SAPIENZA
Università di Roma
Facoltà di Scienze Matematiche Fisiche e Naturali

DOTTORATO DI RICERCA
IN GENETICA E BIOLOGIA MOLECOLARE

XXXI Ciclo
(A.A. 2017/2018)

Common Fragile Sites: updating the causes of their variability in different cell tissues

Dottorando
Klizia Maccaroni

Docente guida
Prof.ssa Franca Pelliccia

Tutore
Dott.ssa Francesca Degrassi

Coordinatore
Prof. Fulvio Cruciani

INDEX

ABSTRACT **1**

INTRODUCTION **2**

Definition of Fragile Sites 2

Classification of Fragile Sites 3

The expression of Fragile Sites in different cell types 7

Fragile Sites and Genomic Instability 8

The replication origins in Common Fragile Sites expression 14

The paucity of replication origins and Genomic Instability 16

Fragile Sites and Transcription 19

AIM OF THE RESEARCH **21**

RESULTS AND DISCUSSION **23**

Molecular characterization of Fragile Sites **25**

1p31.1 Fragile region 25

3q13.3 Fragile region 29

Replication Timing analysis of Fragile Regions **34**

Replication Timing analysis of 1p31.1 Fragile Site 36

Replication Timing analysis of 3q13.3 Fragile Site 40

Genomic Instability at 1p31.1 and 3q13.3 fragile regions **45**

CONCLUSIONS **48**

MATERIALS AND METHODS **50**

REFERENCES **58**

LIST OF PUBLICATIONS **69**

ABSTRACT

Common Fragile Sites (CFSs) are regions in which DNA is prone to gaps, breaks or constrictions visible on metaphase chromosomes when cells are under replicative stressful conditions.
CFSs are characterized by slow/late replication timing mainly due, among other characteristics, to nucleotide sequence which tend to form secondary structure, and to the number of active (or inducible) replication origins. Recent studies indicate that CFSs expression is associated with tissue specificity. In the first part of the work, induction and classification of CFSs in two human lung fibroblast cell line, IMR-90 and MRC-5, has been done. Cytogenetical identification of the most expressed CFSs in both fibroblast cell lines were done: 1p31.1 and 3q13.3, located on chromosome 1 and 3 respectively, are peculiar for this tissue. These regions have typical and confirmed CFSs' characteristics such as expression higher than 3%, high AT levels and enrichment in large genes.
Using genomic databases, searching for causes of their instability were done comparing percentage of repetitive elements among the CFSs, non-fragile regions (NFRA) and standard genomic sequences. These CFSs are characterized by presence of large genes, NEGR1 found in 1p31.1, LSAMP and ARHGAP31 in the most fragile region of 3q13.3, that could be co-responsible for their genomic instability.
Using probes delimitating fragile regions and combining FISH with IF anti-BrdU, analysis of relationship between replication timing and fragility was done. Furthermore, comparison between replication timing, in normal and stressful condition using APH, was done as well. The results obtained for these fragile regions reflect the replication timing impairments typical of fragile sites, in both normal and stressful conditions.
The same probes when used in lymphocytes result in a normal replication timing, moreover also using CFSs probes specific in lymphocytes on fibroblast, results in normal replication timing.
The results from replication timing analysis are strictly correlated with the structural and functional characteristics that are specific of the tissues in which these CFSs are expressed.

INTRODUCTION

Definition of Fragile Sites.

Fragile Sites (FSs) are regions in which DNA is prone to gaps, breakage or constriction that can be visualized on metaphase chromosomes when cells are under replicative stressful condition (Durkin and Glover, 2007); they represent about 1% of whole genome and their length spans from hundreds to thousands of kilobases.
The "fragile sites" definition has been used since 1970 when, on long arm of chromosome 16, recurrent breaks localized and were also under mendelian inheritance (Magenis et al., 1970).
Investigation on fragile sites was under a particular interest in human genetics since some of them, like FS in Xq27.3 region, were recurrent in some families but seemed also associated with some pathologies, as in mental retardation connected to chromosome X abnormalities (Harvey et al, 1977). Thanks to Sutherland, those irregularities on long arm of X chromosome were connected to mental retardation in male individuals in different Australian families; those chromosomal breaks were visible when cells were growth in particular condition. The removal of different elements from culture medium such as folic acid or thymine seemed to affect in some ways the metabolism of nucleotides and DNA replication, resulting in higher X chromosome's abnormalities, connecting mental retardation to this singularity on chromosome X (Sutherland, 1979; Glover, 1981).
The FSs are, moreover, regions in which, following replication, the sister chromatids exchange frequency is even higher compared to non-fragile regions, furthermore they are exogenous viral DNA integration sites, so they are recombinogenic and genetically unstable regions (Glover et al., 1984; Gaddini et al., 1995).
Despite these negative features, FSs can still be found in almost every organism, from prokaryotes to eukaryotes, suggesting evolutive positive selection for these regions and important biological functions for cells and their genome.

Classification of Fragile Sites.

The FSs can be categorized in two main classes based on their inheritance and population frequency:

- Rare Fragile Sites (RFSs),
- Common Fragile Sites (CFSs).

Class	Number of loci	Sequence
Rare fragile sites	31	
Folate-sensitive	24	(CGG)n
Distimycin A	5	AT-rich repeat[a]
BrdU	2	AT-rich repeat
Common fragile sites	87	
Aphidicolin	76	AT-rich
BrdU	7	AT-rich?[b]
5-Azacytidine	4	?[b]

Figure 1. Common and Rare Fragile Sites classification. Here is shown the Fragile Sites classification, the number of associated loci and the characterizing sequences (from Durkin and Glover, 2007).

The **Rare Fragile Sites** are expressed in less than 5% of all individuals in a population (Sutherland, 1984; Schwartz, 2006) and can be further categorized based on cells' culturing condition (figure 1). The folate-sensitive ones are the most expressed RFSs and are elicited when culture medium is folate deprived.
From a structural point of view, nucleotides repetitions expansions are responsible for these regions' fragility; microsatellites or minisatellites with CGG_n sequence, because of repetition number and variation in population frequency, can cause the occurrence of a fragile allele (Mc Murray, 2010) and, in some cases, can also be responsible for hereditary pathologies (Sutherland et al., 1984).
For this class the most known FS is FRAXA associated with X chromosome, which co-localize with FMR1 gene responsible for X-fragile syndrome; the FRAXE RFSs is also associated with mental retardation in bearer individuals, it co-localizes with FMR2 gene (Verkerk et al., 1991; Gu et al., 1996).

The other RFSs are expressed when medium is supplemented with distamycin A, an antibiotic, or Bromodeoxyuridine, a thymine analogue, but are characterized by AT-rich minisatellites expansion; the FRA10B (10q25.2) and FRA16B (16q22.1) are characterized respectively by highly repetitive 42 bp and 33 bp minisatellites expansion (Hewett et al., 1998; Yu et al., 1997).

The **Common Fragile Sites** are found in all normal individuals of a population, they are part of chromosomes' normal structure; they are categorized by differences in frequency for the same site analyzed and, from a molecular point of view, they are not characterized by nucleotides expansion (figure 1; Debatisse et al., 2012).
CFSs characterization started from studies conducted on male individual suffering from Martin-Bell syndrome characterized, among other features, by mental retardation; peripheral blood lymphocytes when cultured in folic acid deprivation conditions expressed a rare fragile site on long arm of X chromosome (Xq27.3), named FRAXA (Martin and Bell, 1943). Those observations made possible the found that another kind of fragile site, named Common Fragile Sites, are expressed when cells are cultured in replicative stressful conditions induced by molecules acting directly on enzymes responsible for DNA replication such as aphidicolin (APH). This molecule at specific concentrations inhibits the polymerase complex, acting on α, δ and ε subunits causing difficult replicative fork progression in specific DNA regions visible as gaps or breaks on metaphase chromosomes upon S phase ending (Glover et al., 1984).
Other molecules can induce CFSs expression acting on replicative machinery at different times; the nucleotides analogues 5-azacitidine and bromodeoxyuridine must be cited, but the CFSs induced by them are less known than those induced by APH. Even Hydroxyurea (HU) can induce CFSs but in a less specific way. Finally, environmental factors such as smoking, caffeine, oxygen deprivation and diet are among factors that increase CFSs expression frequency on human chromosomes (Pelliccia and Rocchi, 1992; Dillon et al., 2010).
As shown in figure 1, the APH induced ones are the human most expressed CFSs, however this classification should not be too strict, as their expression varies accordingly to APH concentration upon

