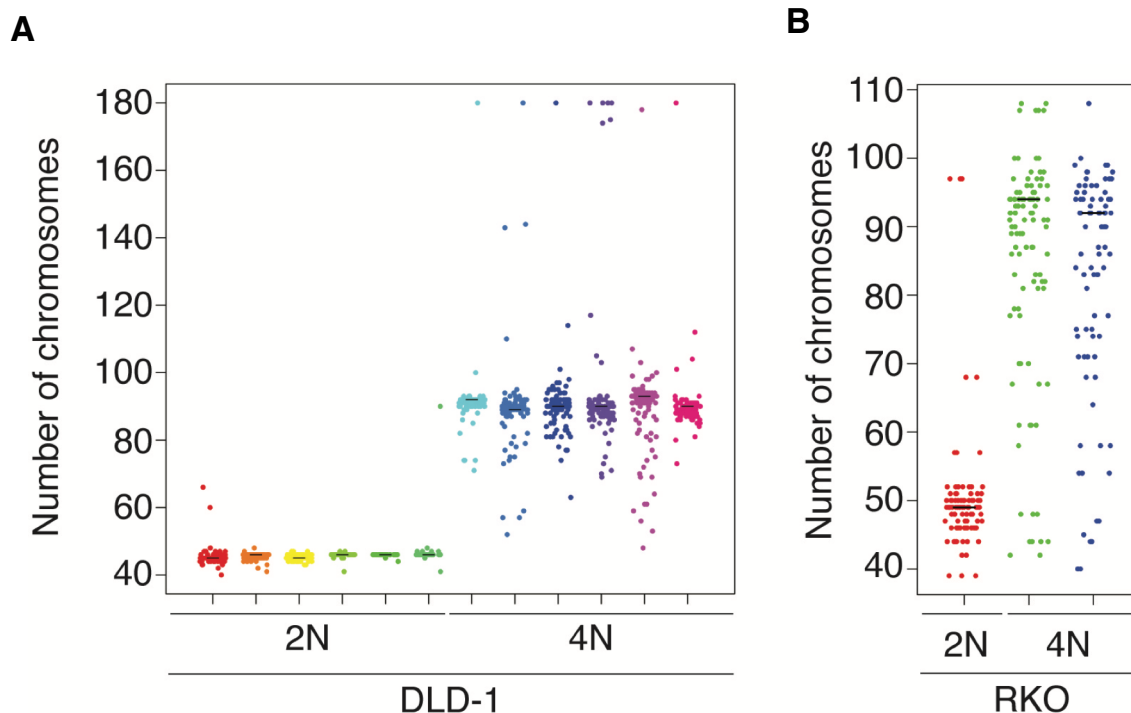


Figure 53. Assessment of intercellular genetic heterogeneity. Dot plot depicting the number of chromosomes in individual cells from (A) six 2N and six 4N DLD-1 clones and (B) one 2N and two 4N RKO clones. Black lines denote the modal chromosome number for each clone ($n = 100$ metaphases / clone).

This heterogeneity was then further validated by counting the number of FISH signals for four locus-specific probes for chromosomes 3, 7, 11 and 13 ($P < 0.0001$) in interphase DLD-1 cells and metaphase spreads (Figure 54A-D).

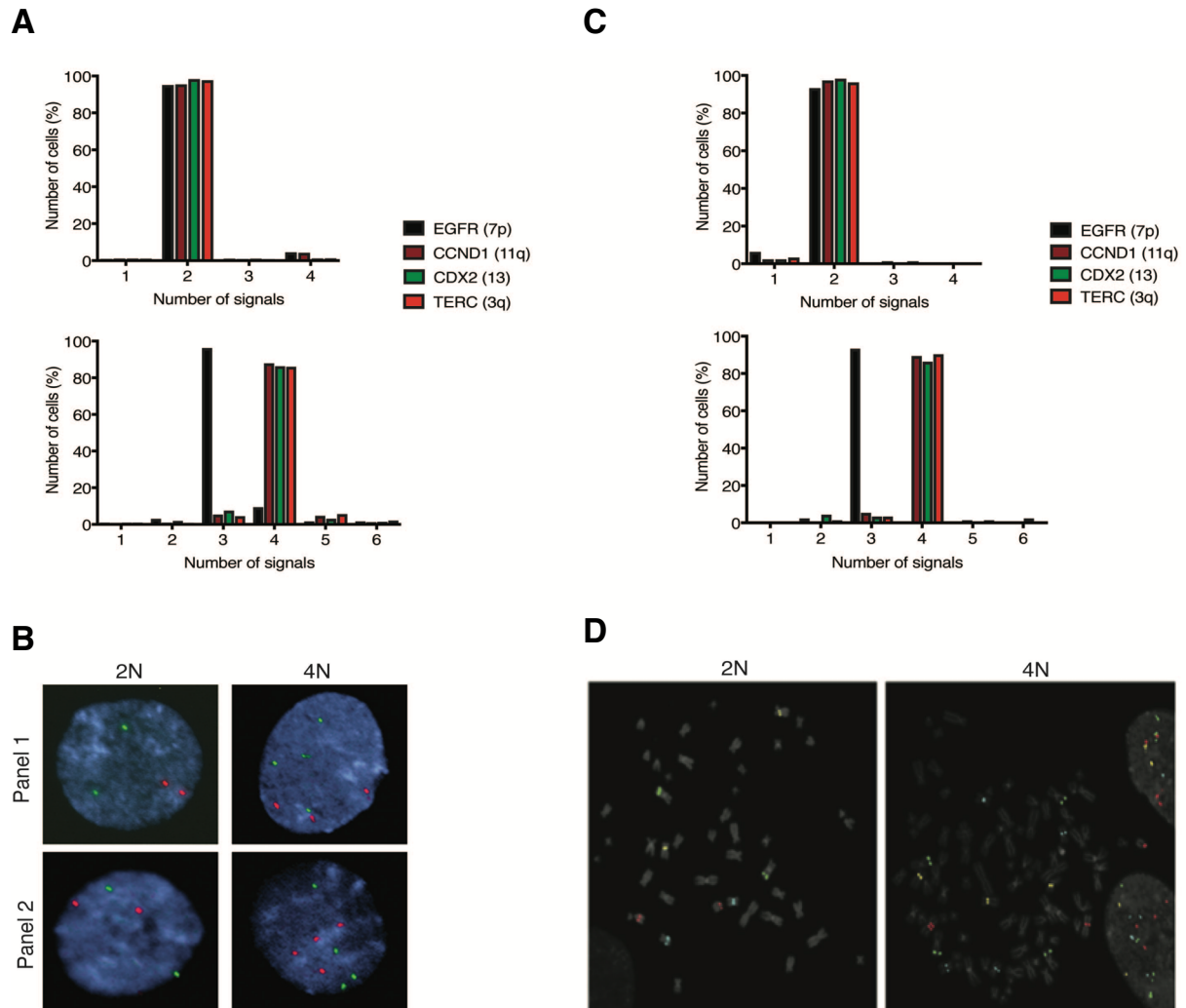


Figure 54. (A) The graphs report the percentage of cells with the corresponding FISH signals for 2N (top) and 4N (bottom) DLD-1 clones after a total of 1,000 nuclei were analyzed and show higher chromosomal variability in 4N cells. (B) Representative FISH images for both sets of clones. FISH panel 1 includes fluorescent-labeled BAC clones for *EGFR* (7p) and *CCND1* (11q), and FISH panel 2 includes fluorescent-labeled BAC clones *TERC* (3q) and *CDX2* (13q). (C) Histograms of the percentage of metaphase cells with the corresponding FISH signals for 2N (top) and 4N clones (bottom) show increased chromosomal variability in 4N cells ($n = 1,000$). (D) Representative metaphase FISH image for 2N (left) and 4N (right) clones.

To determine whether such heterogeneity in the chromosome number masked an underlying selection-induced karyotypic evolution of 4N cells, we performed genome profiling of early (p. 5-10) and late passage (p. 25-30) DLD-1 clones using aCGH. While none of the 2N clones displayed genomic imbalances different from those observed in the parental line, novel non-recurrent genomic imbalances were detected in 75% (nine out of twelve) of the 4N clones (**Figure 55**). However, we did not detect any karyotypic selection over time.

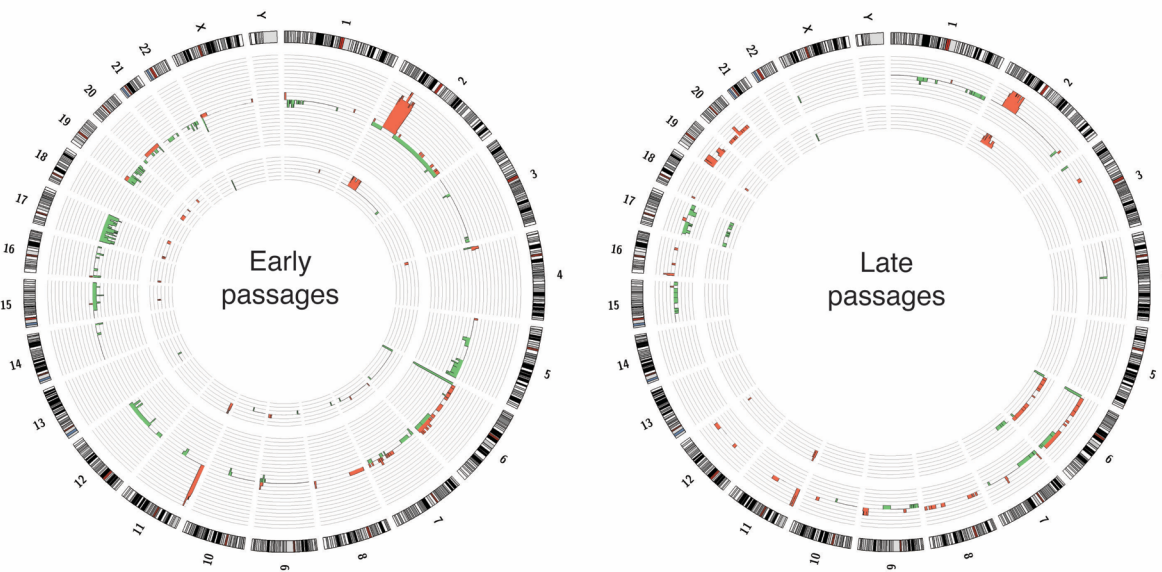


Figure 55. Circos plots illustrate the comparative frequency of copy number detected by aCGH in early DLD-1 passaged 2N (inner histogram; n = 3) and 4N (outer histogram; n = 12) clones, as well as in late passaged 2N (inner histogram; n = 3) and 4N (outer histogram; n = 5) clones. Human chromosome ideogram is laid out at the periphery of the circle. Red blocks indicate copy number gains and green blocks indicate copy number losses. aCGH, array comparative genomic hybridization.

Additionally, for a subset of 4N clones we characterized the presence of structural chromosome alterations by performing SKY (**Figure 56**). Our analysis revealed an average of five *de novo* structural aberrations *per* clone and maximum rates of instability (i.e., number of non-clonal aberrations / number of cells analyzed)²⁹⁷, ranging from 0.18 to 0.3 chromosome rearrangements *per* cell division in the 4N cells.

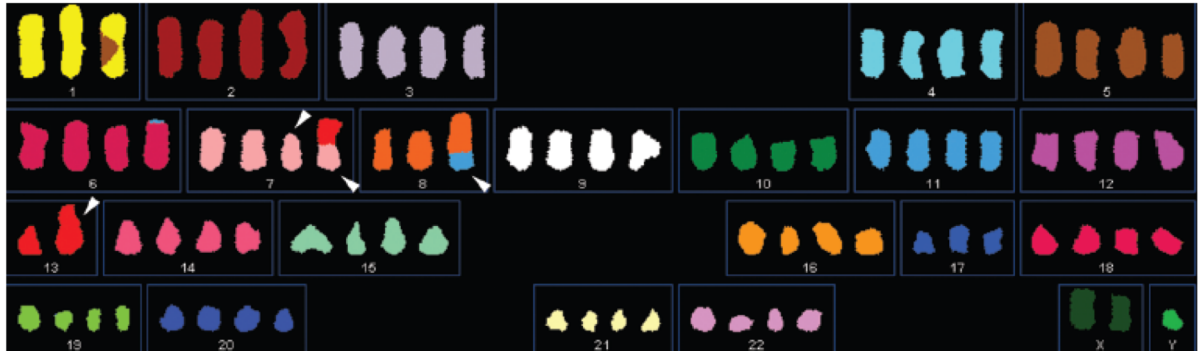


Figure 56. Representative image from SKY analysis performed in DLD-1, indicating with an arrowhead the presence of *de novo* structural chromosomal aberrations in 4N cells. SKY, spectral karyotyping.

The large amount of intercellular heterogeneity and karyotypic changes observed in 4N cells suggested high levels of CIN. To assess this further, we evaluated the formation of micronuclei (MNi) and determined the types and rates of mitotic defects. First, we performed a cytokinesis-block micronucleus assay to evaluate the formation of MNi in binucleated cells and found that 4N cells displayed significantly higher rates of MNi ($P = 0.0001$), which was validated in RKO 4N clones (**Figure 57**).

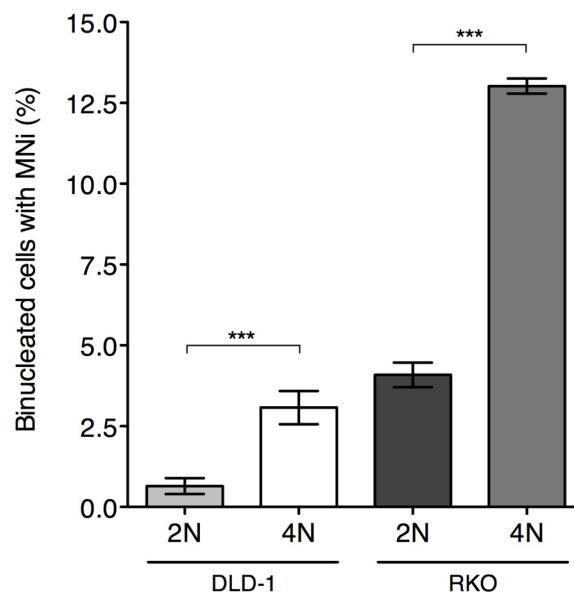


Figure 57. Bar graph showing the percentage of MNi in 2N and 4N binucleated DLD-1 and RKO cells ($n = 1,000$). Data are represented as mean $\pm$ SEM. *** $P < 0.0001$.

MNi can arise from a number of different chromosome missegregation events. Therefore, in order to better characterize the possible underlying causes of CIN in 4N clones, we immunostained the kinetochores and microtubules of mitotic cells (**Figure 58**), and quantified the mitotic defects during prometaphase and anaphase. We found that 4N cells exhibited a 2.3- and 4-fold increase in the levels of both prometaphase and anaphase defects, respectively, compared to their 2N counterparts ($P < 0.0001$ and $P < 0.0001$, respectively) (**Figure 59A and B**).

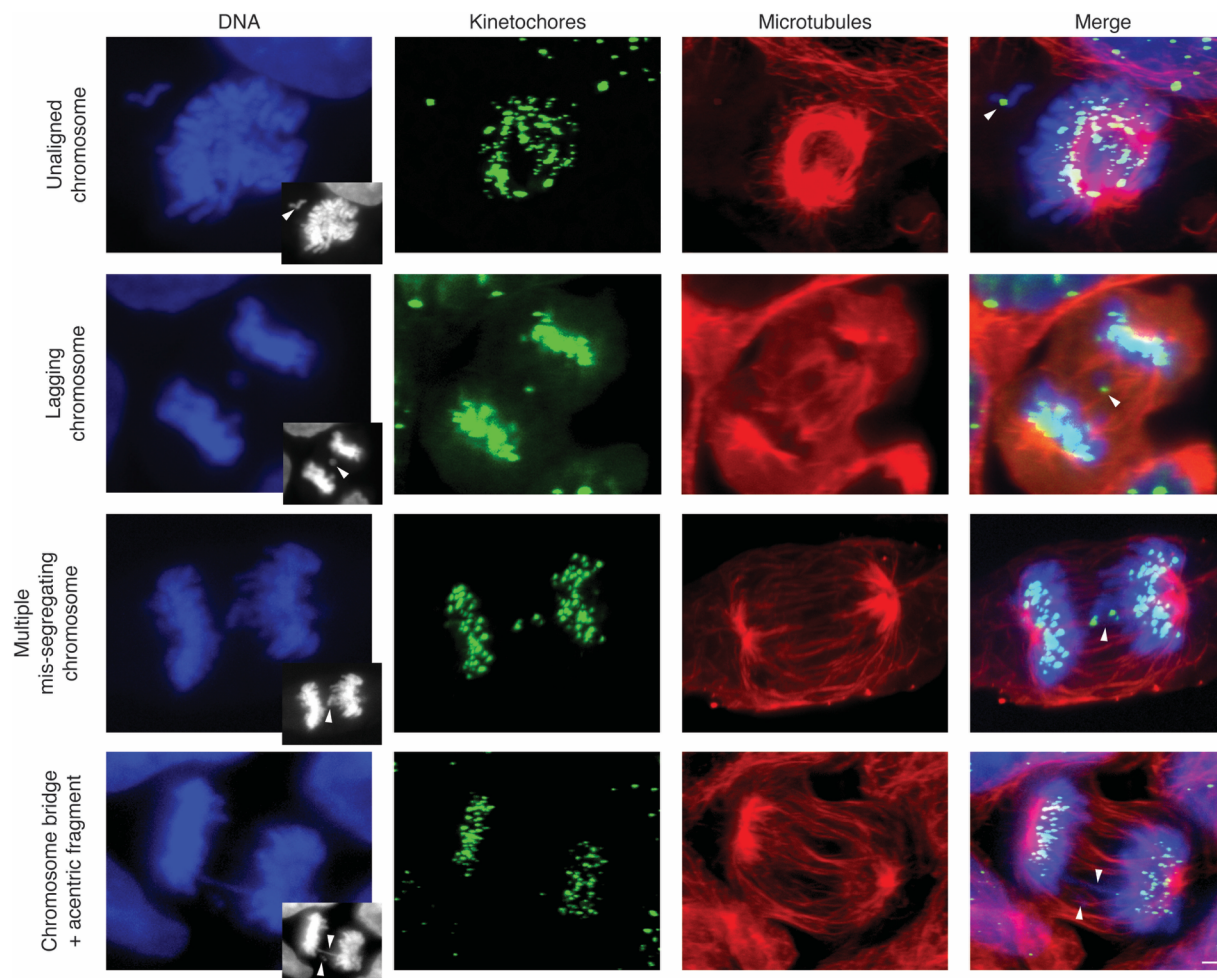


Figure 58. Examples of mitotic defects in prometaphase (top row) and anaphase (rows 2-4) cells immunostained for kinetochores (green) and microtubules (red). The images show examples of the most frequently observed defects (unaligned chromosomes, lagging chromosomes, multiple missegregating chromosomes, and chromosome bridges and acentric fragments). Grey scale images at the bottom right corners of each of the DNA images are single focal planes of DAPI-stained chromosomes shown for easier visualization of the mitotic defects. Arrowheads in the grey scale images and merge panels point to the specific mitotic defect. Scale bar = 10 μm .