cultured cells and, moreover, some genomic regions are more susceptible than others to breaks and lesions.
The exact number of CFSs it is currently material of debate because is dependent by different elements such as concentration and APH action duration, in fact, higher is the replicative stress upon cultured cells and higher will be CFSs expression frequency at the end of DNA replication. It has to be specified that not all CFSs show gaps and/or breaks with the same frequency and that only a small number of CFSs analyzed so far undergoes chromosomic breaks with a frequency higher than 2% (Glover et al., 1984; Yunis and Soreng, 1984); the breaks observed in just 20 CFSs embody alone more than 80% of lesions observed in all knowing lymphocytes' CFSs (Glover et al., 1984). FRA3B (3p14.3) is the most expressed CFS in lymphocytes (Ohta et al., 1996; Inoue et al., 1997; Mimori et al., 1999); the other highly expressed human lymphocytes CFSs are FRA2G (2q31) (Limongi et al., 2003), FRA16D (16q23.2) (Arlt et al., 2002; Ried et al., 2000), FRA6E (6q26) (Denison et al., 2003), FRA7H (7q32.3) (Mishmar et al., 1998) and FRAXB (Xp22.31) (Glover et al., 1984). Their chromosomal localization is shown in figure 2.
Besides showing differences in gaps or breaks expression, the CFSs show other genomic instability associated characteristics. In fact, a high Sister Chromatid Exchange (SCE) frequency is observed (Glover and Stain,1987; Gaddini et al., 1995) and even a very high rate of translocations and deletions in somatic cells hybrid systems (Glover and Stain, 1987; Wang et al., 1993); they are preferentially recombinational spots and plasmid DNA integration sites in transfected cells treated with APH (Rassol et al., 1991), this latter characteristic seems to correspond with the observation that in some tumors and cancer cell lines the CFSs and the viral integration sites are coincident (Wilke et al., 1996). CFSs can also promote chromosomal rearrangements and genomic instability, by inducing breaking-fusion-bridge cycles from breaks, this phenomenon can cause intra-chromosomal genic amplification as observed in different tumors and in vitro in cultured Chinese hamster ovary cells and tumor cell lines (Coquelle et al., 1997).

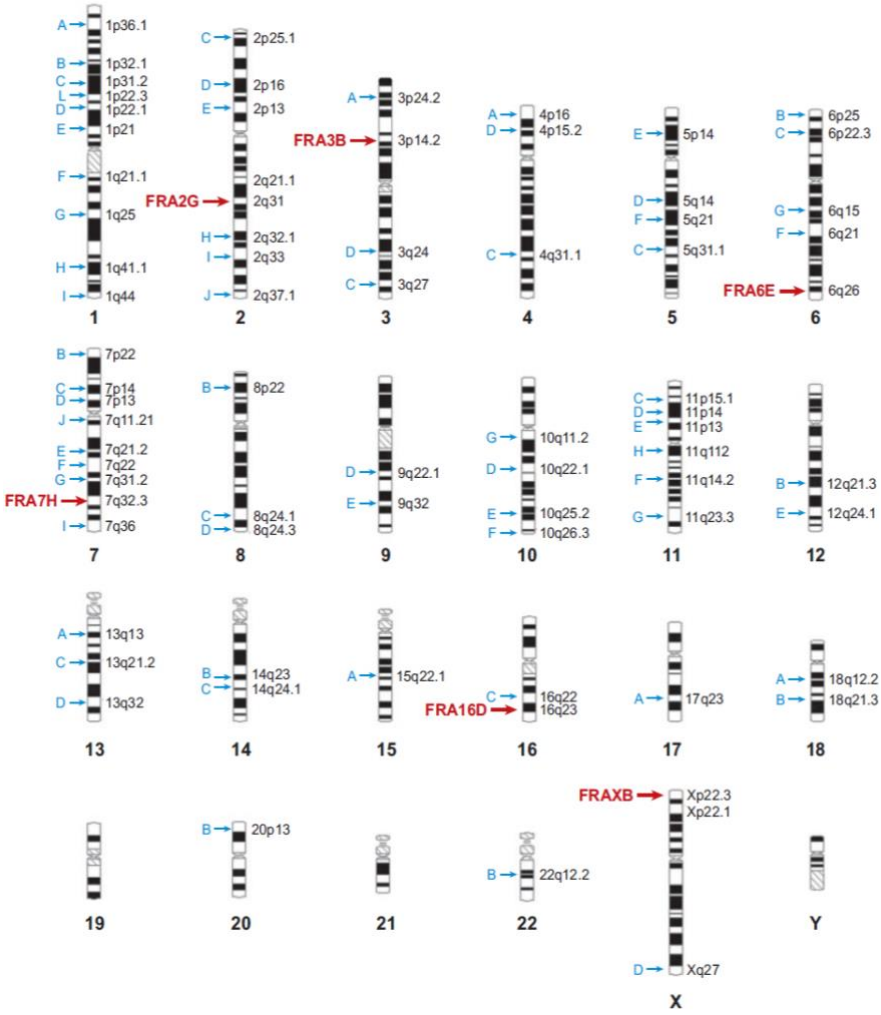


Figure 2. The Common Fragile Sites. The picture shows Common Fragile Sites induced by Aphidicolin and their chromosomes localization; in red are represented the most frequent Common Fragile Sites, in blue the less frequent ones (from Durkin and Glover, 2007)

Some CFSs can be of significant length and given the number of genes within the genome it is not surprising that some of the most expressed CFSs co-localize with one or more genes, this is the case of FRA3B and FRA16D CFSs that co-localize with FHIT and WWOX genes respectively (Ohta et al., 1996; Bednarek et al., 2000). FHIT is a 900

kb long gene with two large introns that centrally co-localizes with FRA3B; this CFS is not characterized by nucleotides repetitions expansion but shows the HPV16 virus integration sites responsible for cervical cancer (Thorland et al., 2000). FRA16D is co-localizing with WWOX gene that spans for over 1 Mb because of two very large introns (Ried et al., 2000).

The expression of Fragile Sites in different cell types.

Historically speaking, FSs have been extensively studied in human lymphocytes, mostly because of the simplicity in eliciting them in this cell type, recently however it has been found that there is a different CFSs expression among different cell types.
The CFSs and their expression can be extremely different among different cell types used for their studies (Murano et al., 1989), despite the pioneering studies on lymphocytes, different FSs have been observed among fibroblasts, epithelial colon cells lines, breast erythrocytes cell lines (Hosseini et al., 2013; le Tallec et al., 2013), adding more complexity to CFSs classification and expression.
Using genome-wide approaches such as Repli-seq combined with classic cytogenetic techniques, new CFSs have been identified on human fibroblast cell lines that differ from the ones localized on human lymphocytes (Le Tallec et al., 2011).
Comparing replication timing of the CFSs in fibroblasts and lymphocytes, in the latter the *core* (the most fragile region in a CFS) of FRA3B is replicated from flanking regions' replicative forks, and they must cover long distances before ending replication, suggesting a sharp paucity in replicative starting sites; in fibroblasts the same *core* region is replicated in late S phase arriving in G2 with more replicative forks localized within *core* region, pointing out a different number of DNA replication starting sites for the same region in different cell types (Letessier et al., 2011). Thus, the fragility depends on replicative molecular mechanisms, the same in every cell type, while differences

has to be found in the chromosomal regions which are different for each cell type (figure 3).

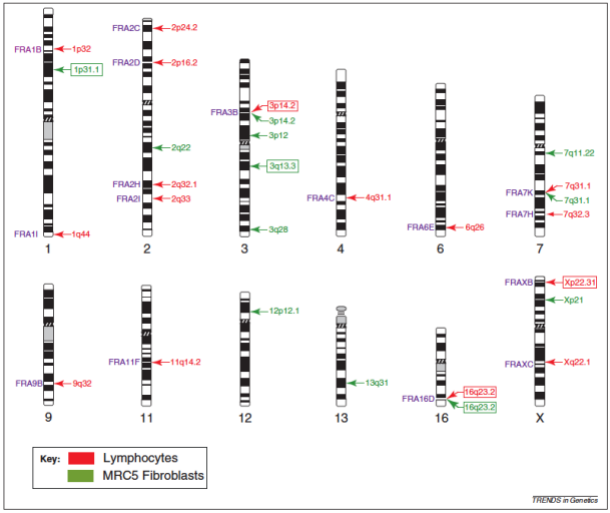


Figure 3. The Fragile Sites in lymphocytes and fibroblasts. The figure shows the sites that have a breaking frequency higher than 1,5%. The regions with higher breaks frequencies of the specific cell types are highlighted with red boxes (fibroblasts) and green boxes (lymphocytes) (From Debatisse et al., 2012).

Apart from fibroblasts and lymphocytes, nowadays the CFSs have been categorized in colon epithelial cell lines, breast cell lines and erythroid cell lines and many of their characterized loci are unstable in different tissues but their fragility could vary consistently from one cell type to another (Le Tallec et al., 2013).

Fragile Sites and Genomic Instability.

Fragile Sites are expressed when cells are under replicative stressful conditions, whether induced by molecules added in culture medium, either by external environmental factors like in CFSs, they can also be induced by chromosomal regions' molecular characteristics, like in RFSs; so fragile sites expression is connected to replication that must be completed without perturbation before cells enter metaphase in order to avoid any transmission of mutations.