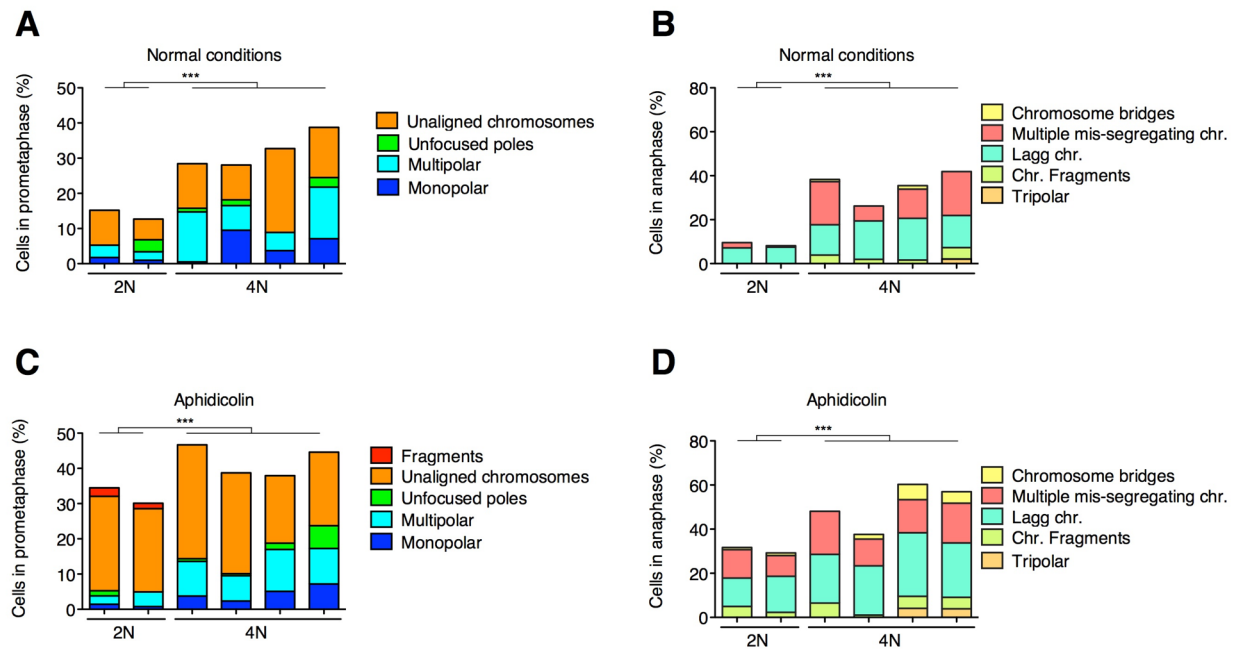


Figure 59. Frequencies of prometaphase (A) and anaphase defects (B) in normal culture conditions for 2N and 4N DLD-1 cells. Frequencies of prometaphase (C) and anaphase defects (D) in 2N and 4N DLD-1 cells treated with aphidicolin (0.2 μ M) for 24 hours. *** $P < 0.0001$.

Specifically, our data revealed unaligned chromosomes to be the most frequent defect in prometaphase (Figure 58, top row and 59A), while anaphase lagging chromosomes (Figure 58, second row and 59B) and multiple missegregating chromosomes (Figure 58, third row and 59B) were observed at a higher frequency during anaphase. The latter chromosome segregation defect consisted of two or more chromosomes lagging behind with no evidence of merotelic attachments, and always displayed kinetochores that faced each other on the two sides of the spindle equator. Moreover, the DNA from these missegregated chromosomes did not appear to be well separated in anaphase. Intriguingly, this defect was mainly found in 4N cells and was rarely observed in the 2N clones. When we assessed the major anaphase defects (i.e., chromosome bridges, lagging chromosomes, and chromosome fragments) in RKO clones, we confirmed that 4N cells exhibited greater levels of anaphase defects than 2N clones (2.5-fold, $P < 0.0001$) (Figure 60). In fact, lagging chromosomes were the most frequently observed mitotic defect in both of our *in vitro* models.

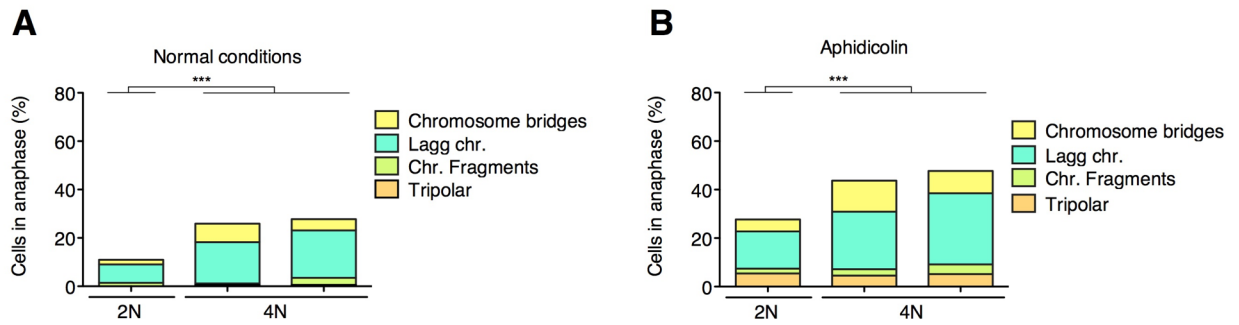


Figure 60. Frequency of anaphase defects in 2N and 4N RKO cells under normal culture conditions (A) and treated with aphidicolin (0.2 μ M) for 24 hours (B). *** $P < 0.0001$.

As it is known that anaphase lagging chromosomes can arise from cells undergoing a transient multipolar stage during prometaphase due to the presence of supernumerary centrosomes^{165, 167}, we sought to count the number of centrosomes in G1 cells by co-immunostaining of γ -tubulin and cyclin D1. Accordingly, we found that 4N clones displayed a larger subpopulation of cells with extra centrosomes compared to 2N clones (mean 14.8 vs. 5.71, $P < 0.0001$ for DLD-1 cells and mean 19.7 vs. 5.66, $P < 0.0001$ for RKO cells) (Figure 61).

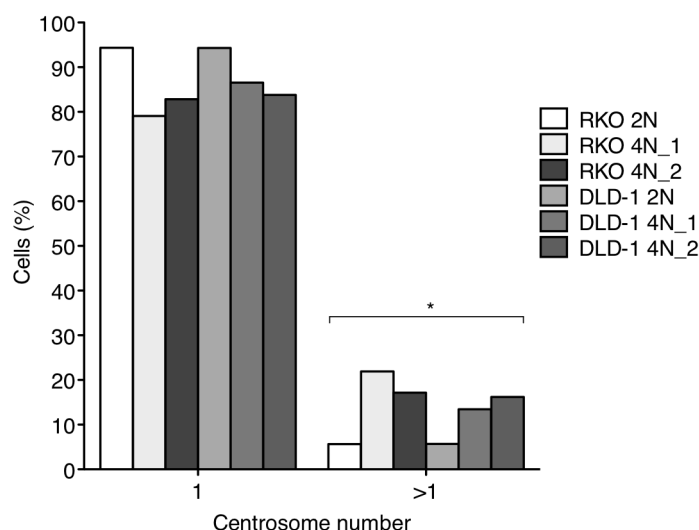


Figure 61. Immunofluorescence staining with anti- γ -tubulin and anti-cyclin D1 antibodies of DLD1 and RKO cells was quantified. Levels are represented as percentage of cells. * $P < 0.05$.

This finding is consistent with the low number of multipolar prometaphases (**Figure 59A**) and can explain, in part, the higher rates of lagging chromosomes observed in 4N clones. Taken together, these results indicate that 4N cells display high rates of chromosome missegregation, which are responsible for the high levels of aneuploidy.

4.5. Replication stress triggers genomic instability

The observation that the major phenotypes of 4N cells are replication stress and genomic instability suggested the possibility of a link between the two. To further examine this, we sought to determine the levels of CIN in response to aphidicolin-induced replication stress in 2N and 4N clones. Accordingly, we evaluated the formation of MNi after aphidicolin treatment and found that aphidicolin significantly increased the frequency of MNi in 2N and 4N cells from both DLD-1 and RKO cell lines ($P < 0.001$ in both cases) (**Figure 62**).

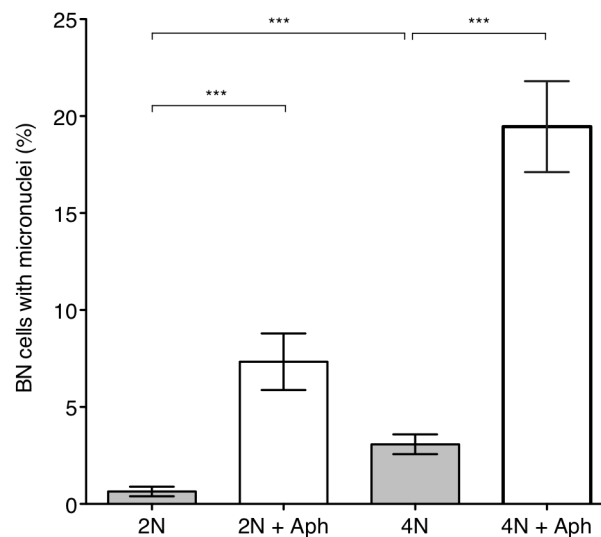


Figure 62. Bar graph showing the percentage of MNi in aphidicolin-untreated and treated 2N and 4N binucleated DLD-1 cells ($n = 1,000$). The aphidicolin treatment was performed at $0.2\mu\text{M}$ for 24 hours. Statistical analysis was performed using a linear model with time and ploidy (4N vs. 2N) as predictors. Data are represented as mean $\pm$ SEM. *** $P < 0.0001$.

Additionally, we found an overall increase in abnormal prometaphases and anaphases in DLD-1 cells ($P < 0.0001$) following aphidicolin treatment. Similarly, significant increases were found in unaligned chromosomes ($\bar{x} = 7.9\%$ vs. 25.2% , $P < 0.0001$ for 2N clones and $\bar{x} = 15.2\%$ vs. 25.3% , $P < 0.0001$ for 4N clones) and lagging chromosomes ($\bar{x} =$

7.33% vs. 14.6%, $P = 0.03$ for 2N clones and $\bar{x} = 16.2\%$ vs. 24.5% $P = 0.005$ for 4N clones) for both 2N and 4N DLD-1 cells (**Figure 59C and D**). Interestingly, we also found that the multiple missegregating chromosome phenotype appeared in the 2N cells upon aphidicolin treatment (**Figure 59D**), suggesting a direct link between replication stress and this chromosome segregation defect. As predicted, higher numbers of chromosome bridges and fragments (**Figure 58**, bottom row) were also identified ($\bar{x} = 0\%$ vs. 4.7% , $P = 0.01$ for 2N clones and $\bar{x} = 3.8\%$ vs. 8.1% , $P = 0.02$ for 4N clones) (**Figure 59D**), which confirmed the initial hypothesis for this experiment. Also, when we assessed the levels of CIN in the RKO clones treated with aphidicolin, we found that they also showed a higher frequency of not only chromosome bridges and fragments but also lagging chromosomes in both the 2N and 4N cells ($P < 0.02$ and $P < 0.01$, respectively) (**Figure 60B**), thus supporting our findings in DLD-1 cells. Collectively, these data indicate that aphidicolin-associated replication stress induces chromosome segregation errors, therefore suggesting that replication stress may contribute to CIN in 4N cells.

4.6. Near-tetraploid cells exhibit enhanced migratory and invasive capabilities in primary tumors

In order to investigate whether whole-genome duplication events confer physiological changes sufficient to affect cellular behavior, we assessed the migratory and the invasive capacities of 2N and 4N cells using a transwell (Boyden chamber) migration assay. DLD-1 near-tetraploid clones displayed increased migratory capacity compared to 2N clones ($P < 0.0001$) (**Figure 63A**). Similar results were obtained in wound-healing assays, in which the average closure area after 48 hours was 13.88 vs. $28.20 \mu\text{m}^2$ for 2N and 4N DLD-1 clones, respectively ($P = 0.05$) (**Figure 63B**). At late passages, 4N cells exhibited even greater migration capabilities ($P < 0.0001$) (**Figure 63B**). Interestingly, when a layer of Matrigel was added to the Boyden chamber in order to assess invasiveness, 4N cells displayed enhanced invasive capabilities ($P = 0.04$) (**Figure 64**).

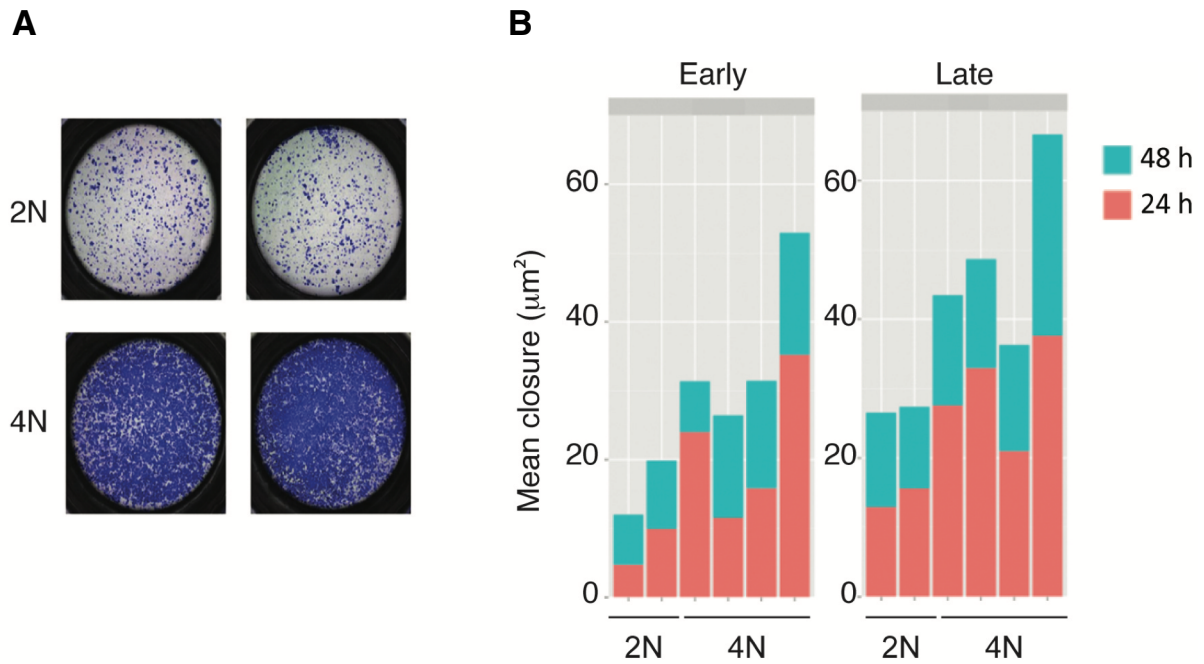


Figure 63. (A) Boyden chamber assay representative images showing the migration capability of the 2N and 4N DLD-1 clones. (B) Plotted is the mean area of the scratch closure resulting from the wound healing assay at 24 and 48 hours post-scratch for both early (p. 5-10) and late (p. 25-30) passaged 2N and 4N clones ($P < 0.05$ and $P < 0.0001$, respectively).

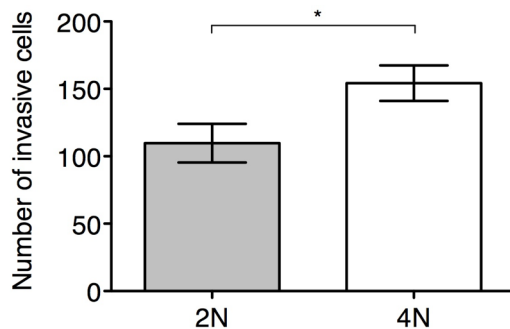


Figure 64. Bar graph showing the quantification of Matrigel invasiveness for 2N and 4N cells. Data are represented as mean number of invasive cells $\pm$ SEM ($n = 3$). $*P < 0.05$.