In some higher eukaryotes like mammals, regions of replicative fork progression perturbations often appear and consequently the replication rate could vary from one region to another. This is due to eukaryotes' genome complexity both in size and structure, these factors can affect replication in different ways, they can be regions difficult to replicate because they are naturally subjected to slower fork progression rate or they are pausing regions (Cha and Kleckner, 2002; Branzei and Foiani, 2010).
In both RFSs and CFSs the replicative fork progression speed is the main cause in their induction, in fact, when the replicative fork progression is analyzed, it is often slowed down, so these regions can be defined as *late replicating regions*.
In RFSs the slower rate in replicative fork progression is due to their molecular characteristics, in fact either CGG_n trinucleotides expansions or repetitive AT-rich regions are present, and it is well known how the replication can be difficult when repetitive regions are present because the replicative machinery is more error prone. Furthermore, these regions can be difficult to replicate because they can lead to the formation of secondary structures known as *hairpin loops* following the intra-chromosomal non-covalent bond of complementary sequences that lead to a slower replication fork progression in attempt to solve them or, in more severe outcomes, to block the replication (Gacy et al., 1995; Hewett et al., 1998).
Even in CFSs a late replication timing is present. Studies analyzing the most expressed CFSs show that some of them can present difficult replicative fork progression regions and, because of that, chromosomal rearrangements and breaks are much more frequent, because they originate from incompletely replicated regions (Hellman et al., 2002; Palakodeti et al., 2004).
This is confirmed by APH mechanism of action upon chromosomal replication; it is used to induce CFSs expression. APH addition can significantly slow the replication of the most expressed lymphocyte's CFSs, the FRA3B, in which almost 16,5% of the fragile regions remain non-replicated upon entering the G2 phase (LeBeau et al., 1998).
Experiments on Xenopus's eggs show that replicative machinery is uncoupled from helicase-topoisomerase I molecular complex when APH is present in culture medium, this leads to an increased fragile

sites expression rate (Pacek et al., 2006; Byun et al., 2005); this experiment shows what happens in fragile regions when replication fork progression is altered.
These experiments allowed the formulation of a valid model to explain the fragile sites' expression mechanism that claims that CFSs are regions in which the replication is slowed down because the replicative machinery is uncoupled from helicase and topoisomerase complexes. Treating cells with camptothecin (CPT), a topoisomerase I inhibitor, the breaks on CFSs induced by APH are almost absent (Arlt and Glover, 2010).
Camptothecin is a chemical used as a chemotherapy drug in different type of tumors, it acts by reversibly binding the topoisomerase I complex preventing DNA re-ligation; high CPT doses are toxic for cells because they cause irreversible breaks on DNA. CPT can activate the S phase checkpoints' mechanisms that can stop the DNA synthesis even for hours after CPT removal (Horwitz et al., 1971; O'Connor et al., 1991).
In normal cells the helicase-topoisomerase I complex proceeds unwinding DNA so the polymerase, which is located behind the complex, begins to replicate unwinded regions, leaving few ssDNA not yet replicated (figure 4a).
When cells are treated with low doses of APH the polymerase slows the replication rate while the helicase-topoisomerase I complex advances in DNA unwinding leaving long regions of exposed non-replicated ssDNA that activates the checkpoint mechanisms (figure 4b).

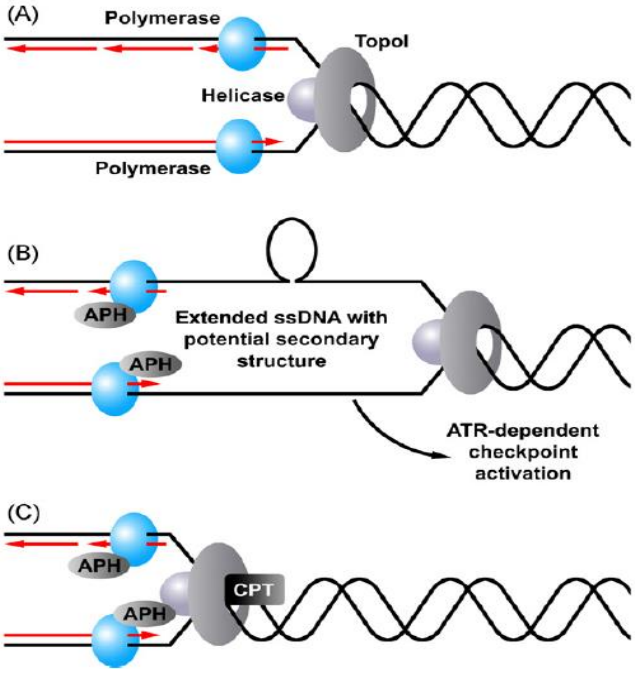


Figure 4. Camptothecin and aphidicolin breaks induction model on CFSs. (A) in normal conditions helicases and topoisomerase I complex proceeds together. (B) with APH addition the replicative complexes slow the replicative rate while the helicase-topoisomerase I proceeds leaving ssDNA non-replicated regions exposed. (C) in camptothecin and aphidicolin treated cells both complexes slow down (from Arlt and Glover, 2010).

These single strand regions could lead to secondary structure formation like hairpins or cruciform-like AT-rich secondary structures typical of common fragile regions leading to further slowdown of replicative complex; in normal conditions the secondary structure produced are identified and resolved, while other escapes from these control mechanisms causing the gaps or breaks to be visible on metaphase chromosomes, especially on CFSs regions (Pacek et al., 2006; Walter et al., 2000). Moreover, in presence of low doses of CPT the mechanisms that would resolve the secondary structures are slower.

Finally, when cells are treated with low doses of CPT and APH simultaneously, both the polymerasic and helicase-topoisomerase I

complexes slow down, the replication proceeds more slower and this seems to determinate a lower frequency of ssDNA regions, decreasing the secondary structure formation and breaks appearing on metaphase chromosomes in CFSs regions (figure 4c).
This shows how the decoupling of polymerase from helicase complex could be the first step in genomic instability associated to CFSs regions.

Fragile Sites could also be part of regions with early replication timing. Recent studies conducted on mouse B lymphocytes show that the cells' growth in stressful condition induced by hydroxyurea (HU) (Barlow et al., 2013) allows a review of replication timing in fragile regions. HU acts on S phase, inhibiting the ribonucleotide reductase, lowing the synthesis of deoxyribonucleotides from ribonucleotides (Szekeres et al., 1997). The damage induced by this drug can be found in an early phase of the cell cycle, causing DNA damage in a different way than APH, triggering the definition of a new class of fragile sites named *Early Replicating Fragile Sites* (ERFSs).
The characteristics associated with this new class of fragile sites are opposite to CFSs, starting from their expression timing during cell cycle: ERFSs in fact are expressed during the earliest phases of cell cycle, while it is known that late replication timing is associated with CFSs expression, moreover ERFSs are expressed in regions enriched in replication origins and coding regions.
ERFSs activate a different damage signaling response pathway: immunofluorescence assays against γ-H2AX, the phosphorylated H2AX histone variant, a marker for dsDNA damaged regions, showed that ERFSs are enriched for this marker, whereas CFSs are not (Seo et al., 2013).
Finally, it has to be said that there are some similarities between the two classes, in fact they are both enriched with CpG islands (Mortusewicz et al., 2013), this implies that more studies must be conducted on both classes in order to further know the characteristics and the similarities between these regions.

Another observation made to explain the genomic instability observed within fragile regions is about the *angle of twist* on both common and rare fragile sites. The angle of twist is the angle between the reference

plane of a base couple before twisting, and the same plane after twisting and it is not always the same between the nucleotides: it is dependent upon nucleotides sequence and it is important in the determination of the stacking strength within the double helix.
The *FlexStab* software is made to analyze the nucleotides' angle of twist and the stacking in the fragile regions; the software on FRA7H (7q32.3) fragile region shows that many flexible regions are present, named "flexibility peaks" (Mishmar, 1998; Mishmar et al., 1999), the same thing has been found on CFSs FRA2G (2q31), FRA3B (3p14.2), FRAXB (Xp22.31), FRA7E (7q21.2) and FRA16D (16q23.2) (Lukusa and Fryns, 2008). This is informing on the fact that even if CFSs are not characterized by nucleotides expansion as in RFSs, all the FSs are characterized by high rate of AT rich sequences in which the fragility must be found. In fact, it has been demonstrated, that the flexible sequences are enriched in AT repetitions that have a similar outcome to AT rich minisatellites in terms of fragility responsible for the fragility of FRA16B and FRA10B RFSs; these sequences can form hairpins or cruciform-like secondary structures that could cause the replicative fork progression stall in order to resolve them, promoting genomic instability and the fragile sites expression (Zlotorynski et al., 2003), or even the activation of checkpoint mechanisms that prevent the mitosis entry as the worst outcome.

The fragility of FSs could be explained by distinctive sequences, visible as gaps or breaks on chromosomes and, by removing them, the genomic instability can be prevented. In fact, studies on tumor cell lines seem to offer a confirmation to this hypothesis; by deleting 500 kb of DNA in FRAXB fragile site the fragility has been totally removed (Arlt et al., 2002). However, subsequent studies on hybrid cells containing chromosome 3 with large deletion on FRA3B fragile regions, did not give the same results in terms of reducing the fragility (Corbin et al., 2002). Recent studies on somatic hybrid cells with a deletion that spans from 200 to 600 kb in FRA3B, removing many flexibility peaks, demonstrated a reduction in fragility, not a complete eradication.
So, the fragile sites instability could be a characteristic of regions delimitating the fragile regions rather than a feature of the fragile site itself.

The replication origins in Common Fragile Sites expression.

DNA replication allows transmission of genetic material from an offspring to another when cell division occurs. In this process replication origins are extremely important; complementary double strands are separated, and bidirectional replication occurs.
In mammals, replication must be fast to replicate correctly all the genomic material, at least 30,000 to 50,000 replication origins are activated in each cell cycle; this value exceeds the number of origins effectively activated during DNA replication, in fact, it is known that not all replication origins are simultaneously activated, but follow an accurate timing related to several factors such as cellular type or the healthy state of the cell (Huberman and Riggs, 1996).
The characteristics of DNA's replication origins can be briefly summarized (figure 5) analyzing them at different levels.
Starting from a sequence point of view, two AT-rich regions are present, known as DUE (*DNA Unwinding Elements*) in which DNA is more easily unwinded, given to the presence of two hydrogen bonds connecting the AT bases: from here, presumably, replication starts; CpG islands are present too, helping double helix unwinding.
From a structural point of view cruciform structure and loops formation have been described as able to alter fork progression speed.
At chromatin level, nucleosome-free regions, histone acetylation and DNase-sensitive sites have been seen, but their direct participation in origin recognition or chromatin organization for transcription is sometimes difficult to estimate. The possible links between transcription factors with replication origin recognition have been described but evidences for their direct interactions remain scarce.

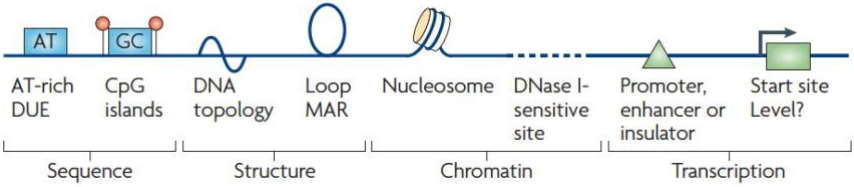


Figure 5. Structural characteristics of DNA replication origins. The figure shows the characteristics of DNA replication origins observed within eukaryotes; many of these characteristics have been described following the metazoan replication origins but they are not shared by all origins (Méchali, 2010).

Not all replication origins in the genome are simultaneously activated because their activation depends on many factors. First of all, replication origins are activated in a very precise moment of S phase, this timing permits a temporal classification of replication origins in *flexible, dormant (*or *inactive)* and *constitutive,* based on when the replicative helicases (*MCM*) are assembled on *pre-Replicative Complex* (*pre-RC*). The pre-RC is a multiprotein complex formed on replication origins during DNA replication initial moments, it is fundamental for the replication starting process; it is composed of six ORC proteins (*ORC 1-6), Cdc6, Cdt1* and the six proteins *MCM* (*MCM2-7*) (Yekezare et al., 2013).
Although not all replication origins are activated, they are activated as fast as possible under particular conditions that could change S phase state; changes able to interfere with cells' internal state leading to replication origins activation can be endogenous or exogenous: DNA damage for example or changes in growth condition of cell itself (Gilbert 2007).
As mentioned before, three classes of replication origins can be described: *flexible, dormant (*or *inactive)* and *constitutive.*
The *constitutive* are used in each DNA replication and are the same in every cell type even if they represent the minority among DNA replication origins.
The *flexibles* are activated stochastically at every cell cycle in different cells, their presence could also be explained by the presence of a 50 kb "initiation zone" in which is possible to find many replication origins activated indifferently (Mesner et al., 2003). The resulting pattern will reflect the sum of all individual situations and the stochastic nature of origin activation in this locus. If some origins are deleted, others nearby become more active or more efficient, reflecting a large choice of origins (Mesner et al., 2003; Kalejta et al., 1998). An explanation to the existence of flexible origins could be that in every cell, proteins responsible for replication to start are limiting and cannot bind every replication origin upon DNA, this will bring to the activation of the solely origins bonded by all molecular component useful to replication firing. Lastly, they could be easily activated by favorable chromatin domain in which origins are localized making their activation advantageous for replication to start, this has been observed in sub-telomeric regions where many replication origins are

located, and it is well known that these regions are rich in heterochromatin structurally complex compared to other regions (Hayashi et al., 2007).
Inactive or *dormant* replication origins are potential origins not activated in normal condition but could been potentially fired following stressful condition or specific cell signals.

The paucity of replication origins and Genomic Instability.

Many fragile sites contain large genomic regions in which replication origins activation is not efficient, this seems to be one of the main causes of late DNA replication and non-replicated regions during mitosis. This late replication timing could be due to the number of replication origins in fragile regions and surrounding area.
This has been extensively studied in lymphocytes, particularly in FRA3B, the most expressed human lymphocytes CFS.
Replication timing along FRA3B is so reduced that replication completion happens only in late G2 in which more than 10% of fragile sites is not replicated (Le Beau et al., 1998; Wang et al., 1999). Analysis of fork progression speed without stressing condition between FRA3B and the rest of the genome shows no significant differences, similarly, when cells are treated with stressing agents such as APH, fork progression's speed is drastically reduced but, again, there are no significant differences between FRA3B and genome. The differences appeared when FRA3B replication origins were identified with DNA combing technique, the presence of a 700 kb long region defined *core* within the FHIT (*Fragile Histidine Triad Protein*) gene lacking in replication origins and coincident with the most fragile region in FRAB can be identified.
When in normal condition, the replication forks delimitating core region covers long distances before ending replication (figure 6a). If stressing agents like APH are present, replication forks activated in core regions are lower than non-treated cells, as well as replicative termination events, the last is a clear index of a defect in replication completion (figure 6b) (Letessier et al., 2011).

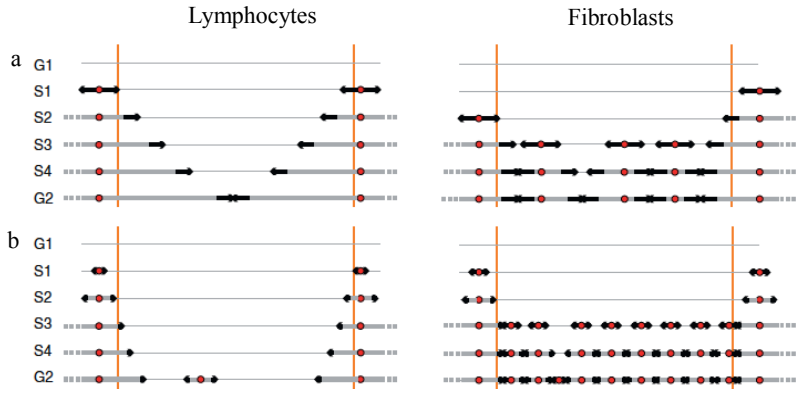


Figure 6. Fragile Sites and replication. Schematic representation of the core region replication in human FRA3B lymphocytes (left) and fibroblasts (right). The cell cycle phases are symbolized, the orange axes delimitate the core region. The upper images show normal conditions, the lower one shows the replication in stressful condition APH-induced (Debatisse et al., 2012).

Studies conducted comparing FRA3B expression in different cell types show a tissue-specificity in fragility that can impact the percentage of fragile sites expression and it is subject to the number of active replication origins, these findings suggest a role of replication origins in fragility along with nucleotide sequence.
From comparison of FRA3B expression between lymphocytes and two type of fibroblasts, MRC-5 and BJ (fibroblasts from fetal lung epithelium and primary epidermis fibroblasts respectively), difference of FRA3B expression is really marked with the highest to be found in lymphocytes. This is explicated by different number of active replication origins found in core flanking regions, showing a tissue-specificity already in DNA replication; in fibroblast it is possible to find a higher number of replication origins flanking core region that, in stressful condition by APH addition are much more activated than in the same lymphocyte's region, yielding to replication completion before mitosis, explaining the lower presence of breaks in fibroblast compared to lymphocytes (figure 6).
This has been observed also in another CFS, FRA6E, in which a 500 kb long region delimitating central portion of PARK2 (*Parkinson Protein* 2) gene shows few replication origins and analysis of replication fork progression's speed gave results similar to FRA3B

and FHIT gene; dissimilarities were given by intrinsic variances of the two fragile sites like number of core's replication origins responsible for both fragile regions fragility, FRA3B, in fact, shows a breaks percentage 5 times higher than FRA6E since the former contains less core's replication origins (Mrasek et al., 2010).
What has been said strengthens the idea that regions in which replicative forks cover long distances are more subjected to replicative fork progression slowing down, moreover, replication origins paucity has an important role in fragile sites' instability indicating a tissue-specificity in presence of active replication origins (Masai et al., 2010).
Regarding replication upon CFSs, a recent evidence shows the role of replicative polymerases able to switch with stalled specialized DNA polymerase δ; using an *in vitro* model of lagging strand replication and DNA template sequences from two different fragile sites, it has been demonstrated that either Pol η or κ, members of the Y-family and extensively studied for their role during DNA lesions, are capable of rescuing a pol δ stalled at repetitive non-B structure by performing synthesis on molecules initially extended, then stalled, by Pol δ within the CFS repeats instead of initiating synthesis on unextended molecules (Barnes et al., 2017). For CFS region that uses a single origin/replication fork during normal conditions may lack a compensatory fork during replication stress conditions (Le Tallec et al., 2011; Letessier et al., 2011). When this is the case, fork restart mechanisms that engage specialized DNA polymerases such as Pols η, κ, and ζ would become indispensable for maintaining CFS stability.