On the other hand, excluding the possibility that the increased size of the 4N cells was associated with mobility, RKO 4N cells showed highly heterogeneous migratory capacities during the wound healing assays. As extra centrosomes have been previously associated with cellular mobility³⁰⁴, we determined the presence of extra centrosomes in the migratory front of the wound-healing assay; however, our results showed that the migratory capacity of these cells was independent of the centrosome number (Figure 65).

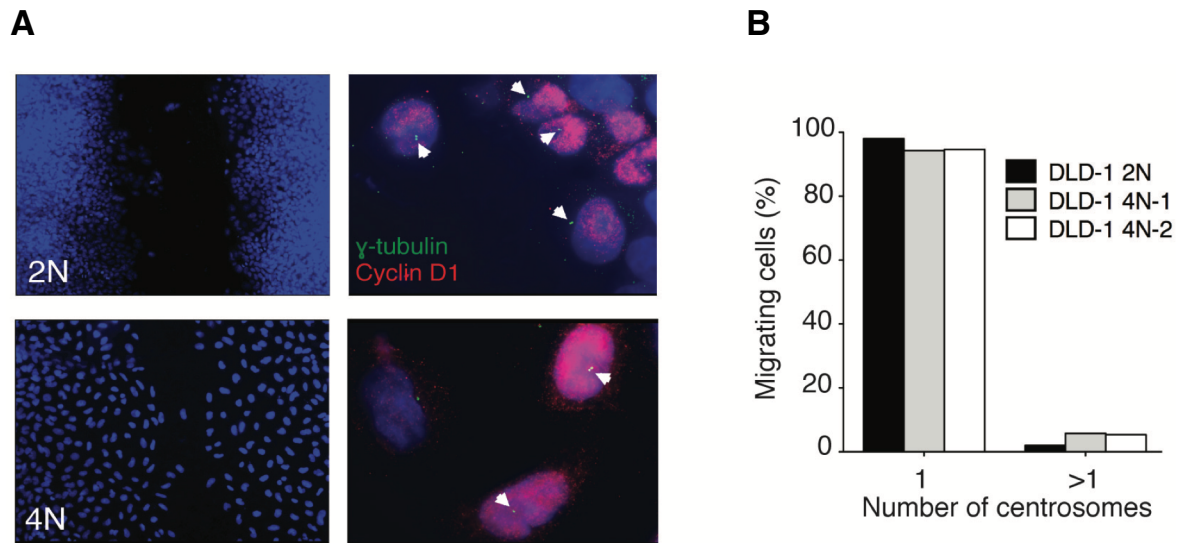


Figure 65. (A) Illustration of the co-immunostaining of γ -tubulin and Cyclin D1 to quantify centrosome number. (B) Representation of the centrosome number in migrating cells during wound healing assay (n = 250 - 350).

Based on these observations, we sought to determine whether these *in vitro* findings could be confirmed *in vivo* and have biological significance by analyzing primary CRC samples. Towards this end, we compared the levels of aneuploidy in the main tumor mass of colon adenocarcinomas with that of their invasive fronts utilizing centromeric FISH probes for two chromosomes unlikely to be involved in copy number changes in CRC (i.e., chromosomes 4 and 6). We found that invasive fronts displayed higher levels of aneuploidy for these two chromosomes ($P < 0.0001$), and there was an enrichment of cells with four or more FISH signals for the two chromosomes under study, suggesting the presence of highly chromosome unstable cells ($P < 0.0001$, for the area under the curve for signals ≥ 4) (**Figure 66A and B**).

Moreover, we observed that the invasive front exhibited a modest 2.6-fold increase in the number of cells with extra centrosomes compared to the main tumor mass ($P < 0.04$) (**Figure 67**), therefore suggesting that *in vivo*, both whole-genome duplication and extra centrosomes are important for tumor cell invasion.

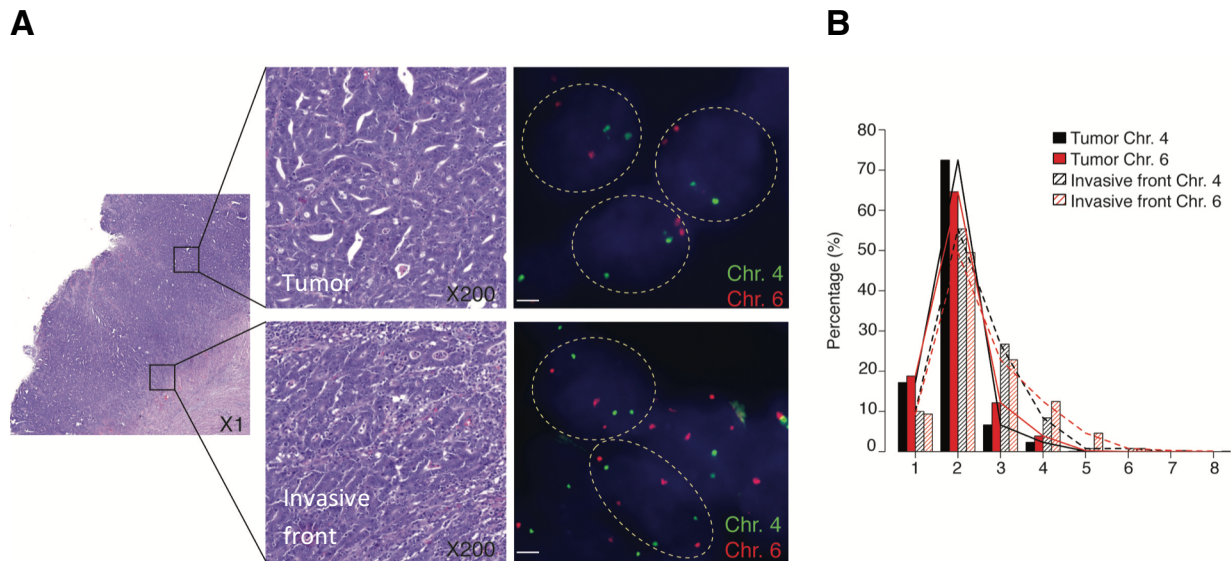


Figure 66. (A) Microscopic assessment of FISH signals using fluorescent probes for centromeres corresponding to chromosomes 4 (FITC) and 6 (Cy3) in a colorectal primary tumor (top) and its associated invasive front (bottom). Single isolated nuclei are outlined with dashed circles. All three nuclei from the primary tumor display two signals *per* probe, whereas three or four signals *per* probe were detected in the nuclei from the invasive front. **(B)** Distribution of centromere probes signals for chromosome 4 and chromosome 6 in nine colorectal primary tumors and their corresponding invasive fronts ($P < 0.0001$).

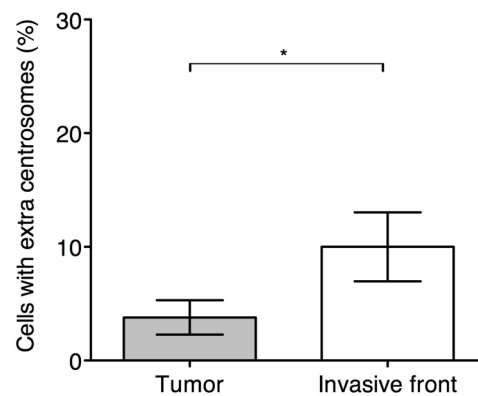


Figure 67. Percentage of cells with extra centrosomes in colorectal primary tumor samples and their corresponding invasive fronts. $*P < 0.05$.

DISCUSSION

1. Aberrant crypt foci and the need for colorectal cancer biomarkers

Colorectal cancer (CRC) often develops over a period of a decade and follows a step-wise progression known as the adenoma-carcinoma sequence¹⁷; thereby letting a perfect scenario for screening and prevention that would result in a decrease in CRC mortality³⁰⁵. In fact, patient survival is highly dependent on the tumor's stage at the time of diagnosis. Of all early lesions, adenoma is the best-known precursor lesion for CRC, and although only about 5% of them progress to carcinoma, the majority of CRC cases arise from adenomas^{49, 51}. Colonoscopy is the gold standard for CRC prevention because resection of precursor lesions (i.e. polypectomy) has demonstrated to reduce CRC mortality³⁰⁶. Nevertheless, a primary prevention strategy, which would prevent the formation of precursor lesions, has not been established yet. Therefore, the identification of biomarkers is crucial for classifying CRC risk groups as well as evaluating the effect of prevention agents as a primary prevention. However, identifying reliable CRC biomarkers has been a persistent challenge^{83, 307} and there is growing interest in finding other lesions that can be used as such.

Over the past few decades, aberrant crypt foci (ACF) have emerged as promising surrogate biomarkers for CRC. They are the first identifiable lesions in the colon, arising following the treatment with colon-specific carcinogens and observed in patients at high-risk for CRC^{19, 20}. Since their discovery, several studies have revealed that both the number and size of ACF are associated with an increased risk of CRC^{25, 31, 308, 309}. Moreover, the presence of dysplasia in ACF has also been reported to be a contributing factor to CRC predisposition³⁴. Additionally, genetic and epigenetic alterations commonly detected in CRC have been found in ACF, supporting their role in colorectal carcinogenesis^{36, 37, 42, 108}. Despite this, the conclusions of these studies remain contentious due to the absence of well-defined patient inclusion criteria and differences in sample size and the techniques used²⁸. Therefore, we sought to characterize a broad cohort of rectal ACF samples from different CRC risk patients in order to determine whether they are precursor lesions and suitable biomarkers for CRC risk. Unfortunately, our results indicated that ACF might solely be preneoplastic lesions but not a reliable biomarker for CRC since we observed inconsistent morphological and molecular findings together with a lack of correlation with CRC risk groups.

1.1. The importance of endoscopic and histological assessments of ACF

The prevalence and number of ACF have been described to progressively increase according to CRC risk^{23, 308, 309}; however, the absence of a well-defined subject sampling frame has raised some concerns about these epidemiological studies²⁸. In fact, in agreement with multi-center and well-standardized studies that have challenged these initial findings^{29, 30, 32}, we did not observe an association between ACF number and the past or future CRC risk in our study. Moreover, the size of ACF increases in an age-dependent manner^{29, 30, 32} and has been reported as a predictive factor of CRC²⁵. In accordance with these observations, we found the number of large ACF to be associated with CRC risk group, whereas none of the other endoscopic features showed to be indicative of CRC. Nevertheless, the lack of association of this endoscopic finding with a specific histopathology diagnosis or molecular alteration provides this observation of uncertain clinical utility.

On the other hand, results from studies examining ACF histology have been inconsistent^{310, 311} due to variability in ACF classification and tissue sampling proceedings. In 2010, the World Health Organization (WHO) established the classification of ACF as either dysplastic or hyperplastic³¹² and in several studies the presence of dysplasia in these lesions have been reported to be predictive of CRC^{23, 25, 34}. However, the prevalence of dysplasia is highly variable even in patients with CRC^{23, 26, 308} and some studies did not even identified dysplastic ACF^{29, 313}. We mainly observed hyperplastic ACF and detected 18.5% of dysplastic ACF. Nonetheless, we did not observe a correlation between dysplastic ACF and CRC risk. In this sense, we found a poor inter- and intra- observer concordance in the histopathological classification of ACF, which highlights the need of a more accurate standardization of the histological diagnosis to use ACF as a precursor lesion or potential biomarker. Nevertheless, the small size of the samples and the lack of orientation during sampling could have influenced our results too.

1.2. Genetic and epigenetic findings in ACF

1.2.1. *KRAS* mutations are commonly found in ACF but are not predictive of CRC

An important consideration when evaluating precursor lesions for CRC is delineating the genomic alterations (genetic and epigenetic) that are found during colorectal carcinogenesis. *APC* gene mutations are one of the earliest and most common genetic alterations during CRC tumorigenesis⁴⁸. In fact, they are observed in both adenomas and carcinomas⁷⁷. However, although several studies have assessed *APC* mutational status in ACF samples, collectively they have concluded these alterations to be infrequent in sporadic ACF^{37, 38, 314} and mainly detected in lesions from FAP patients³⁷. Similarly, *APC* was mutated in only 1.7% of our ACF samples, indicating that in contrast to what happens in adenomas and adenocarcinomas, *APC* gene mutations are not related with ACF formation.

Likewise, *BRAF* mutations are also described during colorectal carcinogenesis, although not as commonly as *APC* mutations. Indeed, they are closely related to serrated lesions, which typically arise in the proximal colon^{315, 316}. Accordingly, mutations in *BRAF* have been observed in ACF with either proximal location or from patients with serrated background³¹⁵⁻³¹⁷, but not in sporadic rectal ACF^{36, 318}, which is in agreement with our observations. Moreover, microsatellite instability (MSI) has been described as the main characteristic of one of the classical pathways driving CRC tumorigenesis. Nevertheless, the presence of MSI in ACF remains controversial. Whereas MSI has been modestly detected in ACF^{108, 319, 320}, it was not found in any of our ACF samples. In fact, the most significant association of MSI with ACF has been reported in Lynch syndrome patients³²¹. Collectively, these results indicate that MSI and *BRAF* mutations are not frequent alterations in ACF.

In contrast to the other previously discussed gene alterations, *KRAS* mutations are consistently reported in ACF samples. Nonetheless, the mutation's rates are discrepant^{37, 38}, possibly due to the different types and sizes of study populations amongst studies. Overall, 16.3% of our ACF samples exhibited *KRAS* mutations although their frequencies were equally distributed within CRC risk groups, in accordance with previous results^{25, 37}. Therefore, these findings indicate that *KRAS* mutations might contribute to ACF formation but they do stratify ACF according to CRC risk.

1.2.2. Enhanced *MGMT* hypermethylation is associated with *KRAS* mutations in ACF

Together with genetic alterations, epigenetics also drive malignant transformation. In the case of CRC, the so-called CIMP pathway is characterized by the hypermethylation of tumor-suppressor and tumor-related genes. *MGMT* is a gene whose promoter methylation is implicated in CRC development since early stages^{118, 119}. Moreover, the hypermethylation of this gene is closely related with *KRAS* mutations^{322, 323}. Our results showed the association between *KRAS* mutations and *MGMT* promoter hypermethylation in ACF, indicating these molecular alterations might be responsible for ACF formation. Nevertheless, those ACF with the combined presence of *KRAS* mutations and *MGMT* promoter hypermethylation were not predictive of CRC risk.

1.2.3. LINE-1 hypomethylation does not occur in ACF and does not depict a field defect for CRC

Tumors usually exhibit not only hypermethylation of individual gene promoters but also global hypomethylation, which gradually increases during carcinogenesis, thereby increasing genomic instability^{99, 103}. The levels of LINE-1 methylation are considered a surrogate marker for global hypomethylation^{111, 112}. Indeed, the hypomethylation of LINE-1 has been stated as a field defect for CRC and is found in the early stages of this type of cancer^{113, 116, 324, 325}. We did not find increased LINE-1 hypomethylation in the normal mucosa of CRC patients, thus supporting an argument against the field defect for CRC. Instead, we observed differential LINE-1 methylation depending on the colon segment, in agreement with what has been recently described³²⁶. Similarly to our results, a previous study described LINE-1 hypomethylation field defect for patients with synchronous CRC but not for those with solitary CRC¹¹⁶, while others reported that global hypomethylation is highly variable in cancer³²⁷. Moreover, other authors observed the field defect but only in patients with serrated polyps³²⁴. Additionally, although LINE-1 hypomethylation has been described in the early stages of CRC^{113, 324}, none of our ACF samples exhibited LINE-1 hypomethylation, indicating that ACF formation is not triggered by this phenomenon and suggesting that genome-wide hypomethylation occurs after appearance of ACF in the colon.

2. The role of chromosome copy number alterations in the adenoma-to-adenocarcinoma sequence

Understanding the molecular changes that lead an adenoma to progress into an adenocarcinoma is critical for establishing effective CRC biomarkers. In this regard, in contrast to the initially proposed pattern of clonal evolution²³⁵, a growing body of evidence suggests that cancer progresses through a branched pattern of evolution. This theory thus implies the coexistence of different subclones within an individual tumor and leads to intratumor heterogeneity^{236, 241-244, 328-330}. Indeed, single cell analysis has shed light on the extent to which intratumor heterogeneity plays a role in tumor initiation and disease progression^{246, 247, 331}. Moreover, chromosomal instability (CIN) and abnormal karyotypes are so ubiquitous in tumors that aneuploidy itself has become a hallmark of cancer³³². In fact, chromosome copy number alterations (CNAs) are known to play a central role in CRC evolution, first appearing during the early stages of CRC progression^{244, 251}. Accordingly, we sought to evaluate the CNAs implicated in the transformation of adenoma-to-adenocarcinoma by performing single cell sequential multi-FISH in non-advanced adenomas (N_AD) and in the adenoma (AD) and adenocarcinoma (ADK) regions of progressed adenomas. Our results revealed that chromosome imbalances were already present at the earliest stages of adenomas and increased during the sequence of colorectal carcinogenesis, providing the fuel for tumor evolution. In progressed adenomas, we identified the gain of chromosomes 7p, 13q and 20q to be the drivers of both the AD-to-ADK transition as well as cancer progression during CRC's branched evolution. Moreover, whole genome duplication (WGD) events were detected not only during ADK progression, but also during the evolution of AD, leading to extensive subclonal diversification. In fact, for some cases a WGD event was the driver of the adenoma transformation into carcinoma.