Fragile Sites and Transcription.

Many Fragile Sites colocalize within large genes (McAvoy et al., 2007), these large genes have many molecular characteristics in common with regions defined as "fragile", in fact they are often localized in late replicating domains or in AT-rich domains and those elements could promote genomic instability (Durkin et al., 2007).
Regarding the transcription of these large genes', it has been demonstrated that could take more than one cell cycle to be completed leading to simultaneous transcription and replication: the two

molecular machineries could collide leading to DNA breaks, that is why in eukaryotes replication and transcription are two temporally well separated events.
Recent studies demonstrated that upon transcription of genes larger than 800 kb the completion goes for a second cell cycle event affecting also part of S phase, this delay is also due to large genes' transcription which starts in late G2 phase, if not even M phase.
Instability observed within these large genes is mainly due to R-loops; the R-loops are RNA:DNA hybrids that form when nascent mRNA strand binds complementary ssDNA resulting from double helix denaturation upon replicative fork progression; this phenomena is possible only if replicative and transcriptional machinery are close to each other. As said before, a large gene is transcribed in G2-M phase and the transcription is completed when RNA Pol II reaches the termination site, when nascent pre-mRNA strand binds to mitotic chromosomes, regions in which non-condensed chromatin could arise visible as CFSs on metaphase chromosomes. Moreover, when pre-mRNA processing is impaired, R-loops are formed and they are among the primary cause of replicative fork progression stall and of double strands breaks of these regions (Gomez-Gonzalez et al., 2009; Tuduri et al., 2009). When DNA replication progression is impaired, for example upon APH addition, breaks depending on R-loops gather near large genes, so RNA:DNA hybrids are formed when transcription machinery collides with slowed-down or blocked replicative machinery (figure 7).

Topoisomerase I (TOP I) has a major role in R-loops resolving. TOP I is an essential enzyme for replication, functioning cutting one of the two DNA strands, relaxing and annealing it when supercoiled DNA forms: if not resolved, breaks upon double strand helix could arise impairing normal replicative fork progression (Wang 2002); when TOP I is absent, chromosomal breaks, both in fragile and non-fragile regions, arise even in cells in which replication is not impaired (Tuduri et al., 2009). So, instability observed within CFSs could come from slow replication or from TOP I loss or malfunctioning (Helmrich et al., 2011).

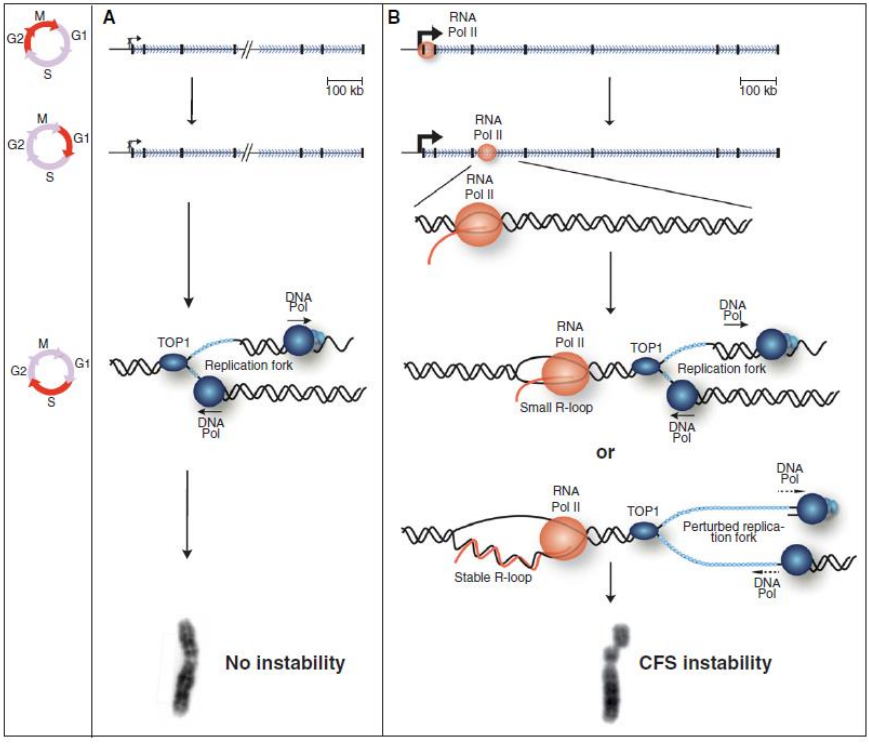


Figure 7. Collision between replicative and transcriptional machineries bring to R-loops formation. Large genes begin to be transcribed in G2 or M phase, when the machinery meets a stressed replicative fork R-loops are formed, bringing to CFSs formation. The right image shows the current model in large genes' R-loops formation, at left normal condition upon silent genes are shown (Helmrich et al., 2011).

AIM OF THE RESEARCH

Common Fragile Sites (CFSs) are important elements in promoting genomic instability in human genome.
They are characterized by impairments in replication due to AT-rich nucleotide sequences that tend to form secondary structures and by the number of active (or *activable*) replication origins responsible for the replication of the entire region: all these characteristics make the replication timing of CFSs to be *late* or *delayed*. Many pathologies, including different type of tumors and neurodegenerative diseases in particular, can be generated by aberrations involving breaks at CFSs regions.
Replication timing can vary among different cell types mostly because of the number of active replication origins so, given the connection between CFSs and replication, I searched if a connection between tissue specificity and CFSs expression was present. In fact, historically speaking, CFSs have been discovered and extensively studied in human lymphocytes, mostly because of the simplicity in eliciting them in this cell type, giving the possibility to understand many of the molecular and structural characteristics of the CFSs. An explanation to the tissue specificity of CFSs expression will be helpful in understanding the correlation between their localization and chromosomal rearrangements in tumor and diseases from different cell types.
I started the analysis from two fibroblast cell lines, both from fetal lung epithelium, MRC-5 and IMR-90. Comparing the CFSs induced by APH in fibroblasts with the most expressed CFSs in lymphocytes, among the others, were 1p31.1 on short arm of chromosome 1 and 3q13.3 on long arm of chromosome 3, exclusively expressed in fibroblasts.
Their molecular characterization was performed using BAC clones and fluorescent *in situ* hybridization (FISH) in order to define their exact cytogenetic localization. Next, the database consulting was used to characterize the nucleotides composition, such as AT/GC percentage, and the relative content in LINEs, Alu, miRNAs and LTR elements to see if their presence could be involved in promoting

fragility. The same elements were confronted with non-fragile regions (NFRA) and standard genome sequences with similar GC content. Even the presence of actively transcribing genes was investigated.

In the second part of my work I analyzed the CFSs replication in order to investigate the relationship between DNA replication timing and fragility.
Using FISH combined with IF anti-BrdU on interphasic nuclei, it was possible to observe their replicative state using fragile regions from lymphocytes on fibroblasts as controls and it was possible to analyze their peculiarities in replication timing confronting fibroblasts CFSs replication in lymphocytes in which these regions are non-fragile.

In the last part of my work I searched on Mitelman database if these fragile regions were subjected to chromosomal instability in different type of tumors and diseases.

RESULTS AND DISCUSSION

We induced CFSs in both MRC-5 and IMR-90 fibroblast cell lines cultured with medium added with APH to induce replicative stress. The breaks have been identified with Giemsa staining, recognized and localized cytogenetically with R-banding (figure 8).

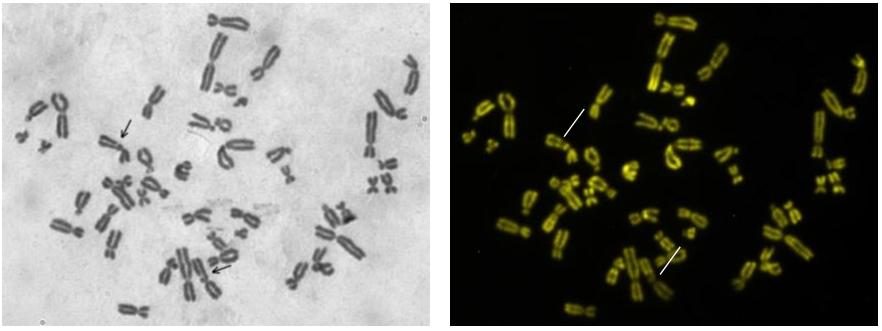


Figure 8. IMR-90 metaphase stained with Giemsa (left) and CMA3 (right). The arrows show breaks on chromosomes 7 (7q11.2; upper left) and 3 (3q13.3; down).