2.1. Copy number alterations result in heterogeneity during CRC evolution

Intratumor heterogeneity has been revealed in several cancer genomic studies assessing gene mutations, DNA methylation and CNAs^{236, 242, 244, 247, 329, 331, 333, 334}. These findings indicate that the levels of genetic heterogeneity found in a tumor can be driven by different mechanisms, including CIN, which is commonly reported in CRC and can promote tumor initiation, progression and response to therapy^{122, 335-339}. Moreover, another recent

study found the patterns of CNAs to be the source of CRC heterogeneity and define the disease's landscape better than gene mutations³⁴⁰. Similarly, we observed the chromosome imbalances and heterogeneity increased in a stepwise manner from early lesions to ADK, recapitulating the evolution of CRC and even differentiating non-advanced adenomas from those that have progressed into carcinoma. Indeed, previous studies analyzing the bulk of samples excised from colorectal lesions have shown the recurrence of chromosomes alterations to be higher in adenocarcinomas than in adenomas³⁴¹.

2.2. Specific patterns of chromosomal alterations are selected during CRC

When considering specific CNAs observed in CRC, the gains of 7p, 13q and 20q, and the loss at 18q were the most common events^{252, 341-346}. In fact, a recent study defined 20q as the most recurrent chromosome gain in CRC, while 18q was the most frequently lost³⁴⁰. The presence of genes with central roles in cancer, such as oncogenes or tumor suppressor genes, explain the alterations of these genomic regions^{337, 344}. The gain of chromosome 20q has been described as being involved in the AD-to-ADK transition^{252, 341} and, perhaps unsurprisingly, it contains several genes with reported oncogenic function. For example, *AURKA* is located on 20q13.2 and is related to the regulation of chromosome segregation and cell growth through its function in microtubules and centrosomes. Another oncogene located in the long arm of chromosome 20 is *ZNF217*, whose overexpression correlates with the enhanced cell migration and invasion observed in CRC^{347, 348}. On the other hand, approximately 70% of CRC are found to exhibit the allelic loss of 18q.21, although the rates of this copy number deletion remain relatively low³⁴⁹. This chromosome contains the tumor suppressor genes *SMAD2* and *SMAD4*, which inhibit tumor growth and invasion through the regulation of the factor- β pathway, and are also reportedly mutated at a low frequency in CRC³⁵⁰. In agreement with these findings, the gain of 20q was the most recurrent chromosome imbalance found in our cohort of progressed adenomas, while the loss of 18q was the most frequent deletion, although observed at a much lower frequency than other chromosome imbalances. Conversely, chromosome 20q exhibited the highest frequency for both gains and losses in non-advanced adenomas, suggesting there was not selection for specific chromosome patterns during that stage of CRC. We also found the imbalances in chromosomes 7p and 13q to be important during CRC. Chromosome 7p contains genes that influence the RAS and WNT pathways, such as *EGFR*, and other genes related with disease

progression and prognosis, such as *PODXL*, *MET* or *WNT2*³⁵¹⁻³⁵⁵. Furthermore, *CDX2*, *CDK8* and *LN2* are genes associated with CRC progression and located on chromosome 13q³⁵⁶⁻³⁵⁹.

2.3. Depicting the phylogenetic evolution of CRC

Several studies have estimated intratumor heterogeneity by either deep sequencing or single cell based approaches allowing the inference of the evolutionary history of tumors^{236, 242, 244, 329}. In this sense, FISHtree software enables the assembly of phylogenetic trees from FISH data³⁰⁰. A previous study used multi-region CRC sampling and deep sequencing to show branched evolution for CRC and determined founder CNAs²⁴⁴. Utilizing our FISH approach, we observed branched evolution for our CRC samples and further confirmed the roles of chromosomes 7p, 13q and 20q in leading CRC progression. Interestingly, our work also revealed branched evolution in the CRC precursor lesions, thus providing novel insight into the phylogenetic transition of an AD into ADK. Moreover, we found the initial imbalance in the chromosomes driving the progression to occur early in the evolution, and to be present in the common ancestor clone and lead the AD-to-ADK transition. In fact, we observed that it was the sequential copy number increase that provided the majority of the progression. These findings are in agreement with the theory of neutral evolution proposed by Sottoriva and colleagues^{330, 360}, which also proposed the absence of a strong selection during evolution and the high prevalence of low-frequency clones, as our data suggests. Moreover, our study highlighted the importance of these low-frequency clones in two different situations: the common ancestor did not always come from the main population of the AD, and the fittest clone in the ADK was not always part of the main cell population. In this regard, other authors reported that intratumor heterogeneity could lead to the underestimation of a tumor's genetic composition and interfere with both therapeutic response and biomarker discovery^{236, 245, 330, 360, 361}.

Genome doubling events have been described to occur in several types of tumors³⁶². Indeed, a recent study found epithelial cancers to have the highest incidence of WGD events, accounting for over 50% of the cases¹⁸¹. Genome doubling events are believed to be clonal, and during lung cancer evolution have been shown to occur before tumor diversification^{242, 334}. Furthermore, another study detected a tetraploid subpopulation in the phylogenetic branch preceding metastasis in renal cancer²³⁶. Accordingly, we found WGD

events to occur during the AD-to-ADK transition and also during the progression of not only ADK but also AD. Therefore, we defined the different possible timings for WGD during CRC evolution.

3. Tetraploidy and cancer

As demonstrated by the previous findings, chromosome alterations are early contributors to the malignant transformation of CRC. In fact, CIN and aneuploidy are considered to be drivers of carcinogenesis³³². Moreover, WGD are present in many tumors and the fact that a significant percentage of cancers exhibit near-triploid karyotypes supports the notion that a tetraploid intermediate promotes the generation of highly aneuploid karyotypes within a tumor³⁶². Indeed, tetraploid cells are prone to both tolerate and propagate CIN, thereby increasing the cell's adaptability and buffering the detrimental consequences of aneuploidy^{190, 363-366}. Furthermore, it has been reported that tetraploid cells are more tumorigenic than their diploid counterparts¹⁵⁴, and their presence has been observed during early stages of cancer¹⁹²⁻¹⁹⁴. Nevertheless, although the role of aneuploidy in cancer has been described in great depth²²³, the effects of polyploidization on cellular behavior remain poorly understood. Therefore, we sought to characterize isogenic cell lines of different ploidy status in order to assess the differential features conferred by tetraploidy. Our results indicated that near-tetraploid cells displayed replication stress, which promoted genomic instability, including structural and numerical CIN, leading to intercellular heterogeneity. Furthermore, we also found polyploidization to be associated with an increased capacity for mobility, which was supported by our observation of tetraploid and highly unstable cells in invasive fronts of colon primary tumors.

3.1. The role of p53 in the generation of tetraploidy

Several studies have suggested the existence of a pathway limiting the generation of tetraploid cells, with the p53-dependent arrest being the best described¹⁹⁷⁻¹⁹⁹. Despite this, the role of p53 in the generation of tetraploid cells remains a matter of debate. While some authors have suggested that it is the drug treatment used to produce these tetraploid cells which promotes their arrest^{200, 201}, others have postulated that a non-functional p53 pathway will allow more efficient and rapid proliferation of tetraploid cells because

chromosome missegregation after tetraploidy formation triggers a p53-dependent arrest¹⁹⁷. In fact, abrogation of p53 function in tetraploid mouse mammary epithelial cells generates tumorigenic chromosomally unstable cells¹⁵⁴. However, a recent study demonstrated that p53 proficient polyploidy cells are able to progress due to the overexpression of oncogenes, such as *CCND2*, indicating that the abrogation of p53 is not essential for tetraploidy generation³⁶⁷. Accordingly, whereas we found that our p53-deficient DLD-1 cells spontaneously generated tetraploidy, we could not detect the spontaneous formation of a tetraploid subpopulation from our p53-proficient RKO cells. Also, after inducing cytokinesis failure in the RKO cells we observed an arrest as well as a low-frequency of binucleated cells formation, even though some tetraploid cells were finally able to progress. Therefore, our results indicate that p53 deficiency favors the formation of tetraploid cells, although it is not completely essential for the generation of polyploid cells.

3.2. Chromosomal instability is a common feature of near-tetraploid cells

Our data also showed that near-tetraploid cells exhibit increased levels of genomic instability and chromosomal heterogeneity, which was in agreement with previous studies^{366, 368}. Although tetraploid cells also arise in normal organs³⁶⁹, it is considered to be an unstable intermediate state that leads to aneuploid karyotypes³⁶⁸. Moreover, as stated before, it is near-tetraploid and near-triploid karyotypes that are frequently observed in human cancers³⁶². Our 4N cells also displayed high rates of missegregation defects, mainly lagging chromosomes, which help explain the high levels of chromosomal heterogeneity and indicates a degree of tolerance to CIN. Indeed, recent studies have shown that even the presence of an individual extra chromosome increases mitotic defects and overall genomic instability^{126, 127}. Furthermore, while extra centrosomes tend to cluster together, they generate a transient multipolar stage during prometaphase, thus triggering the formation of merotelic attachments^{165, 167}. In accordance with these findings, we found an increase in the number of cells with supernumerary centrosomes in our 4N cells, which partially explain the increase in lagging chromosomes. Additionally, our observation of monopolar spindles during prometaphase might indicate a delayed spindle bipolarization, which would promote the formation of lagging chromosomes, as previously described³⁷⁰.

3.3. Near-tetraploid cells experience replication stress

Despite the high levels of mitotic defects we observed in 4N cells, our gene expression data also revealed the upregulation of genes related to replication stress (RS), which also contributes to genomic instability and CIN^{127, 258, 264, 265, 280, 282}. In fact, oncogene activation during cancer progression induces RS through several mechanisms, including the decrease in nucleotides pool levels, and increases DNA damage^{268, 371}. Our data revealed that 4N clones displayed an upregulation of multiple genes implicated in the RS checkpoint, including *ATRIP*, *PCNA*, and several members of the *MCM* family. Consistent with this, we observed greater levels of the S-phase checkpoint protein pCHK1 in both sets of 4N clones. Indeed, the levels of pCHK1 were almost undetectable in the RKO 2N cells, while the newly generated RKO 4N cells displayed a significant RS response, reinforcing the idea that replication stress is a consequence rather than a cause of tetraploidy, contrary to what has been previously stated³⁷¹. Moreover, RS was also revealed by a S-phase specific delay together with increased RPA levels and overexpression of *RRM2* in our 4N clones, which may indicate an adaptive mechanism to cope with RS. Interestingly, since RS can be detrimental for cancer cells that are dependent on the activation of the S-phase checkpoint to survive under high levels of genomic instability^{255, 279, 289}, 4N cell growth was compromised upon the inhibition of ATR.

Furthermore, RS promotes the generation of DNA damage during replication, leading to the activation of the DNA damage response (DDR) and the consequent phosphorylation of H2AX and formation of 53BP1 foci^{261, 262}. Indeed, a recent study showed that cells with extra chromosomes exhibit high levels of DNA damage associated with RS¹²⁷. In this regard, we detected greater levels of γ H2AX and large 53BP1 foci in our 4N cells in comparison to their diploid counterparts. However, the levels of γ H2AX exceeded the DNA damage expected due to the double DNA content of 4N cells, reaching a 2.7-fold increase. Moreover, the fact that almost the entire 4N population was positive for γ H2AX in S-phase might have led to the significant activation of the RS response as well as the dependence on it of the 4N cells. In fact, we detected γ H2AX in a significant proportion of 4N cells in M-phase, more than the expected for their DNA content, indicating that incomplete DNA synthesis during S-phase was carried through to the subsequent G1-phase³⁷¹⁻³⁷³. Collectively, our results indicate that tetraploidy leads to RS and, subsequently, increases

DNA damage and genomic instability, which may help explain the presence of near-tetraploid karyotypes in cancer³⁶², despite its potentially deleterious effects.

3.4. Replication stress leads to both structural and numerical chromosome abnormalities

The association of RS with structural chromosome aberrations has been widely recognized^{127, 282, 374}; however, the contribution of RS to numerical chromosome abnormalities has yet to be completely understood. We observed the induction of RS upon aphidicolin treatment not only increased the levels of anaphase bridges and acentric fragments, but also the frequency of lagging chromosomes, suggesting a novel and direct link between RS and numerical CIN. Interestingly and along the same line, the activation of the DDR was recently associated with the formation of lagging chromosomes³⁷⁵.

Moreover, it has been recently described that chromosome bridges do not always break during anaphase. Instead, they might also lag behind in a position proximal the spindle equator of the mitotic cell, thus causing the missegregation of the bridged chromosomes or even the formation of a micronucleus³⁷⁶. These observations might explain the appearance of what we considered multiple missegregating chromosomes in our 4N cells and after the induction of RS with aphidicolin. Therefore, indicating that RS can also promote the formation of chromosome bridges that not only lead to structural chromosome errors but also numerical ones.

3.5. The benefit of tetraploidy on cancer cell behavior

Whereas aneuploidy has been widely correlated with proliferation impairment in physiological conditions^{221, 226, 234, 377}, recent studies have demonstrated that it may also improve cellular fitness and confer an adaptive potential under selective environments^{222, 234, 378, 379}. Additionally, although the consequences of tetraploidization have not been fully studied, some authors reported increased resistance to chemotherapeutic drugs and anchorage-independent growth in tetraploid cells, indicating that genome doubling events can provide benefits for the malignant growth³⁶⁶. We observed not only growth impairment in our 4N cells, but also enhanced migration and invasive capacities. These somewhat conflicting results highlight the diverse implications of having extra chromosomes, which

themselves carry different individual genes that can impact the cell's behavior and lead to phenotypic advantages. In fact, our gene expression data showed an upregulation of genes associated with mobility in 4N cells. Moreover, when we assessed the ploidy of invasive fronts from primary colon tumor samples, we observed higher levels of karyotypic variability along with increased percentage of 4N cells in comparison to the main tumor mass.

In addition, the presence of supernumerary centrosomes in cancer cells has also been associated with aggressiveness³⁸⁰. However, the *in vitro* migration capabilities were independent of extra centrosomes and although the population of cells with extra centrosomes was larger in the invasive fronts than in the main tumor mass, it was still relatively low (10%) and could not by itself completely explain the degree of invasiveness and the presence of karyotypic variability. Nevertheless, additional studies are needed in order to fully assess the mobility capabilities of 4N cells *in vivo*.

CONCLUSIONS

Aberrant crypt foci and LINE-1 methylation as predictive biomarkers for CRC development

1. The inability to reliably classify ACF based on their histology hinders their use as surrogate markers for CRC.
2. Endoscopic and morphologic characteristics do not indicate ACF are suitable biomarkers for CRC risk.
3. Only a few of the genetic and epigenetic CRC alterations are detected in ACF and they do not correlate with the CRC risk group.
4. ACF might merely be preneoplastic and not an intermediate endpoint for colorectal carcinogenesis.
5. ACF do not exhibit global DNA hypomethylation and our results do not support LINE-1 methylation field defect for sporadic CRC.

Copy number alterations and whole genome duplication events during CRC evolution

1. Chromosome imbalances progressively increase during colorectal carcinogenesis, leading to extensive intratumor heterogeneity.
2. The amount and patterns of chromosome alterations distinguish non-advanced adenomas from progressed adenomas.
3. The AD-to-ADK transition and CRC progression are defined by specific chromosome alterations, mainly involving the sequential gain of either 7p, 13q or 20q, or the combination of them.
4. WGD events are not only implicated in CRC progression but also in tumor initiation, thus demonstrating the crucial role of genome doubling events in CRC evolution.

5. The first gain in driver chromosomes is acquired during the early stages of CRC formation. Moreover, low-frequency clones and the absence of strong selection define the landscape of this type of cancer.
6. Increased DNA ploidy influences the level of genomic instability and the existence of heterogeneity in a population. Moreover, near-tetraploid cells are able to progress and show tolerance to aneuploidy.
7. Near-tetraploid cells systematically undergo replication stress, which increases the levels of both structural and numerical CIN.
8. Although near-tetraploid cells exhibit a growth impairment due to replication stress and genomic instability, they show greater migratory and invasive capabilities than their diploid counterparts, thus supporting their being observed in the invasive fronts of primary tumors.

BIBLIOGRAPHY

1. Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin.* 2014;64(2):104-17. Epub 2014/03/19.
2. Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. *Cancer Epidemiol Biomarkers Prev.* 2009;18(6):1688-94. Epub 2009/06/10.
3. Stock C, Pulte D, Haug U, et al. Subsite-specific colorectal cancer risk in the colorectal endoscopy era. *Gastrointest Endosc.* 2012;75(3):621-30. Epub 2012/02/22.
4. Welch HG, Robertson DJ. Colorectal Cancer on the Decline--Why Screening Can't Explain It All. *N Engl J Med.* 2016;374(17):1605-7. Epub 2016/04/28.
5. Brenner H, Bouvier AM, Foschi R, et al. Progress in colorectal cancer survival in Europe from the late 1980s to the early 21st century: the EURO CARE study. *Int J Cancer.* 2012;131(7):1649-58. Epub 2011/05/25.
6. Taylor DP, Burt RW, Williams MS, et al. Population-based family history-specific risks for colorectal cancer: a constellation approach. *Gastroenterology.* 2010;138(3):877-85. Epub 2009/11/26.
7. Jess T, Rungoe C, Peyrin-Biroulet L. Risk of colorectal cancer in patients with ulcerative colitis: a meta-analysis of population-based cohort studies. *Clin Gastroenterol Hepatol.* 2012;10(6):639-45. Epub 2012/02/01.
8. Liang PS, Chen TY, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *Int J Cancer.* 2009;124(10):2406-15. Epub 2009/01/15.
9. Fedirko V, Tramacere I, Bagnardi V, et al. Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. *Ann Oncol.* 2011;22(9):1958-72. Epub 2011/02/11.
10. Chan DS, Lau R, Aune D, et al. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS One.* 2011;6(6):e20456. Epub 2011/06/16.
11. Ma Y, Yang Y, Wang F, et al. Obesity and risk of colorectal cancer: a systematic review of prospective studies. *PLoS One.* 2013;8(1):e53916. Epub 2013/01/26.
12. Jiang Y, Ben Q, Shen H, et al. Diabetes mellitus and incidence and mortality of colorectal cancer: a systematic review and meta-analysis of cohort studies. *Eur J Epidemiol.* 2011;26(11):863-76. Epub 2011/09/23.