The putative CFSs identified are present in fibroblasts and not in lymphocytes apart from FRA3B and FRA16D, the two most expressed CFSs in human genome.
To be classified as a CFS, a break must be expressed with a frequency higher than 3%.
In our fibroblast cell lines, we chose to analyze the following putative CFSs:

- **1p31.1** located on chromosome 1 with a 19% frequency of expression in MRC-5 and 5% frequency of expression in IMR-90;
- **3q13.3** located on chromosome 3 with an 8% frequency of expression in MRC-5 and 27% frequency of expression in IMR-90.

The histogram below shows the most expressed breaks in both fibroblast cell lines (figure 9).

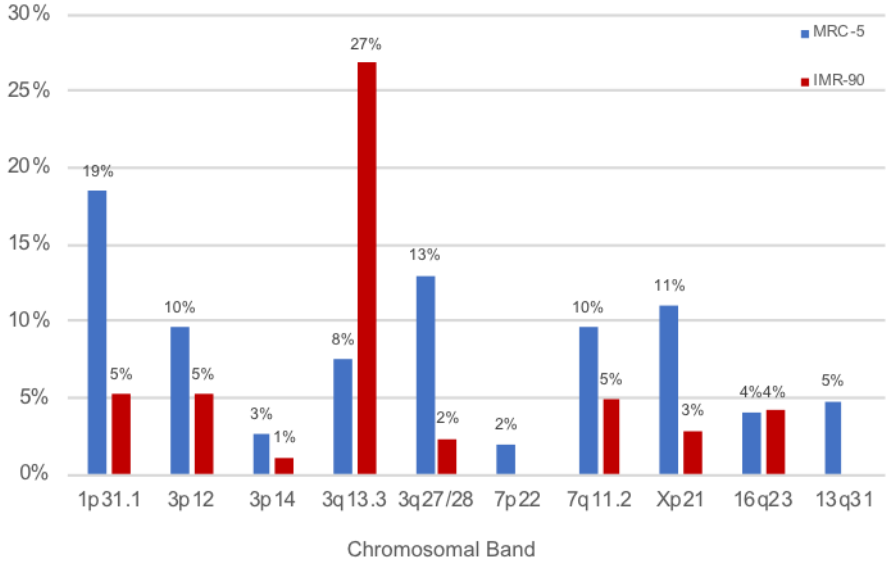


Figure 9. CFSs expression in fibroblasts. The histogram shows all the breaks expressed in both MRC-5 and IMR-90 cell lines; the breaks with frequency lower than 3% have not been considered.

The results obtained show a different frequency of expression for the same CFS in both fibroblast cell lines used for our analysis.
It is well known that different individuals among the same population show variations in frequency for the expression of the same fragile site. In this specific case, these data may be explained by differences in gender, developmental stage and also, possibly, by a different transcriptional activity and replication profile in both cell lines.

Molecular characterization of Fragile Sites.

1p31.1 Fragile Region.

The 1p31.1 is located on the short arm of chromosome 1 for about 4 Mb.

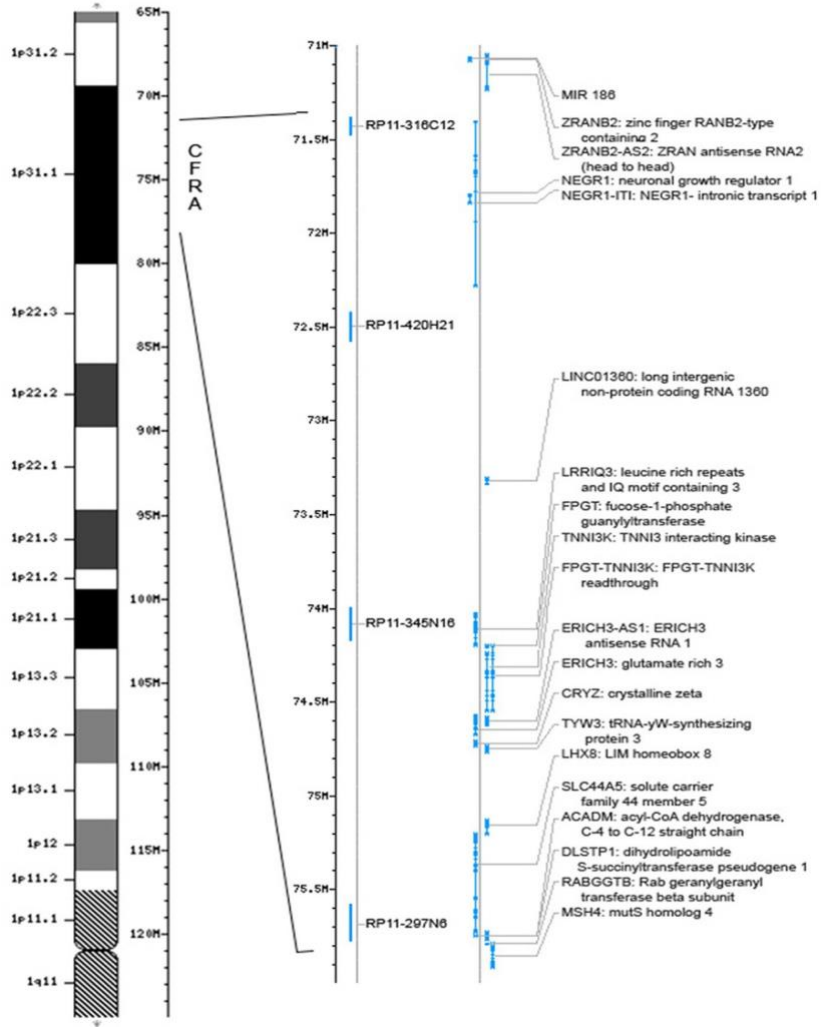


Figure 10. Schematic representation of 1p31.3 fragile region. The BAC clones used for the characterization, along with some genes, are shown (adapted from www.ncbi.nlm.nih.gov/mapview).

The FISH signals with the two chosen probes (*distal* RP11-316C12; *proximal* RP11-297N6) localize the fragile region in a G-band (figure 10).

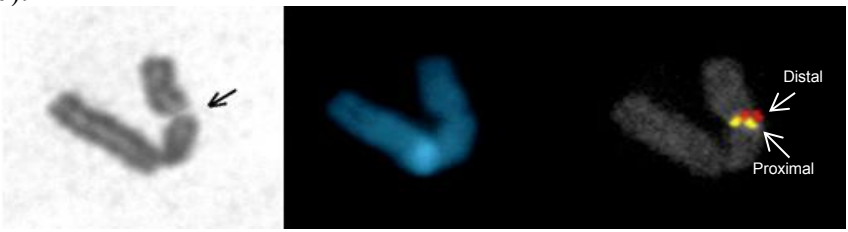


Figure 11. 1p31.1 fragile region. On the left, break on short arm of chromosome 1 (1p31.1) is identified with Giemsa staining; in the central figure the same chromosome is visualized with DAPI; in the right figure distal and proximal probes are localized, in the same chromosome, through FISH experiment.

The Genome Data Viewer NCBI database (build 37) locates 52 different genes along the fragile region (figure 10), many of these genes regulate important cellular features; a long non-coding RNA is also present (*LINC01360*), and a miR (*miR186*) known to function as tumor suppressor in many solid tumors can also be found.
It is known the role of large genes in promoting fragile sites' instability (Smith et al., 2007) due to R-loops formation; in the most fragile region is located an 886 kb long gene, NEGR1 (*Neuronal Growth Regulator 1*). This gene is highly expressed in brain and is involved in protein metabolism pathways and post-translational modification-synthesis of GPI-anchored proteins; an important paralog of this gene is LSAMP, located in 3q13.3, the other fragile region analyzed in fibroblasts.
Using Replication Domain database (www.replicationdomain.org), the replicative state of the gene was analyzed (figure 12); the replication pattern shows a late replication timing. In fact, this is in concordance with studies which hypothesize that large genes replicate late (Helmrich et al., 2011) and have a replication origin scarcity and, as a consequence, are particularly sensitive to replication stress because stalling of converging forks cannot be rescued by dormant

origins, preventing these regions from completing replication before entry into mitosis (Oestergaard and Lisby, 2017).
The fact that this large gene shows a late replication timing suggests