13. Elmunzer BJ, Hayward RA, Schoenfeld PS, et al. Effect of flexible sigmoidoscopy-based screening on incidence and mortality of colorectal cancer: a systematic review and meta-analysis of randomized controlled trials. *PLoS Med.* 2012;9(12):e1001352. Epub 2012/12/12.
14. Pinol V, Castells A, Andreu M, et al. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA.* 2005;293(16):1986-94. Epub 2005/04/28.
15. Burt RW. Colon cancer screening. *Gastroenterology.* 2000;119(3):837-53. Epub 2000/09/13.
16. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology.* 2007;50(1):113-30. Epub 2007/01/06.
17. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med.* 1988;319(9):525-32. Epub 1988/09/01.
18. Gupta AK, Pinsky P, Rall C, et al. Reliability and accuracy of the endoscopic appearance in the identification of aberrant crypt foci. *Gastrointest Endosc.* 2009;70(2):322-30. Epub 2009/06/23.
19. Bird RP. Observation and quantification of aberrant crypts in the murine colon treated with a colon carcinogen: preliminary findings. *Cancer Lett.* 1987;37(2):147-51. Epub 1987/10/30.
20. Pretlow TP, Barrow BJ, Ashton WS, et al. Aberrant crypts: putative preneoplastic foci in human colonic mucosa. *Cancer Res.* 1991;51(5):1564-7. Epub 1991/03/01.
21. Drew DA, Devers TJ, O'Brien MJ, et al. HD chromoendoscopy coupled with DNA mass spectrometry profiling identifies somatic mutations in microdissected human proximal aberrant crypt foci. *Mol Cancer Res.* 2014;12(6):823-9. Epub 2014/03/22.
22. Yokota T, Sugano K, Kondo H, et al. Detection of aberrant crypt foci by magnifying colonoscopy. *Gastrointest Endosc.* 1997;46(1):61-5. Epub 1997/07/01.
23. Hurlstone DP, Karajeh M, Sanders DS, et al. Rectal aberrant crypt foci identified using high-magnification-chromoscopic colonoscopy: biomarkers for flat and depressed neoplasia. *Am J Gastroenterol.* 2005;100(6):1283-9. Epub 2005/06/03.

24. Al-Henhena N, Khalifa SA, Ying RP, et al. Chemopreventive effects of *Strobilanthes crispus* leaf extract on azoxymethane-induced aberrant crypt foci in rat colon. *Sci Rep*. 2015;5:13312. Epub 2015/08/27.
25. Takayama T KS, Takahashi Y, et al. Aberrant crypt foci of the colon as precursors of adenoma and cancer. *N Engl J Med*. 1998;339 (18):1277-84
26. Adler DG, Gostout CJ, Sorbi D, et al. Endoscopic identification and quantification of aberrant crypt foci in the human colon. *Gastrointest Endosc*. 2002;56(5):657-62. Epub 2002/10/25.
27. Roncucci L, Modica S, Pedroni M, et al. Aberrant crypt foci in patients with colorectal cancer. *Br J Cancer*. 1998;77(12):2343-8. Epub 1998/07/02.
28. Stevens RG, Swede H, Rosenberg DW. Epidemiology of colonic aberrant crypt foci: review and analysis of existing studies. *Cancer Lett*. 2007;252(2):171-83. Epub 2006/12/22.
29. Cho NL, Redston M, Zauber AG, et al. Aberrant crypt foci in the adenoma prevention with celecoxib trial. *Cancer Prev Res (Phila)*. 2008;1(1):21-31. Epub 2009/01/14.
30. Mutch MG, Schoen RE, Fleshman JW, et al. A multicenter study of prevalence and risk factors for aberrant crypt foci. *Clin Gastroenterol Hepatol*. 2009;7(5):568-74. Epub 2009/05/07.
31. Stevens RG, Swede H, Heinen CD, et al. Aberrant crypt foci in patients with a positive family history of sporadic colorectal cancer. *Cancer Lett*. 2007;248(2):262-8. Epub 2006/09/05.
32. Rudolph RE, Dominitz JA, Lampe JW, et al. Risk factors for colorectal cancer in relation to number and size of aberrant crypt foci in humans. *Cancer Epidemiol Biomarkers Prev*. 2005;14(3):605-8. Epub 2005/03/16.
33. Bird RP, Good CK. The significance of aberrant crypt foci in understanding the pathogenesis of colon cancer. *Toxicol Lett*. 2000;112-113:395-402. Epub 2000/03/18.
34. Siu IM, Pretlow TG, Amini SB, et al. Identification of dysplasia in human colonic aberrant crypt foci. *Am J Pathol*. 1997;150(5):1805-13. Epub 1997/05/01.
35. Pretlow TP, Brasitus TA, Fulton NC, et al. K-ras mutations in putative preneoplastic lesions in human colon. *J Natl Cancer Inst*. 1993;85(24):2004-7. Epub 1993/12/15.
36. Beach R, Chan AO, Wu TT, et al. BRAF mutations in aberrant crypt foci and hyperplastic polyposis. *Am J Pathol*. 2005;166(4):1069-75. Epub 2005/03/29.

37. Takayama T OM, Hayashi T, Miyanishi K, Nobuoka A, Nakajima T, Satoh T, Takimoto R, Kato J, Sakamaki S, Niitsu Y. Analysis of K-ras, APC, and beta-catenin in aberrant crypt foci in sporadic adenoma, cancer, and familial adenomatous polyposis. *Gastroenterology*. 2001;121(3):599-611. Epub 611.
38. Smith AJ SH, Penner M, Hay K, Mitri A, Bapat BV, Gallinger S. Somatic APC and K-ras codon 12 mutations in aberrant crypt foci from human colons. *Cancer Res*. 1994;54(21):5527-30.
39. Carr NJ SL, Neiderau C, editors. Aberrant crypt foci. *Pathology and genetics: Tumors of the digestive system* 2000.
40. Pretlow TP, Pretlow TG. Mutant KRAS in aberrant crypt foci (ACF): initiation of colorectal cancer? *Biochim Biophys Acta*. 2005;1756(2):83-96. Epub 2005/10/13.
41. Suehiro Y, Hinoda Y. Genetic and epigenetic changes in aberrant crypt foci and serrated polyps. *Cancer Sci*. 2008;99(6):1071-6. Epub 2008/04/04.
42. Alrawi SJ, Carroll RE, Hill HC, et al. Genomic instability of human aberrant crypt foci measured by inter-(simple sequence repeat) PCR and array-CGH. *Mutat Res*. 2006;601(1-2):30-8. Epub 2006/06/30.
43. Morson BC. Evolution of cancer of the colon and rectum. *Cancer*. 1974;34(3):suppl:845-9. Epub 1974/09/01.
44. Schoen RE, Mutch M, Rall C, et al. The natural history of aberrant crypt foci. *Gastrointest Endosc*. 2008;67(7):1097-102. Epub 2008/01/08.
45. Spring KJ, Zhao ZZ, Karamatic R, et al. High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. *Gastroenterology*. 2006;131(5):1400-7. Epub 2006/11/15.
46. Hamilton SR, Aaltonen LA. *Pathology and Genetics of Tumours of the Digestive System*. IARC Press. 2000.
47. Dik VK, Moons LM, Siersema PD. Endoscopic innovations to increase the adenoma detection rate during colonoscopy. *World J Gastroenterol*. 2014;20(9):2200-11. Epub 2014/03/08.
48. Powell SM, Zilz N, Beazer-Barclay Y, et al. APC mutations occur early during colorectal tumorigenesis. *Nature*. 1992;359(6392):235-7. Epub 1992/09/17.
49. Shinya H, Wolff WI. Morphology, anatomic distribution and cancer potential of colonic polyps. *Ann Surg*. 1979;190(6):679-83. Epub 1979/12/01.

50. Fenoglio CM, Lane N. The anatomical precursor of colorectal carcinoma. *Cancer*. 1974;34(3):suppl:819-23. Epub 1974/09/01.
51. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. *Cancer*. 1975;36(6):2251-70. Epub 1975/12/01.
52. Burt RW, Bishop DT, Lynch HT, et al. Risk and surveillance of individuals with heritable factors for colorectal cancer. WHO Collaborating Centre for the Prevention of Colorectal Cancer. *Bull World Health Organ*. 1990;68(5):655-65. Epub 1990/01/01.
53. Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. *Eur J Cancer*. 2006;42(2):216-27. Epub 2005/12/13.
54. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med*. 1993;329(27):1977-81. Epub 1993/12/30.
55. DiSario JA, Foutch PG, Mai HD, et al. Prevalence and malignant potential of colorectal polyps in asymptomatic, average-risk men. *Am J Gastroenterol*. 1991;86(8):941-5. Epub 1991/08/01.
56. Boland CR, Hofstad B, Harpaz N. Colon Polyps: Incidence, Growth and Pathology. Wayne JD, Rex DK, Williams CB, eds. Massachusetts: Blackwell Publishing, Inc.; 2003.
57. Vogelstein B, Kinzler KW. The genetic basis of human cancer. New York McGraw-Hill Medical; 2002.
58. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology*. 2010;138(6):2088-100. Epub 2010/04/28.
59. Higuchi T, Sugihara K, Jass JR. Demographic and pathological characteristics of serrated polyps of colorectum. *Histopathology*. 2005;47(1):32-40. Epub 2005/06/29.
60. Snover DC. Update on the serrated pathway to colorectal carcinoma. *Hum Pathol*. 2011;42(1):1-10. Epub 2010/09/28.
61. Bettington M, Walker N, Clouston A, et al. The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology*. 2013;62(3):367-86. Epub 2013/01/24.
62. Cooper HS. Surgical pathology of endoscopically removed malignant polyps of the colon and rectum. *Am J Surg Pathol*. 1983;7(7):613-23. Epub 1983/10/01.

63. Coverlizza S, Risio M, Ferrari A, et al. Colorectal adenomas containing invasive carcinoma. Pathologic assessment of lymph node metastatic potential. *Cancer*. 1989;64(9):1937-47. Epub 1989/11/01.
64. Morson BC, Whiteway JE, Jones EA, et al. Histopathology and prognosis of malignant colorectal polyps treated by endoscopic polypectomy. *Gut*. 1984;25(5):437-44. Epub 1984/05/01.
65. Kitajima K, Fujimori T, Fujii S, et al. Correlations between lymph node metastasis and depth of submucosal invasion in submucosal invasive colorectal carcinoma: a Japanese collaborative study. *J Gastroenterol*. 2004;39(6):534-43. Epub 2004/07/06.
66. Beaton C, Twine CP, Williams GL, et al. Systematic review and meta-analysis of histopathological factors influencing the risk of lymph node metastasis in early colorectal cancer. *Colorectal Dis*. 2013;15(7):788-97. Epub 2013/01/22.
67. Bosch SL, Teerenstra S, de Wilt JH, et al. Predicting lymph node metastasis in pT1 colorectal cancer: a systematic review of risk factors providing rationale for therapy decisions. *Endoscopy*. 2013;45(10):827-34. Epub 2013/07/26.
68. Quirke P, Risio M, Lambert R, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Quality assurance in pathology in colorectal cancer screening and diagnosis. *Endoscopy*. 2012;44 Suppl 3:SE116-30. Epub 2012/10/04.
69. Watanabe T, Itabashi M, Shimada Y, et al. Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2010 for the treatment of colorectal cancer. *Int J Clin Oncol*. 2012;17(1):1-29. Epub 2011/10/18.
70. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61(5):759-67. Epub 1990/06/01.
71. Issa JP. Colon cancer: it's CIN or CIMP. *Clin Cancer Res*. 2008;14(19):5939-40. Epub 2008/10/03.
72. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. *Gastroenterology*. 2010;138(6):2059-72. Epub 2010/04/28.
73. Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*. 1993;363(6429):558-61. Epub 1993/06/10.

74. Wheeler JM, Bodmer WF, Mortensen NJ. DNA mismatch repair genes and colorectal cancer. *Gut*. 2000;47(1):148-53. Epub 2000/06/22.
75. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015;21(11):1350-6. Epub 2015/10/13.
76. Groden J, Thliveris A, Samowitz W, et al. Identification and characterization of the familial adenomatous polyposis coli gene. *Cell*. 1991;66(3):589-600. Epub 1991/08/09.
77. De Filippo C, Luceri C, Caderni G, et al. Mutations of the APC gene in human sporadic colorectal cancers. *Scand J Gastroenterol*. 2002;37(9):1048-53. Epub 2002/10/11.
78. Aberle H, Bauer A, Stappert J, et al. beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J*. 1997;16(13):3797-804. Epub 1997/07/01.
79. Molenaar M, van de Wetering M, Oosterwegel M, et al. XTcf-3 transcription factor mediates beta-catenin-induced axis formation in *Xenopus* embryos. *Cell*. 1996;86(3):391-9. Epub 1996/08/09.
80. He TC, Sparks AB, Rago C, et al. Identification of c-MYC as a target of the APC pathway. *Science*. 1998;281(5382):1509-12. Epub 1998/09/04.
81. Andreyev HJ, Norman AR, Cunningham D, et al. Kirsten ras mutations in patients with colorectal cancer: the multicenter "RASCAL" study. *J Natl Cancer Inst*. 1998;90(9):675-84. Epub 1998/05/20.
82. Armaghany T, Wilson JD, Chu Q, et al. Genetic alterations in colorectal cancer. *Gastrointest Cancer Res*. 2012;5(1):19-27. Epub 2012/05/11.
83. Walther A, Johnstone E, Swanton C, et al. Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer*. 2009;9(7):489-99. Epub 2009/06/19.
84. Caron de Fromentel C, Soussi T. TP53 tumor suppressor gene: a model for investigating human mutagenesis. *Genes Chromosomes Cancer*. 1992;4(1):1-15. Epub 1992/01/01.
85. Lane DP, Benchimol S. p53: oncogene or anti-oncogene? *Genes Dev*. 1990;4(1):1-8. Epub 1990/01/01.
86. Leslie A, Carey FA, Pratt NR, et al. The colorectal adenoma-carcinoma sequence. *Br J Surg*. 2002;89(7):845-60. Epub 2002/06/26.
87. Meling GI, Lothe RA, Borresen AL, et al. The TP53 tumour suppressor gene in colorectal carcinomas. I. Genetic alterations on chromosome 17. *Br J Cancer*. 1993;67(1):88-92. Epub 1993/01/01.

88. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 1997;390(6659):465-71. Epub 1997/12/11.
89. Duff EK, Clarke AR. Smad4 (DPC4)--a potent tumour suppressor? *Br J Cancer*. 1998;78(12):1615-9. Epub 1998/12/23.
90. Popat S, Houlston RS. A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis. *Eur J Cancer*. 2005;41(14):2060-70. Epub 2005/08/30.
91. Rowan A, Halford S, Gaasenbeek M, et al. Refining molecular analysis in the pathways of colorectal carcinogenesis. *Clin Gastroenterol Hepatol*. 2005;3(11):1115-23. Epub 2005/11/08.
92. Alhopuro P, Alazzouzi H, Sammalkorpi H, et al. SMAD4 levels and response to 5-fluorouracil in colorectal cancer. *Clin Cancer Res*. 2005;11(17):6311-6. Epub 2005/09/08.
93. Holliday R. The inheritance of epigenetic defects. *Science*. 1987;238(4824):163-70. Epub 1987/10/09.
94. Handel AE, Ebers GC, Ramagopalan SV. Epigenetics: molecular mechanisms and implications for disease. *Trends Mol Med*. 2010;16(1):7-16. Epub 2009/12/22.
95. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A*. 2006;103(5):1412-7. Epub 2006/01/25.
96. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet*. 2009;41(2):178-86. Epub 2009/01/20.
97. Doi A, Park IH, Wen B, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet*. 2009;41(12):1350-3. Epub 2009/11/03.
98. Bestor TH. The DNA methyltransferases of mammals. *Hum Mol Genet*. 2000;9(16):2395-402. Epub 2000/09/27.
99. Ibrahim AE, Arends MJ, Silva AL, et al. Sequential DNA methylation changes are associated with DNMT3B overexpression in colorectal neoplastic progression. *Gut*. 2011;60(4):499-508. Epub 2010/11/12.
100. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet*. 2007;8(4):286-98. Epub 2007/03/07.

101. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*. 1983;301(5895):89-92. Epub 1983/01/06.
102. Feinberg AP, Tycko B. The history of cancer epigenetics. *Nat Rev Cancer*. 2004;4(2):143-53. Epub 2004/01/21.
103. Rodriguez J, Frigola J, Vendrell E, et al. Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers. *Cancer Res*. 2006;66(17):8462-9468. Epub 2006/09/05.
104. Eden A, Gaudet F, Waghmare A, et al. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*. 2003;300(5618):455. Epub 2003/04/19.
105. Bestor TH. Transposons reanimated in mice. *Cell*. 2005;122(3):322-5. Epub 2005/08/13.
106. Fraga MF, Herranz M, Espada J, et al. A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer Res*. 2004;64(16):5527-34. Epub 2004/08/18.
107. Esteller M, Fraga MF, Guo M, et al. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum Mol Genet*. 2001;10(26):3001-7. Epub 2001/12/26.
108. Chan AO BR, Houlihan PS, Issa JP, Hamilton SR, Rashid A. CpG island methylation in aberrant crypt foci of the colorectum. *Am J Pathol*. 2002;160(5):1823-30.
109. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet*. 2006;38(7):787-93. Epub 2006/06/29.
110. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. Epub 2001/03/10.
111. Kazazian HH, Jr. Mobile elements: drivers of genome evolution. *Science*. 2004;303(5664):1626-32. Epub 2004/03/16.
112. Howard G, Eiges R, Gaudet F, et al. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*. 2008;27(3):404-8. Epub 2007/07/11.
113. Sunami E, de Maat M, Vu A, et al. LINE-1 hypomethylation during primary colon cancer progression. *PLoS One*. 2011;6(4):e18884. Epub 2011/05/03.

114. Chalitchagorn K, Shuangshoti S, Hourpai N, et al. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. *Oncogene*. 2004;23(54):8841-6. Epub 2004/10/14.
115. Ogino S, Nosho K, Kirkner GJ, et al. A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. *J Natl Cancer Inst*. 2008;100(23):1734-8. Epub 2008/11/27.
116. Kamiyama H, Suzuki K, Maeda T, et al. DNA demethylation in normal colon tissue predicts predisposition to multiple cancers. *Oncogene*. 2012;31(48):5029-37. Epub 2012/02/09.
117. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. *Nat Rev Cancer*. 2004;4(4):296-307. Epub 2004/04/02.
118. Lee KH, Lee JS, Nam JH, et al. Promoter methylation status of hMLH1, hMSH2, and MGMT genes in colorectal cancer associated with adenoma-carcinoma sequence. *Langenbecks Arch Surg*. 2011;396(7):1017-26. Epub 2011/06/28.
119. Hibi K GT, Mizukami H, Kitamura Y, Sakata M, Saito M, Ishibashi K, Kigawa G, Nemoto H, Sanada Y. MGMT gene is aberrantly methylated from the early stages of colorectal cancers. *Hepatogastroenterology*. 2009;56(96):1642-4.
120. Shen L, Kondo Y, Rosner GL, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. *J Natl Cancer Inst*. 2005;97(18):1330-8. Epub 2005/09/22.
121. von Hansermann D. Ueber asymmetrische zelltheilung in epitelkrebsen und deren biologische bedeutung. *Virchows Arch A Pathol Anat*. 1890;119:299-326.
122. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. *Nature*. 1997;386(6625):623-7. Epub 1997/04/10.
123. Haruki N, Saito H, Harano T, et al. Molecular analysis of the mitotic checkpoint genes BUB1, BUBR1 and BUB3 in human lung cancers. *Cancer Lett*. 2001;162(2):201-5. Epub 2001/01/09.
124. Yoon DS, Wersto RP, Zhou W, et al. Variable levels of chromosomal instability and mitotic spindle checkpoint defects in breast cancer. *Am J Pathol*. 2002;161(2):391-7. Epub 2002/08/07.
125. Sheltzer JM, Blank HM, Pfau SJ, et al. Aneuploidy drives genomic instability in yeast. *Science*. 2011;333(6045):1026-30. Epub 2011/08/20.

126. Nicholson JM, Macedo JC, Mattingly AJ, et al. Chromosome mis-segregation and cytokinesis failure in trisomic human cells. *Elife*. 2015;4. Epub 2015/05/06.
127. Passerini V, Ozeri-Galai E, de Pagter MS, et al. The presence of extra chromosomes leads to genomic instability. *Nat Commun*. 2016;7:10754. Epub 2016/02/16.
128. Thompson SL, Compton DA. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J Cell Biol*. 2010;188(3):369-81. Epub 2010/02/04.
129. Potapova TA, Zhu J, Li R. Aneuploidy and chromosomal instability: a vicious cycle driving cellular evolution and cancer genome chaos. *Cancer Metastasis Rev*. 2013;32(3-4):377-89. Epub 2013/05/28.
130. Musacchio A, Salmon ED. The spindle-assembly checkpoint in space and time. *Nat Rev Mol Cell Biol*. 2007;8(5):379-93. Epub 2007/04/12.
131. Hardwick KG, Murray AW. Mad1p, a phosphoprotein component of the spindle assembly checkpoint in budding yeast. *J Cell Biol*. 1995;131(3):709-20. Epub 1995/11/01.
132. Hoyt MA, Totis L, Roberts BT. *S. cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. *Cell*. 1991;66(3):507-17. Epub 1991/08/09.
133. Li R, Murray AW. Feedback control of mitosis in budding yeast. *Cell*. 1991;66(3):519-31. Epub 1991/08/09.
134. Roberts BT, Farr KA, Hoyt MA. The *Saccharomyces cerevisiae* checkpoint gene BUB1 encodes a novel protein kinase. *Mol Cell Biol*. 1994;14(12):8282-91. Epub 1994/12/01.
135. Weiss E, Winey M. The *Saccharomyces cerevisiae* spindle pole body duplication gene MPS1 is part of a mitotic checkpoint. *J Cell Biol*. 1996;132(1-2):111-23. Epub 1996/01/01.
136. Howell BJ, Moree B, Farrar EM, et al. Spindle checkpoint protein dynamics at kinetochores in living cells. *Curr Biol*. 2004;14(11):953-64. Epub 2004/06/09.
137. Kim SH, Lin DP, Matsumoto S, et al. Fission yeast Slp1: an effector of the Mad2-dependent spindle checkpoint. *Science*. 1998;279(5353):1045-7. Epub 1998/03/07.
138. Sudakin V, Chan GK, Yen TJ. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *J Cell Biol*. 2001;154(5):925-36. Epub 2001/09/06.
139. Cohen-Fix O, Peters JM, Kirschner MW, et al. Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes Dev*. 1996;10(24):3081-93. Epub 1996/12/15.

140. Li M, Fang X, Wei Z, et al. Loss of spindle assembly checkpoint-mediated inhibition of Cdc20 promotes tumorigenesis in mice. *J Cell Biol.* 2009;185(6):983-94. Epub 2009/06/17.
141. Michel LS, Liberal V, Chatterjee A, et al. MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells. *Nature.* 2001;409(6818):355-9. Epub 2001/02/24.
142. Dobles M, Liberal V, Scott ML, et al. Chromosome missegregation and apoptosis in mice lacking the mitotic checkpoint protein Mad2. *Cell.* 2000;101(6):635-45. Epub 2000/07/13.
143. Kops GJ, Foltz DR, Cleveland DW. Lethality to human cancer cells through massive chromosome loss by inhibition of the mitotic checkpoint. *Proc Natl Acad Sci U S A.* 2004;101(23):8699-704. Epub 2004/05/26.
144. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. *Science.* 2006;314(5797):268-74. Epub 2006/09/09.
145. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science.* 2007;318(5853):1108-13. Epub 2007/10/13.
146. Kato T, Daigo Y, Aragaki M, et al. Overexpression of MAD2 predicts clinical outcome in primary lung cancer patients. *Lung Cancer.* 2011;74(1):124-31. Epub 2011/03/08.
147. Tanaka K, Nishioka J, Kato K, et al. Mitotic checkpoint protein hsMAD2 as a marker predicting liver metastasis of human gastric cancers. *Jpn J Cancer Res.* 2001;92(9):952-8. Epub 2001/09/27.
148. Nasmyth K. Cohesin: a catenase with separate entry and exit gates? *Nat Cell Biol.* 2011;13(10):1170-7. Epub 2011/10/05.
149. Barber TD, McManus K, Yuen KW, et al. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci U S A.* 2008;105(9):3443-8. Epub 2008/02/27.
150. Solomon DA, Kim T, Diaz-Martinez LA, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science.* 2011;333(6045):1039-43. Epub 2011/08/20.
151. Iwaizumi M, Shinmura K, Mori H, et al. Human Sgo1 downregulation leads to chromosomal instability in colorectal cancer. *Gut.* 2009;58(2):249-60. Epub 2008/07/19.
152. Wirth KG, Wutz G, Kudo NR, et al. Separase: a universal trigger for sister chromatid disjunction but not chromosome cycle progression. *J Cell Biol.* 2006;172(6):847-60. Epub 2006/03/15.

153. Greenman C, Stephens P, Smith R, et al. Patterns of somatic mutation in human cancer genomes. *Nature*. 2007;446(7132):153-8. Epub 2007/03/09.
154. Fujiwara T, Bandi M, Nitta M, et al. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature*. 2005;437(7061):1043-7. Epub 2005/10/14.
155. Dai W, Wang Q, Liu T, et al. Slippage of mitotic arrest and enhanced tumor development in mice with BubR1 haploinsufficiency. *Cancer Res*. 2004;64(2):440-5. Epub 2004/01/28.
156. Basto R, Brunk K, Vinadogrova T, et al. Centrosome amplification can initiate tumorigenesis in flies. *Cell*. 2008;133(6):1032-42. Epub 2008/06/17.
157. Peloponese JM, Jr., Haller K, Miyazato A, et al. Abnormal centrosome amplification in cells through the targeting of Ran-binding protein-1 by the human T cell leukemia virus type-1 Tax oncoprotein. *Proc Natl Acad Sci U S A*. 2005;102(52):18974-9. Epub 2005/12/21.
158. Wang X, Yang Y, Duan Q, et al. sSgo1, a major splice variant of Sgo1, functions in centriole cohesion where it is regulated by Plk1. *Dev Cell*. 2008;14(3):331-41. Epub 2008/03/12.
159. Duelli D, Lazebnik Y. Cell fusion: a hidden enemy? *Cancer Cell*. 2003;3(5):445-8. Epub 2003/06/05.
160. Boveri T. Zur Frage der Entstehung Maligner Tumoren. 1914.
161. Ghadimi BM, Sackett DL, Difilippantonio MJ, et al. Centrosome amplification and instability occurs exclusively in aneuploid, but not in diploid colorectal cancer cell lines, and correlates with numerical chromosomal aberrations. *Genes Chromosomes Cancer*. 2000;27(2):183-90. Epub 1999/12/29.
162. Nigg EA. Origins and consequences of centrosome aberrations in human cancers. *Int J Cancer*. 2006;119(12):2717-23. Epub 2006/10/04.
163. Lingle WL, Barrett SL, Negron VC, et al. Centrosome amplification drives chromosomal instability in breast tumor development. *Proc Natl Acad Sci U S A*. 2002;99(4):1978-83. Epub 2002/02/07.
164. Neben K, Giesecke C, Schweizer S, et al. Centrosome aberrations in acute myeloid leukemia are correlated with cytogenetic risk profile. *Blood*. 2003;101(1):289-91. Epub 2002/10/24.

165. Ganem NJ, Godinho SA, Pellman D. A mechanism linking extra centrosomes to chromosomal instability. *Nature*. 2009;460(7252):278-82. Epub 2009/06/10.
166. Quintyne NJ, Reing JE, Hoffelder DR, et al. Spindle multipolarity is prevented by centrosomal clustering. *Science*. 2005;307(5706):127-9. Epub 2005/01/08.
167. Silkworth WT, Nardi IK, Scholl LM, et al. Multipolar spindle pole coalescence is a major source of kinetochore mis-attachment and chromosome mis-segregation in cancer cells. *PLoS One*. 2009;4(8):e6564. Epub 2009/08/12.
168. Rieder CL, Schultz A, Cole R, et al. Anaphase onset in vertebrate somatic cells is controlled by a checkpoint that monitors sister kinetochore attachment to the spindle. *J Cell Biol*. 1994;127(5):1301-10. Epub 1994/12/01.
169. Rieder CL, Cole RW, Khodjakov A, et al. The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *J Cell Biol*. 1995;130(4):941-8. Epub 1995/08/01.
170. Hauf S, Cole RW, LaTerra S, et al. The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. *J Cell Biol*. 2003;161(2):281-94. Epub 2003/04/23.
171. Cimini D, Howell B, Maddox P, et al. Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. *J Cell Biol*. 2001;153(3):517-27. Epub 2001/05/02.
172. Cimini D, Cameron LA, Salmon ED. Anaphase spindle mechanics prevent mis-segregation of merotelically oriented chromosomes. *Curr Biol*. 2004;14(23):2149-55. Epub 2004/12/14.
173. Thompson SL, Compton DA. Chromosome missegregation in human cells arises through specific types of kinetochore-microtubule attachment errors. *Proc Natl Acad Sci U S A*. 2011;108(44):17974-8. Epub 2011/10/15.
174. Thompson SL, Compton DA. Examining the link between chromosomal instability and aneuploidy in human cells. *J Cell Biol*. 2008;180(4):665-72. Epub 2008/02/20.
175. Bakhoun SF, Thompson SL, Manning AL, et al. Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nat Cell Biol*. 2009;11(1):27-35. Epub 2008/12/09.
176. Bakhoun SF, Genovese G, Compton DA. Deviant kinetochore microtubule dynamics underlie chromosomal instability. *Curr Biol*. 2009;19(22):1937-42. Epub 2009/11/03.

177. Papp B, Pal C, Hurst LD. Dosage sensitivity and the evolution of gene families in yeast. *Nature*. 2003;424(6945):194-7. Epub 2003/07/11.
178. Torres EM, Williams BR, Amon A. Aneuploidy: cells losing their balance. *Genetics*. 2008;179(2):737-46. Epub 2008/06/19.
179. Biesterfeld S, Gerres K, Fischer-Wein G, et al. Polyploidy in non-neoplastic tissues. *J Clin Pathol*. 1994;47(1):38-42. Epub 1994/01/01.
180. Edgar BA, Orr-Weaver TL. Endoreplication cell cycles: more for less. *Cell*. 2001;105(3):297-306. Epub 2001/05/12.
181. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol*. 2012;30(5):413-21. Epub 2012/05/01.
182. Brito DA, Rieder CL. Mitotic checkpoint slippage in humans occurs via cyclin B destruction in the presence of an active checkpoint. *Curr Biol*. 2006;16(12):1194-200. Epub 2006/06/20.
183. Duncan AW, Hanlon Newell AE, Bi W, et al. Aneuploidy as a mechanism for stress-induced liver adaptation. *J Clin Invest*. 2012;122(9):3307-15. Epub 2012/08/07.
184. Ogle BM, Cascalho M, Platt JL. Biological implications of cell fusion. *Nat Rev Mol Cell Biol*. 2005;6(7):567-75. Epub 2005/06/16.
185. Zimmet J, Ravid K. Polyploidy: occurrence in nature, mechanisms, and significance for the megakaryocyte-platelet system. *Exp Hematol*. 2000;28(1):3-16. Epub 2000/02/05.
186. Duelli DM, Hearn S, Myers MP, et al. A primate virus generates transformed human cells by fusion. *J Cell Biol*. 2005;171(3):493-503. Epub 2005/11/09.
187. Duelli DM, Padilla-Nash HM, Berman D, et al. A virus causes cancer by inducing massive chromosomal instability through cell fusion. *Curr Biol*. 2007;17(5):431-7. Epub 2007/02/27.
188. Eggert US, Mitchison TJ, Field CM. Animal cytokinesis: from parts list to mechanisms. *Annu Rev Biochem*. 2006;75:543-66. Epub 2006/06/08.
189. Storchova Z, Breneman A, Cande J, et al. Genome-wide genetic analysis of polyploidy in yeast. *Nature*. 2006;443(7111):541-7. Epub 2006/10/07.
190. Ganem NJ, Storchova Z, Pellman D. Tetraploidy, aneuploidy and cancer. *Curr Opin Genet Dev*. 2007;17(2):157-62. Epub 2007/02/28.

191. Lv L, Zhang T, Yi Q, et al. Tetraploid cells from cytokinesis failure induce aneuploidy and spontaneous transformation of mouse ovarian surface epithelial cells. *Cell Cycle*. 2012;11(15):2864-75. Epub 2012/07/18.
192. Galipeau PC, Cowan DS, Sanchez CA, et al. 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett's esophagus. *Proc Natl Acad Sci U S A*. 1996;93(14):7081-4. Epub 1996/07/09.
193. Levine DS, Rabinovitch PS, Haggitt RC, et al. Distribution of aneuploid cell populations in ulcerative colitis with dysplasia or cancer. *Gastroenterology*. 1991;101(5):1198-210. Epub 1991/11/01.
194. Olaharski AJ, Sotelo R, Solorza-Luna G, et al. Tetraploidy and chromosomal instability are early events during cervical carcinogenesis. *Carcinogenesis*. 2006;27(2):337-43. Epub 2005/08/27.
195. Andreassen PR, Lohez OD, Lacroix FB, et al. Tetraploid state induces p53-dependent arrest of nontransformed mammalian cells in G1. *Mol Biol Cell*. 2001;12(5):1315-28. Epub 2001/05/22.
196. Margolis RL, Lohez OD, Andreassen PR. G1 tetraploidy checkpoint and the suppression of tumorigenesis. *J Cell Biochem*. 2003;88(4):673-83. Epub 2003/02/11.
197. Kuffer C, Kuznetsova AY, Storchova Z. Abnormal mitosis triggers p53-dependent cell cycle arrest in human tetraploid cells. *Chromosoma*. 2013;122(4):305-18. Epub 2013/04/30.
198. Ganem NJ, Cornils H, Chiu SY, et al. Cytokinesis failure triggers hippo tumor suppressor pathway activation. *Cell*. 2014;158(4):833-48. Epub 2014/08/16.
199. Castedo M, Coquelle A, Vivet S, et al. Apoptosis regulation in tetraploid cancer cells. *EMBO J*. 2006;25(11):2584-95. Epub 2006/05/06.
200. Uetake Y, Sluder G. Cell cycle progression after cleavage failure: mammalian somatic cells do not possess a "tetraploidy checkpoint". *J Cell Biol*. 2004;165(5):609-15. Epub 2004/06/09.
201. Wong C, Stearns T. Mammalian cells lack checkpoints for tetraploidy, aberrant centrosome number, and cytokinesis failure. *BMC Cell Biol*. 2005;6(1):6. Epub 2005/02/17.
202. Gisselsson D, Jonson T, Petersen A, et al. Telomere dysfunction triggers extensive DNA fragmentation and evolution of complex chromosome abnormalities in human malignant tumors. *Proc Natl Acad Sci U S A*. 2001;98(22):12683-8. Epub 2001/10/25.

203. Gisselsson D, Pettersson L, Hoglund M, et al. Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. *Proc Natl Acad Sci U S A*. 2000;97(10):5357-62. Epub 2000/05/11.
204. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature*. 1998;396(6712):643-9. Epub 1999/01/01.
205. Mullins JM, Biesele JJ. Terminal phase of cytokinesis in D-98s cells. *J Cell Biol*. 1977;73(3):672-84. Epub 1977/06/01.
206. Pampalona J, Frias C, Genesca A, et al. Progressive telomere dysfunction causes cytokinesis failure and leads to the accumulation of polyploid cells. *PLoS Genet*. 2012;8(4):e1002679. Epub 2012/05/10.
207. Steigemann P, Wurzenberger C, Schmitz MH, et al. Aurora B-mediated abscission checkpoint protects against tetraploidization. *Cell*. 2009;136(3):473-84. Epub 2009/02/11.
208. Cimini D, Antoccia A, Tanzarella C, et al. Topoisomerase II inhibition in mitosis produces numerical and structural chromosomal aberrations in human fibroblasts. *Cytogenet Cell Genet*. 1997;76(1-2):61-7. Epub 1997/01/01.
209. Clarke DJ, Johnson RT, Downes CS. Topoisomerase II inhibition prevents anaphase chromatid segregation in mammalian cells independently of the generation of DNA strand breaks. *J Cell Sci*. 1993;105 (Pt 2):563-9. Epub 1993/06/01.
210. Downes CS, Mullinger AM, Johnson RT. Inhibitors of DNA topoisomerase II prevent chromatid separation in mammalian cells but do not prevent exit from mitosis. *Proc Natl Acad Sci U S A*. 1991;88(20):8895-9. Epub 1991/10/15.
211. Hauf S, Waizenegger IC, Peters JM. Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science*. 2001;293(5533):1320-3. Epub 2001/08/18.
212. Chestukhin A, Pfeffer C, Milligan S, et al. Processing, localization, and requirement of human separase for normal anaphase progression. *Proc Natl Acad Sci U S A*. 2003;100(8):4574-9. Epub 2003/04/04.
213. Janssen A, van der Burg M, Szuhai K, et al. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science*. 2011;333(6051):1895-8. Epub 2011/10/01.

214. Terradas M, Martin M, Tusell L, et al. DNA lesions sequestered in micronuclei induce a local defective-damage response. *DNA Repair (Amst)*. 2009;8(10):1225-34. Epub 2009/08/18.
215. Xu B, Sun Z, Liu Z, et al. Replication stress induces micronuclei comprising of aggregated DNA double-strand breaks. *PLoS One*. 2011;6(4):e18618. Epub 2011/04/29.
216. Zhang CZ, Spektor A, Cornils H, et al. Chromothripsis from DNA damage in micronuclei. *Nature*. 2015;522(7555):179-84. Epub 2015/05/29.
217. Li M, Fang X, Baker DJ, et al. The ATM-p53 pathway suppresses aneuploidy-induced tumorigenesis. *Proc Natl Acad Sci U S A*. 2010;107(32):14188-93. Epub 2010/07/29.
218. Boveri T. Über mehrpolige mitosen als mittel zur analyse des zellkerns. 1902.
219. Gao C, Furge K, Koeman J, et al. Chromosome instability, chromosome transcriptome, and clonal evolution of tumor cell populations. *Proc Natl Acad Sci U S A*. 2007;104(21):8995-9000. Epub 2007/05/23.
220. Upender MB, Habermann JK, McShane LM, et al. Chromosome transfer induced aneuploidy results in complex dysregulation of the cellular transcriptome in immortalized and cancer cells. *Cancer Res*. 2004;64(19):6941-9. Epub 2004/10/07.
221. Sheltzer JM, Ko JH, Replogle JM, et al. Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell*. 2017;31(2):240-55. Epub 2017/01/17.
222. Pavelka N, Rancati G, Zhu J, et al. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*. 2010;468(7321):321-5. Epub 2010/10/22.
223. Ried T, Hu Y, Difilippantonio MJ, et al. The consequences of chromosomal aneuploidy on the transcriptome of cancer cells. *Biochim Biophys Acta*. 2012;1819(7):784-93. Epub 2012/03/20.
224. Gemoll T, Habermann JK, Becker S, et al. Chromosomal aneuploidy affects the global proteome equilibrium of colorectal cancer cells. *Anal Cell Pathol (Amst)*. 2013;36(5-6):149-61. Epub 2014/01/28.
225. Sheltzer JM, Torres EM, Dunham MJ, et al. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci U S A*. 2012;109(31):12644-9. Epub 2012/07/18.
226. Torres EM, Sokolsky T, Tucker CM, et al. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science*. 2007;317(5840):916-24. Epub 2007/08/19.

227. Durrbaum M, Kuznetsova AY, Passerini V, et al. Unique features of the transcriptional response to model aneuploidy in human cells. *BMC Genomics*. 2014;15:139. Epub 2014/02/20.
228. Stingele S, Stoehr G, Peplowska K, et al. Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. *Mol Syst Biol*. 2012;8:608. Epub 2012/09/13.
229. Donnelly N, Storchova Z. Dynamic karyotype, dynamic proteome: buffering the effects of aneuploidy. *Biochim Biophys Acta*. 2014;1843(2):473-81. Epub 2013/12/04.
230. Oromendia AB, Amon A. Aneuploidy: implications for protein homeostasis and disease. *Dis Model Mech*. 2014;7(1):15-20. Epub 2014/01/08.
231. Donnelly N, Passerini V, Durrbaum M, et al. HSF1 deficiency and impaired HSP90-dependent protein folding are hallmarks of aneuploid human cells. *EMBO J*. 2014;33(20):2374-87. Epub 2014/09/11.
232. Selmecki A, Forche A, Berman J. Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science*. 2006;313(5785):367-70. Epub 2006/07/22.
233. Rancati G, Pavelka N, Fleharty B, et al. Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell*. 2008;135(5):879-93. Epub 2008/12/02.
234. Rutledge SD, Douglas TA, Nicholson JM, et al. Selective advantage of trisomic human cells cultured in non-standard conditions. *Sci Rep*. 2016;6:22828. Epub 2016/03/10.
235. Nowell PC. The clonal evolution of tumor cell populations. *Science*. 1976;194(4260):23-8. Epub 1976/10/01.
236. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012;366(10):883-92. Epub 2012/03/09.
237. Harris JF, Chambers AF, Hill RP, et al. Metastatic variants are generated spontaneously at a high rate in mouse KHT tumor. *Proc Natl Acad Sci U S A*. 1982;79(18):5547-51. Epub 1982/09/01.
238. Geido E, Sciutto A, Rubagotti A, et al. Combined DNA flow cytometry and sorting with k-ras2 mutation spectrum analysis and the prognosis of human sporadic colorectal cancer. *Cytometry*. 2002;50(4):216-24. Epub 2002/09/05.

239. Mekenkamp LJ, Tol J, Dijkstra JR, et al. Beyond KRAS mutation status: influence of KRAS copy number status and microRNAs on clinical outcome to cetuximab in metastatic colorectal cancer patients. *BMC Cancer*. 2012;12:292. Epub 2012/07/19.
240. Tellez-Gabriel M, Ory B, Lamoureux F, et al. Tumour Heterogeneity: The Key Advantages of Single-Cell Analysis. *Int J Mol Sci*. 2016;17(12). Epub 2016/12/22.
241. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature*. 2011;469(7330):356-61. Epub 2010/12/17.
242. de Bruin EC, McGranahan N, Mitter R, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science*. 2014;346(6206):251-6. Epub 2014/10/11.
243. Nik-Zainal S, Van Loo P, Wedge DC, et al. The life history of 21 breast cancers. *Cell*. 2012;149(5):994-1007. Epub 2012/05/23.
244. Uchi R, Takahashi Y, Niida A, et al. Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. *PLoS Genet*. 2016;12(2):e1005778. Epub 2016/02/20.
245. Yap TA, Gerlinger M, Futreal PA, et al. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med*. 2012;4(127):127ps10. Epub 2012/03/31.
246. Heselmeyer-Haddad K, Berroa Garcia LY, Bradley A, et al. Single-cell genetic analysis of ductal carcinoma in situ and invasive breast cancer reveals enormous tumor heterogeneity yet conserved genomic imbalances and gain of MYC during progression. *Am J Pathol*. 2012;181(5):1807-22. Epub 2012/10/16.
247. Heselmeyer-Haddad KM, Berroa Garcia LY, Bradley A, et al. Single-cell genetic analysis reveals insights into clonal development of prostate cancers and indicates loss of PTEN as a marker of poor prognosis. *Am J Pathol*. 2014;184(10):2671-86. Epub 2014/08/19.
248. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res*. 2012;72(19):4875-82. Epub 2012/09/25.
249. Novelli MR, Williamson JA, Tomlinson IP, et al. Polyclonal origin of colonic adenomas in an XO/XY patient with FAP. *Science*. 1996;272(5265):1187-90. Epub 1996/05/24.
250. Thirlwell C, Will OC, Domingo E, et al. Clonality assessment and clonal ordering of individual neoplastic crypts shows polyclonality of colorectal adenomas. *Gastroenterology*. 2010;138(4):1441-54, 54 e1-7. Epub 2010/01/28.

251. Maffei M, Mongera S, Terpstra L, et al. Chromosome 20 aberrations at the diploid-aneuploid transition in sporadic colorectal cancer. *Cytogenet Genome Res.* 2014;144(1):9-14. Epub 2014/10/18.
252. Carvalho B, Postma C, Mongera S, et al. Multiple putative oncogenes at the chromosome 20q amplicon contribute to colorectal adenoma to carcinoma progression. *Gut.* 2009;58(1):79-89. Epub 2008/10/03.
253. Boos D, Frigola J, Diffley JF. Activation of the replicative DNA helicase: breaking up is hard to do. *Curr Opin Cell Biol.* 2012;24(3):423-30. Epub 2012/03/20.
254. Blow JJ, Ge XQ, Jackson DA. How dormant origins promote complete genome replication. *Trends Biochem Sci.* 2011;36(8):405-14. Epub 2011/06/07.
255. Lopez-Contreras AJ, Fernandez-Capetillo O. The ATR barrier to replication-born DNA damage. *DNA Repair (Amst).* 2010;9(12):1249-55. Epub 2010/11/03.
256. Zeman MK, Cimprich KA. Causes and consequences of replication stress. *Nat Cell Biol.* 2014;16(1):2-9. Epub 2013/12/25.
257. Couch FB, Bansbach CE, Driscoll R, et al. ATR phosphorylates SMARCAL1 to prevent replication fork collapse. *Genes Dev.* 2013;27(14):1610-23. Epub 2013/07/23.
258. Gaillard H, Garcia-Muse T, Aguilera A. Replication stress and cancer. *Nat Rev Cancer.* 2015;15(5):276-89. Epub 2015/04/25.
259. Eykelenboom JK, Harte EC, Canavan L, et al. ATR activates the S-M checkpoint during unperturbed growth to ensure sufficient replication prior to mitotic onset. *Cell Rep.* 2013;5(4):1095-107. Epub 2013/11/26.
260. Jazayeri A, Falck J, Lukas C, et al. ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol.* 2006;8(1):37-45. Epub 2005/12/06.
261. Lukas C, Savic V, Bekker-Jensen S, et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat Cell Biol.* 2011;13(3):243-53. Epub 2011/02/15.
262. Szilard RK, Jacques PE, Laramée L, et al. Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. *Nat Struct Mol Biol.* 2010;17(3):299-305. Epub 2010/02/09.
263. Barlow JH, Faryabi RB, Callen E, et al. Identification of early replicating fragile sites that contribute to genome instability. *Cell.* 2013;152(3):620-32. Epub 2013/01/29.

264. Burrow AA, Williams LE, Pierce LC, et al. Over half of breakpoints in gene pairs involved in cancer-specific recurrent translocations are mapped to human chromosomal fragile sites. *BMC Genomics*. 2009;10:59. Epub 2009/02/03.
265. Le Tallec B, Millot GA, Blin ME, et al. Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. *Cell Rep*. 2013;4(3):420-8. Epub 2013/08/06.
266. Macheret M, Halazonetis TD. DNA replication stress as a hallmark of cancer. *Annu Rev Pathol*. 2015;10:425-48. Epub 2015/01/27.
267. Dereli-Oz A, Versini G, Halazonetis TD. Studies of genomic copy number changes in human cancers reveal signatures of DNA replication stress. *Mol Oncol*. 2011;5(4):308-14. Epub 2011/06/07.
268. Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. *Science*. 2008;319(5868):1352-5. Epub 2008/03/08.
269. Ekholm-Reed S, Mendez J, Tedesco D, et al. Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J Cell Biol*. 2004;165(6):789-800. Epub 2004/06/16.
270. Pruitt SC, Bailey KJ, Freeland A. Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. *Stem Cells*. 2007;25(12):3121-32. Epub 2007/08/25.
271. Shima N, Alcaraz A, Liachko I, et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet*. 2007;39(1):93-8. Epub 2006/12/05.
272. Bester AC, Roniger M, Oren YS, et al. Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell*. 2011;145(3):435-46. Epub 2011/05/03.
273. Dominguez-Sola D, Ying CY, Grandori C, et al. Non-transcriptional control of DNA replication by c-Myc. *Nature*. 2007;448(7152):445-51. Epub 2007/06/29.
274. Jones RM, Mortusewicz O, Afzal I, et al. Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress. *Oncogene*. 2013;32(32):3744-53. Epub 2012/09/05.

275. Aird KM, Zhang G, Li H, et al. Suppression of nucleotide metabolism underlies the establishment and maintenance of oncogene-induced senescence. *Cell Rep.* 2013;3(4):1252-65. Epub 2013/04/09.
276. Aye Y, Li M, Long MJ, et al. Ribonucleotide reductase and cancer: biological mechanisms and targeted therapies. *Oncogene.* 2015;34(16):2011-21. Epub 2014/06/10.
277. Aguilera A, Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell.* 2012;46(2):115-24. Epub 2012/05/01.
278. Ginno PA, Lott PL, Christensen HC, et al. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell.* 2012;45(6):814-25. Epub 2012/03/06.
279. Lecona E, Fernandez-Capetillo O. Replication stress and cancer: it takes two to tango. *Exp Cell Res.* 2014;329(1):26-34. Epub 2014/09/27.
280. Chan KL, Palmai-Pallag T, Ying S, et al. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol.* 2009;11(6):753-60. Epub 2009/05/26.
281. Naim V, Rosselli F. The FANC pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nat Cell Biol.* 2009;11(6):761-8. Epub 2009/05/26.
282. Burrell RA, McGranahan N, Bartek J, et al. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature.* 2013;501(7467):338-45. Epub 2013/09/21.
283. Wilhelm T, Magdalou I, Barascu A, et al. Spontaneous slow replication fork progression elicits mitosis alterations in homologous recombination-deficient mammalian cells. *Proc Natl Acad Sci U S A.* 2014;111(2):763-8. Epub 2013/12/19.
284. Malkova A, Ira G. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev.* 2013;23(3):271-9. Epub 2013/06/25.
285. Hoglund A, Nilsson LM, Muralidharan SV, et al. Therapeutic implications for the induced levels of Chk1 in Myc-expressing cancer cells. *Clin Cancer Res.* 2011;17(22):7067-79. Epub 2011/09/22.
286. Verlinden L, Vanden Bempt I, Eelen G, et al. The E2F-regulated gene Chk1 is highly expressed in triple-negative estrogen receptor /progesterone receptor /HER-2 breast carcinomas. *Cancer Res.* 2007;67(14):6574-81. Epub 2007/07/20.