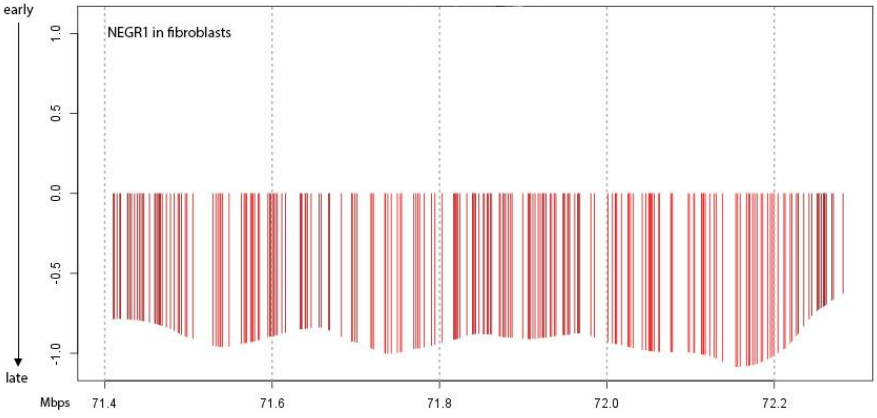


Figure 12. Schematic representation of region 71,402,942 - 72,282,594 bp of chromosome 1 where NEGR1 can be found; the graph shows the late replication timing of NEGR1 gene (from replicationdomain.org).

that could promote genomic instability by forming, probably, R-loops. For molecular characterization www.repeatmasker.org public database was used to analyze the sequence in order to find characteristics to explain the fragility; AT content, LINEs, Alu, LTRs

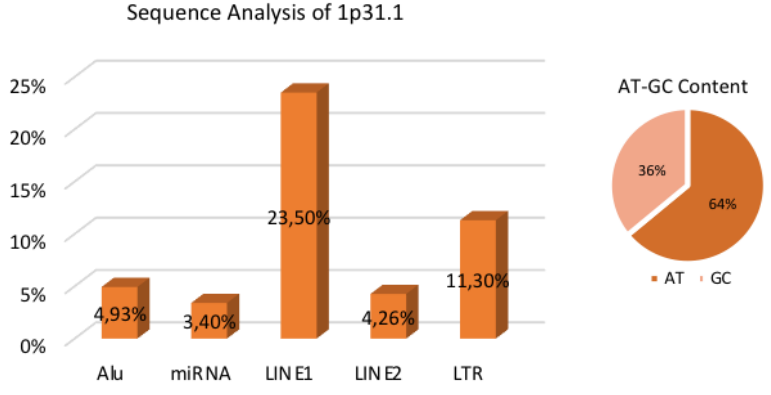
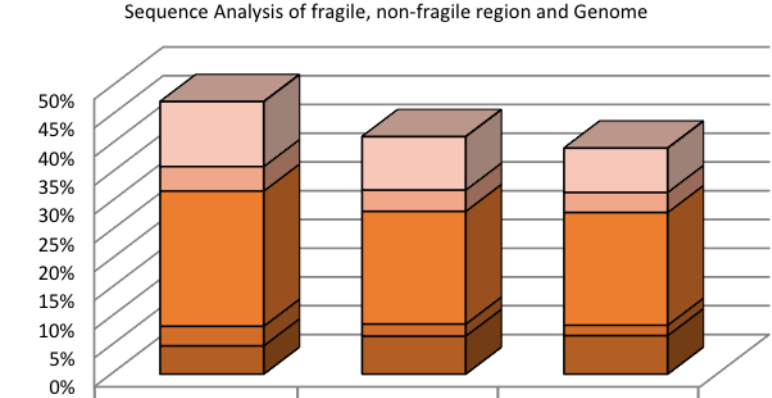


Figure 13. The histogram (left) shows the percentage of Alu, miRNAs, LINE1-2, LTR elements along the 1p31.1 region; on the right GC-AT content is shown.

and miRNAs have been investigated for their role in promoting genomic plasticity in higher eukaryotes (figure 13; left**)**.
The sequence analysis was done by examining all the sequences along the fragile region, moving from telomeric to centromeric extremity. The repeatmasker.org database results showed that this region is not enriched for Alu, miRNAs and LINE2 elements but a high percentage of LINE 1 elements (23,5%) is observed; this enrichment finds validation in literature data since LINEs elements are preferentially transposed in AT rich regions and are AT-rich themselves, in concordance with high levels of AT (64%) in 1p31.1 fragile site, which is a very high value (figure 13; right**)**.
The sequence analysis is then compared with non-fragile regions (NFRA) and standard genomic sequences with similar AT content (Smit 1999), to investigate the presence of any difference between different regions. The comparisons showed no significative differences between 1p31.1 fragile region, NFRA and genomic sequences (figure 14), apart from slightly enrichment in LINE1, LTR and miRNAs.



	1p31.1	NFRA	Human Genome
LTR	11,30%	9,30%	7,70%
LINE 2	4,26%	3,70%	3,50%
LINE 1	23,50%	19,60%	19,60%
miRNA	3,40%	2,10%	1,80%
Alu	4,93%	6,60%	6,70%

Figure 14. Sequence analysis and comparison between fragile region (1p31.1), non-fragile region (NFRA) and standard Genomic Sequence (Human Genome).

The results suggest a marginal role of these elements in promoting genomic instability of 1p31.1 fragile site, in concordance with literature data in which the fragile regions are not enriched of these elements, even in other tissues.

3q13.3 Fragile Region.

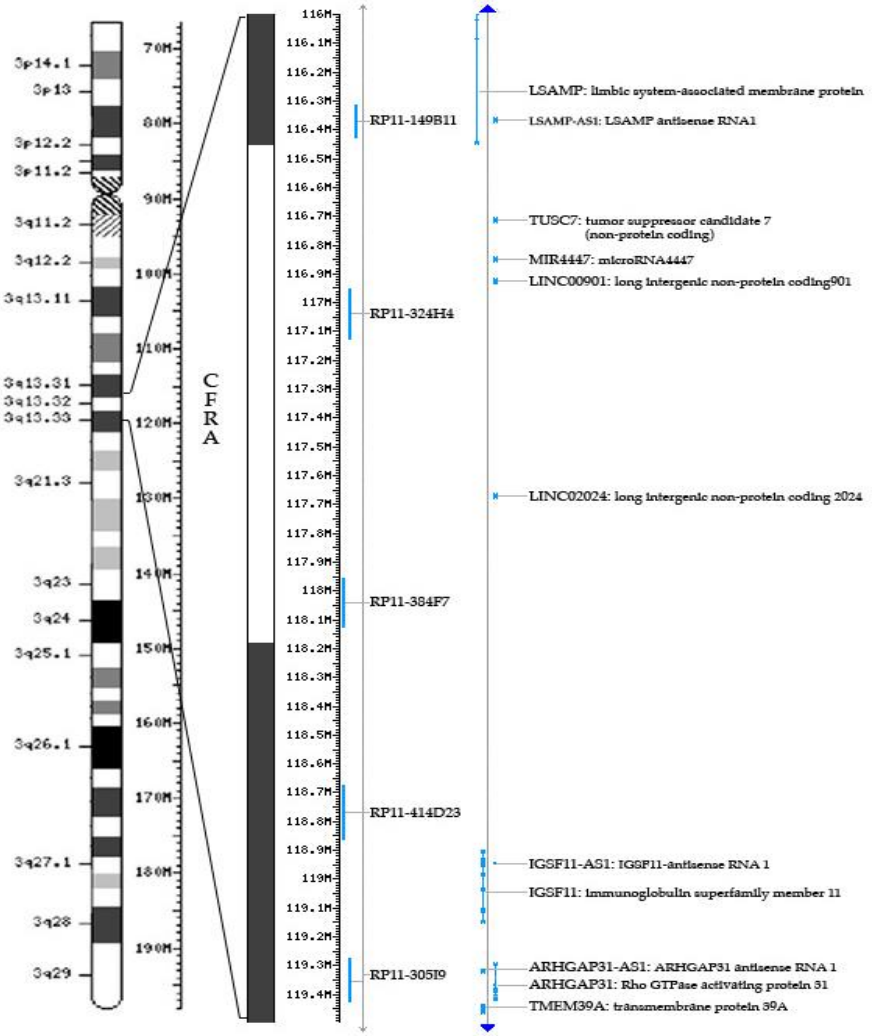


Figure 15. Schematic representation of 3q13.3 fragile region; the BACs used for characterization are shown along with some genes (adapted from www.ncbi.nlm.nih.gov/mapview).