287. Charrier JD, Durrant SJ, Golec JM, et al. Discovery of potent and selective inhibitors of ataxia telangiectasia mutated and Rad3 related (ATR) protein kinase as potential anticancer agents. *J Med Chem*. 2011;54(7):2320-30. Epub 2011/03/19.
288. Reaper PM, Griffiths MR, Long JM, et al. Selective killing of ATM- or p53-deficient cancer cells through inhibition of ATR. *Nat Chem Biol*. 2011;7(7):428-30. Epub 2011/04/15.
289. Toledo LI, Murga M, Zur R, et al. A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. *Nat Struct Mol Biol*. 2011;18(6):721-7. Epub 2011/05/10.
290. Koelzer VH, Zlobec I, Lugli A. Tumor budding in colorectal cancer--ready for diagnostic practice? *Hum Pathol*. 2016;47(1):4-19. Epub 2015/10/20.
291. Torabi K, Wangsa D, Ponsa I, et al. Transcription-dependent radial distribution of TCF7L2 regulated genes in chromosome territories. *Chromosoma*. 2017. Epub 2017/03/28.
292. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics*. 2007;23(6):657-63. Epub 2007/01/20.
293. Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19(9):1639-45. Epub 2009/06/23.
294. Smyth GK. Limma: linear models for microarray data. Gentleman R, Carey V, Dudoit S, et al., eds. New York: Springer; 2005.
295. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-50. Epub 2005/10/04.
296. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9(7):671-5. Epub 2012/08/30.
297. Camps J, Ponsa I, Ribas M, et al. Comprehensive measurement of chromosomal instability in cancer cells: combination of fluorescence in situ hybridization and cytokinesis-block micronucleus assay. *FASEB J*. 2005;19(7):828-30. Epub 2005/03/12.
298. Schrock E, du Manoir S, Veldman T, et al. Multicolor spectral karyotyping of human chromosomes. *Science*. 1996;273(5274):494-7. Epub 1996/07/26.
299. Schaffer LG, Slovak ML, Campbell LJ. An International System for Human Cytogenetic Nomenclature: Karger; 2009.

300. Chowdhury SA, Shackney SE, Heselmeyer-Haddad K, et al. Phylogenetic analysis of multiprobe fluorescence in situ hybridization data from tumor cell populations. *Bioinformatics*. 2013;29(13):i189-98. Epub 2013/07/03.
301. Auer GU, Caspersson TO, Wallgren AS. DNA content and survival in mammary carcinoma. *Anal Quant Cytol*. 1980;2(3):161-5. Epub 1980/09/01.
302. Di Nicolantonio F, Martini M, Molinari F, et al. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol*. 2008;26(35):5705-12. Epub 2008/11/13.
303. Patai AV, Molnar B, Tulassay Z, et al. Serrated pathway: alternative route to colorectal cancer. *World J Gastroenterol*. 2013;19(5):607-15. Epub 2013/02/23.
304. Ogden A, Rida PC, Aneja R. Heading off with the herd: how cancer cells might maneuver supernumerary centrosomes for directional migration. *Cancer Metastasis Rev*. 2013;32(1-2):269-87. Epub 2012/11/02.
305. Holme O, Bretthauer M, Fretheim A, et al. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. *Cochrane Database Syst Rev*. 2013(9):CD009259. Epub 2013/10/03.
306. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med*. 2012;366(8):687-96. Epub 2012/02/24.
307. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer*. 2013;108(3):479-85. Epub 2013/01/10.
308. Seike K, Koda K, Oda K, et al. Assessment of rectal aberrant crypt foci by standard chromoscopy and its predictive value for colonic advanced neoplasms. *Am J Gastroenterol*. 2006;101(6):1362-9. Epub 2006/06/15.
309. Anderson JC, Swede H, Rustagi T, et al. Aberrant crypt foci as predictors of colorectal neoplasia on repeat colonoscopy. *Cancer Causes Control*. 2012;23(2):355-61. Epub 2011/12/22.
310. Di Gregorio C, Losi L, Fante R, et al. Histology of aberrant crypt foci in the human colon. *Histopathology*. 1997;30(4):328-34. Epub 1997/04/01.
311. Shpitz B, Bomstein Y, Mekori Y, et al. Aberrant crypt foci in human colons: distribution and histomorphologic characteristics. *Hum Pathol*. 1998;29(5):469-75. Epub 1998/05/22.

312. Bosman FT, Carneiro, F., Hruban, R.H., Theise, N.D. WHO Classification of Tumours of the Digestive System. Paul Kleihues LHS, ed.: IARC Press; 2010.
313. Pinsky PF, Fleshman J, Mutch M, et al. One year recurrence of aberrant crypt foci. *Cancer Prev Res (Phila)*. 2010;3(7):839-43. Epub 2010/06/24.
314. Otori K, Konishi M, Sugiyama K, et al. Infrequent somatic mutation of the adenomatous polyposis coli gene in aberrant crypt foci of human colon tissue. *Cancer*. 1998;83(5):896-900. Epub 1998/09/10.
315. Mo A, Jackson S, Varma K, et al. Distinct Transcriptional Changes and Epithelial-Stromal Interactions Are Altered in Early-Stage Colon Cancer Development. *Mol Cancer Res*. 2016;14(9):795-804. Epub 2016/06/30.
316. Rosenberg DW YS, Pleau DC, Greenspan EJ, Stevens RG, Rajan TV, Heinen CD, Levine J, Zhou Y, O'Brien MJ. Mutations in BRAF and KRAS differentially distinguish serrated versus non-serrated hyperplastic aberrant crypt foci in humans. *Cancer Res*. 2007;67(8):3551-4. Epub 3554.
317. Inoue A, Okamoto K, Fujino Y, et al. B-RAF mutation and accumulated gene methylation in aberrant crypt foci (ACF), sessile serrated adenoma/polyp (SSA/P) and cancer in SSA/P. *Br J Cancer*. 2015;112(2):403-12. Epub 2014/10/15.
318. Fransen K, Klintenas M, Osterstrom A, et al. Mutation analysis of the BRAF, ARAF and RAF-1 genes in human colorectal adenocarcinomas. *Carcinogenesis*. 2004;25(4):527-33. Epub 2003/12/23.
319. Augenlicht LH, Richards C, Corner G, et al. Evidence for genomic instability in human colonic aberrant crypt foci. *Oncogene*. 1996;12(8):1767-72. Epub 1996/04/18.
320. Heinen CD, Shivapurkar N, Tang Z, et al. Microsatellite instability in aberrant crypt foci from human colons. *Cancer Res*. 1996;56(23):5339-41. Epub 1996/12/01.
321. Staffa L, Echterdiek F, Nelius N, et al. Mismatch repair-deficient crypt foci in Lynch syndrome--molecular alterations and association with clinical parameters. *PLoS One*. 2015;10(3):e0121980. Epub 2015/03/31.
322. Menigatti M, Pedroni M, Verrone AM, et al. O6-methylguanine-DNA methyltransferase promoter hypermethylation in colorectal carcinogenesis. *Oncol Rep*. 2007;17(6):1421-7. Epub 2007/05/10.

323. Nagasaka T, Goel A, Notohara K, et al. Methylation pattern of the O6-methylguanine-DNA methyltransferase gene in colon during progressive colorectal tumorigenesis. *Int J Cancer*. 2008;122(11):2429-36. Epub 2008/02/02.
324. Pavicic W, Joensuu EI, Nieminen T, et al. LINE-1 hypomethylation in familial and sporadic cancer. *J Mol Med (Berl)*. 2012;90(7):827-35. Epub 2012/01/10.
325. Yamada A, Minamiguchi S, Sakai Y, et al. Colorectal advanced neoplasms occur through dual carcinogenesis pathways in individuals with coexisting serrated polyps. *PLoS One*. 2014;9(5):e98059. Epub 2014/05/23.
326. Kaz AM, Wong CJ, Dzieciatkowski S, et al. Patterns of DNA methylation in the normal colon vary by anatomical location, gender, and age. *Epigenetics*. 2014;9(4):492-502. Epub 2014/01/15.
327. Estecio MR, Gharibyan V, Shen L, et al. LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability. *PLoS One*. 2007;2(5):e399. Epub 2007/05/04.
328. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481(7381):306-13. Epub 2012/01/20.
329. Navin N, Krasnitz A, Rodgers L, et al. Inferring tumor progression from genomic heterogeneity. *Genome Res*. 2010;20(1):68-80. Epub 2009/11/12.
330. Sottoriva A, Spiteri I, Piccirillo SG, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A*. 2013;110(10):4009-14. Epub 2013/02/16.
331. Gao R, Davis A, McDonald TO, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet*. 2016;48(10):1119-30. Epub 2016/08/16.
332. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-74. Epub 2011/03/08.
333. Siegmund KD, Marjoram P, Woo YJ, et al. Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. *Proc Natl Acad Sci U S A*. 2009;106(12):4828-33. Epub 2009/03/06.
334. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med*. 2017. Epub 2017/04/27.
335. Brosens RP, Belt EJ, Haan JC, et al. Deletion of chromosome 4q predicts outcome in stage II colon cancer patients. *Anal Cell Pathol (Amst)*. 2010;33(2):95-104. Epub 2010/10/23.

336. Nowak MA, Komarova NL, Sengupta A, et al. The role of chromosomal instability in tumor initiation. *Proc Natl Acad Sci U S A*. 2002;99(25):16226-31. Epub 2002/11/26.
337. Postma C, Koopman M, Buffart TE, et al. DNA copy number profiles of primary tumors as predictors of response to chemotherapy in advanced colorectal cancer. *Ann Oncol*. 2009;20(6):1048-56. Epub 2009/01/20.
338. R SD, Gutteridge A, Swanton C, et al. Modelling the evolution of genetic instability during tumour progression. *Evol Appl*. 2013;6(1):20-33. Epub 2013/02/12.
339. Sapkota Y, Ghosh S, Lai R, et al. Germline DNA copy number aberrations identified as potential prognostic factors for breast cancer recurrence. *PLoS One*. 2013;8(1):e53850. Epub 2013/01/24.
340. Mamlouk S, Childs LH, Aust D, et al. DNA copy number changes define spatial patterns of heterogeneity in colorectal cancer. *Nat Commun*. 2017;8:14093. Epub 2017/01/26.
341. Meijer GA, Hermsen MA, Baak JP, et al. Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. *J Clin Pathol*. 1998;51(12):901-9. Epub 1999/03/10.
342. Loo LW, Tiirikainen M, Cheng I, et al. Integrated analysis of genome-wide copy number alterations and gene expression in microsatellite stable, CpG island methylator phenotype-negative colon cancer. *Genes Chromosomes Cancer*. 2013;52(5):450-66. Epub 2013/01/24.
343. Diep CB, Kleivi K, Ribeiro FR, et al. The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. *Genes Chromosomes Cancer*. 2006;45(1):31-41. Epub 2005/09/08.
344. Lin CH, Lin JK, Chang SC, et al. Molecular profile and copy number analysis of sporadic colorectal cancer in Taiwan. *J Biomed Sci*. 2011;18:36. Epub 2011/06/08.
345. Ried T, Knutzen R, Steinbeck R, et al. Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. *Genes Chromosomes Cancer*. 1996;15(4):234-45. Epub 1996/04/01.
346. Sheffer M, Bacolod MD, Zuk O, et al. Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer. *Proc Natl Acad Sci U S A*. 2009;106(17):7131-6. Epub 2009/04/11.

347. Quinlan KG, Verger A, Yaswen P, et al. Amplification of zinc finger gene 217 (ZNF217) and cancer: when good fingers go bad. *Biochim Biophys Acta*. 2007;1775(2):333-40. Epub 2007/06/19.
348. Zhang ZC, Zheng LQ, Pan LJ, et al. ZNF217 is overexpressed and enhances cell migration and invasion in colorectal carcinoma. *Asian Pac J Cancer Prev*. 2015;16(6):2459-63. Epub 2015/04/01.
349. Takagi Y, Koumura H, Futamura M, et al. Somatic alterations of the SMAD-2 gene in human colorectal cancers. *Br J Cancer*. 1998;78(9):1152-5. Epub 1998/11/20.
350. Fleming NI, Jorissen RN, Mouradov D, et al. SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer Res*. 2013;73(2):725-35. Epub 2012/11/10.
351. Flora M, Piana S, Bassano C, et al. Epidermal growth factor receptor (EGFR) gene copy number in colorectal adenoma-carcinoma progression. *Cancer Genet*. 2012;205(12):630-5. Epub 2012/11/28.
352. Larsson A, Fridberg M, Gaber A, et al. Validation of podocalyxin-like protein as a biomarker of poor prognosis in colorectal cancer. *BMC Cancer*. 2012;12:282. Epub 2012/07/10.
353. Zhao D, Wang SH, Feng Y, et al. Intratumoral c-Met expression is associated with vascular endothelial growth factor C expression, lymphangiogenesis, and lymph node metastasis in oral squamous cell carcinoma: implications for use as a prognostic marker. *Hum Pathol*. 2011;42(10):1514-23. Epub 2011/05/03.
354. Lu Y, Jingyan G, Baorong S, et al. Expression of EGFR, Her2 predict lymph node metastasis (LNM)-associated metastasis in colorectal cancer. *Cancer Biomark*. 2012;11(5):219-26. Epub 2012/12/12.
355. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology*. 2008;135(4):1079-99. Epub 2008/09/09.
356. Salari K, Spulak ME, Cuff J, et al. CDX2 is an amplified lineage-survival oncogene in colorectal cancer. *Proc Natl Acad Sci U S A*. 2012;109(46):E3196-205. Epub 2012/11/01.
357. Bai YQ, Miyake S, Iwai T, et al. CDX2, a homeobox transcription factor, upregulates transcription of the p21/WAF1/CIP1 gene. *Oncogene*. 2003;22(39):7942-9. Epub 2003/09/13.

358. Firestein R, Shima K, Nosh K, et al. CDK8 expression in 470 colorectal cancers in relation to beta-catenin activation, other molecular alterations and patient survival. *Int J Cancer*. 2010;126(12):2863-73. Epub 2009/10/01.
359. Camps J, Pitt JJ, Emons G, et al. Genetic amplification of the NOTCH modulator LNX2 upregulates the WNT/beta-catenin pathway in colorectal cancer. *Cancer Res*. 2013;73(6):2003-13. Epub 2013/01/16.
360. Williams MJ, Werner B, Barnes CP, et al. Identification of neutral tumor evolution across cancer types. *Nat Genet*. 2016;48(3):238-44. Epub 2016/01/19.
361. Gerlinger M, Swanton C. How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br J Cancer*. 2010;103(8):1139-43. Epub 2010/09/30.
362. Zack TI, Schumacher SE, Carter SL, et al. Pan-cancer patterns of somatic copy number alteration. *Nat Genet*. 2013;45(10):1134-40. Epub 2013/09/28.
363. Gordon DJ, Resio B, Pellman D. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet*. 2012;13(3):189-203. Epub 2012/01/25.
364. Storchova Z, Kuffer C. The consequences of tetraploidy and aneuploidy. *J Cell Sci*. 2008;121(Pt 23):3859-66. Epub 2008/11/21.
365. Storchova Z, Pellman D. From polyploidy to aneuploidy, genome instability and cancer. *Nat Rev Mol Cell Biol*. 2004;5(1):45-54. Epub 2004/01/07.
366. Kuznetsova AY, Seget K, Moeller GK, et al. Chromosomal instability, tolerance of mitotic errors and multidrug resistance are promoted by tetraploidization in human cells. *Cell Cycle*. 2015;14(17):2810-20. Epub 2015/07/08.
367. Potapova TA, Seidel CW, Box AC, et al. Transcriptome analysis of tetraploid cells identifies cyclin D2 as a facilitator of adaptation to genome doubling in the presence of p53. *Mol Biol Cell*. 2016;27(20):3065-84. Epub 2016/08/26.
368. Dewhurst SM, McGranahan N, Burrell RA, et al. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discov*. 2014;4(2):175-85. Epub 2014/01/18.
369. Gentric G, Desdouets C, Celton-Morizur S. Hepatocytes polyploidization and cell cycle control in liver physiopathology. *Int J Hepatol*. 2012;2012:282430. Epub 2012/11/15.
370. Silkworth WT, Nardi IK, Paul R, et al. Timing of centrosome separation is important for accurate chromosome segregation. *Mol Biol Cell*. 2012;23(3):401-11. Epub 2011/12/02.

371. Ichijima Y, Yoshioka K, Yoshioka Y, et al. DNA lesions induced by replication stress trigger mitotic aberration and tetraploidy development. *PLoS One*. 2010;5(1):e8821. Epub 2010/01/26.
372. Chanoux RA, Yin B, Urtishak KA, et al. ATR and H2AX cooperate in maintaining genome stability under replication stress. *J Biol Chem*. 2009;284(9):5994-6003. Epub 2008/12/04.
373. Harrigan JA, Belotserkovskaya R, Coates J, et al. Replication stress induces 53BP1-containing OPT domains in G1 cells. *J Cell Biol*. 2011;193(1):97-108. Epub 2011/03/30.
374. Lamm N, Ben-David U, Golan-Lev T, et al. Genomic Instability in Human Pluripotent Stem Cells Arises from Replicative Stress and Chromosome Condensation Defects. *Cell Stem Cell*. 2016;18(2):253-61. Epub 2015/12/17.
375. Bakhoum SF, Kabeche L, Murnane JP, et al. DNA-damage response during mitosis induces whole-chromosome missegregation. *Cancer Discov*. 2014;4(11):1281-9. Epub 2014/08/12.
376. Pampalona J, Roscioli E, Silkworth WT, et al. Chromosome Bridges Maintain Kinetochore-Microtubule Attachment throughout Mitosis and Rarely Break during Anaphase. *PLoS One*. 2016;11(1):e0147420. Epub 2016/01/20.
377. Williams BR, Prabhu VR, Hunter KE, et al. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science*. 2008;322(5902):703-9. Epub 2008/11/01.
378. Selmecki AM, Maruvka YE, Richmond PA, et al. Polyploidy can drive rapid adaptation in yeast. *Nature*. 2015;519(7543):349-52. Epub 2015/03/04.
379. Lee AJ, Endesfelder D, Rowan AJ, et al. Chromosomal instability confers intrinsic multidrug resistance. *Cancer Res*. 2011;71(5):1858-70. Epub 2011/03/03.
380. Godinho SA, Picone R, Burute M, et al. Oncogene-like induction of cellular invasion from centrosome amplification. *Nature*. 2014;510(7503):167-71. Epub 2014/04/18.