The fragile region on long arm of chromosome 3 spans for more than 4 Mb in 3q13.3 chromosomal region; the probes chosen for FISH analysis (*distal* RP11-324H4; *proximal* RP11-305I9) localize 3q13.3 between a G and a R-band and could partially explain the high level of expression of this fragile site, in fact it is known that instability associated with FSs is higher at AT-GC interface (G/R-bands) because of a greater difference in corresponding twist angles (Mishmar et al., 1999).
The database analysis revealed the presence of 64 genes and two of them localize in the most fragile region; LSAMP (*Limbic System-Associated Membrane Protein)* and ARHGAP31 (*Rho GTPase activating protein 31*) are 1,33 Mb and 126 kbs long respectively and could be involved in fragility of 3q13.3 fragile site. miRNAs and long non-coding RNAs can be found as well (figure 15).
LSAMP, mapping in the fragile site proximal region, encodes for a member of immunoglobulin LAMP, OBCAM and neurotrimin (IgLON) family of proteins, contributes to the guidance of developing axons and remodeling of mature circuits in the limbic system. Known to function as tumor suppressor, its expression is high in brain, bladder and prostate. The replication timing analysis showed a delayed replication (figure 16), that could promote fragility by forming R-loops with mechanisms explained before.

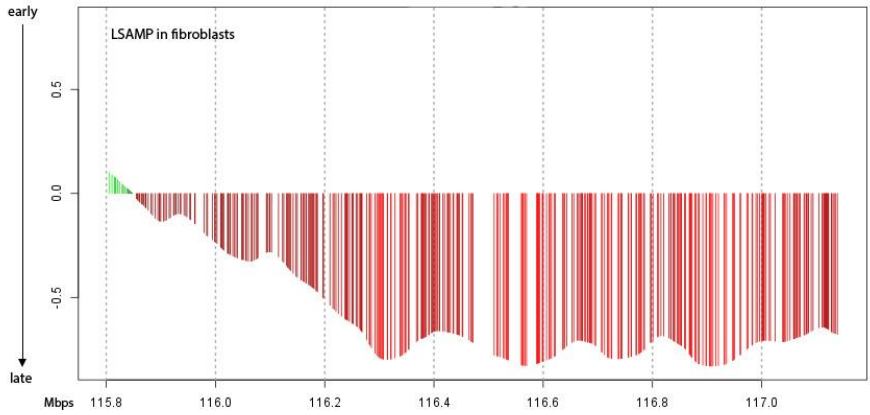


Figure 16. Replication timing analysis of LSAMP gene; region from 115,802,363 - 117,139,389 bp shows LSAMP gene replication timing analysis (from replicationdomain.org).

Since of its 1,33 Mb length, LSAMP gene could promote instability of 3q13.3 region; supportive to this hypothesis is replication timing analysis of the region. Since its late replication timing, and its high expression in lung tissue, this could explain the fragile region's high expression frequency by promoting the formation of R-loops as consequence of transcriptional and replicative machineries encounter (Helmrich et al., 2011), and by lacking activable replication origins to complete replication timing, which is a similar outcome hypothesized for NEGR1 gene on chromosome 1.
ARHGAP31 is a GTPase Activating Protein coding gene located in fragile site distal region; the encoded protein is a GAP shown to regulate two GTPases involved in protein trafficking and cell growth, required for cell spreading, polarized lamellipodia formation and cell migration. Chromosomal aberrations in this gene cause the Adams-Oliver Syndrome with abnormality in skin development and malformations of the limb among the primary features (Isrie et al., 2014). The gene dimensions could not promote the region genomic instability, but since its early replication timing (figure 17) and its high expression in lung tissue, encounters between replicative and transcriptional machineries could happen.

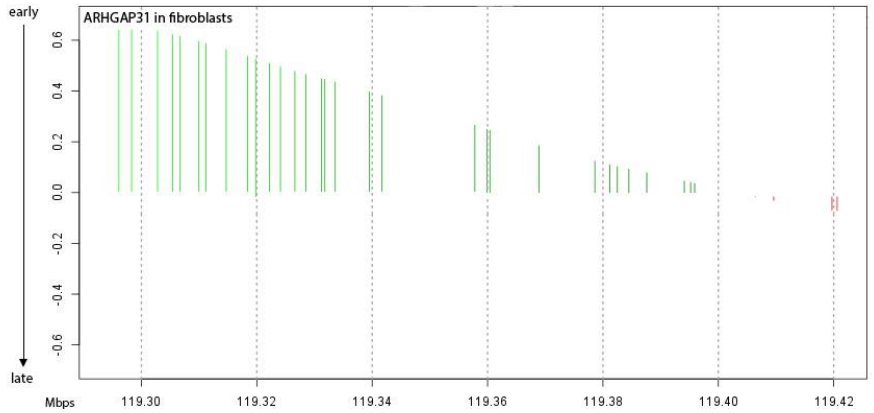


Figure 17. Replication timing analysis of ARHGAP31 gene (region 119,294,289-119,420,714 bp) show an early replication timing of the region that could promote fragile site expression (from replicationdomain.org).

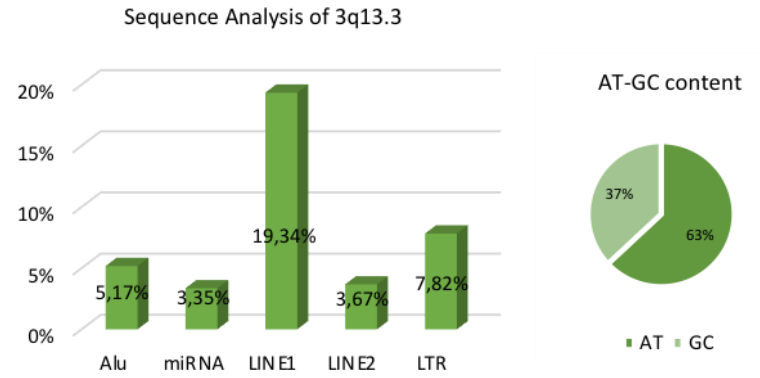
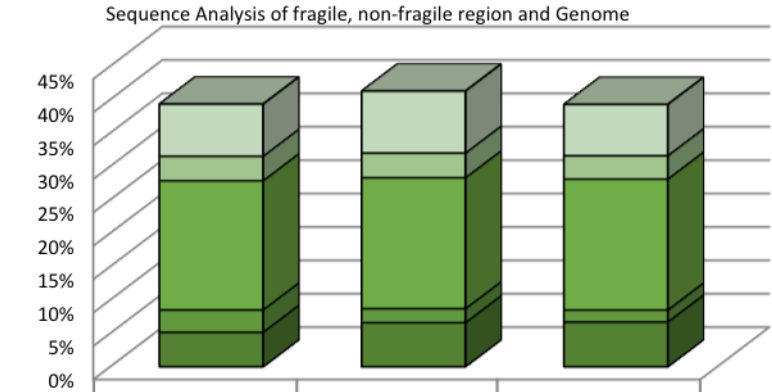


Figure 18. The histogram (left) shows the percentage of Alus, miRNAs, LINE1-2, LTR elements along the 3q13.3 region; on the right GC-AT content is shown

Next, there is the analysis of elements characterizing the region, with www.repeatmasker.org public database: AT content, LINEs, Alu, LTRs and miRNAs were analyzed.



	3q13.3	NFRA	Human Genome
LTR	7,82%	9,30%	7,70%
LINE 2	3,67%	3,70%	3,50%
LINE 1	19,34%	19,60%	19,60%
miRNA	3,35%	2,10%	1,80%

Figure 19. Sequence analysis and comparison between fragile region (3q13.3), non-fragile region (NFRA) and standard Genomic Sequence (Human Genome).

Again, the region is not enriched for these elements apart from LINE1 as shown in figure 18. The repetitive and regulative elements are not responsible for instability associated within this region.
The comparison between the fragile region, NFRA and standard genomic sequences, again showed no significant differences (figure 19) confirming the previous hypothesis.

Replication Timing analysis of Fragile Regions.

An important characteristic of CFSs is altered replication timing. To characterize putative CFSs' replication timing, analysis has been conducted on interphasic nuclei combining FISH experiments with immunofluorescence against BrdU to identify active replicating cells. FISH probes chosen by online genomic database screening, delimitate putative CFSs edges: one is proximal to the centromere and the other one is distal from the centromere as shown in figure 20 on chromosome 1 CFSs.

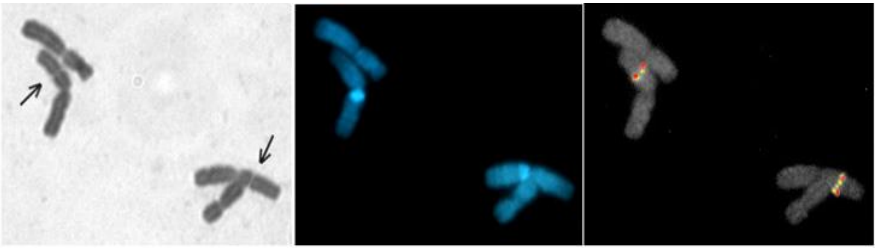


Figure 20. The figure shows 1p31.1 fragile region. On the left, break on short arm of chromosome 1 (black arrow) is identified with Giemsa staining; in the central figure the same chromosome is visualized with DAPI; on the right, distal (red spot) and proximal probes (yellow spot) are localized, in the same chromosome, through FISH experiment.

Five temporal S-phase stages, from early (I stage) to late (V stage), can be identified analyzing the pattern of BrdU incorporation in replicating cells (figure 21).

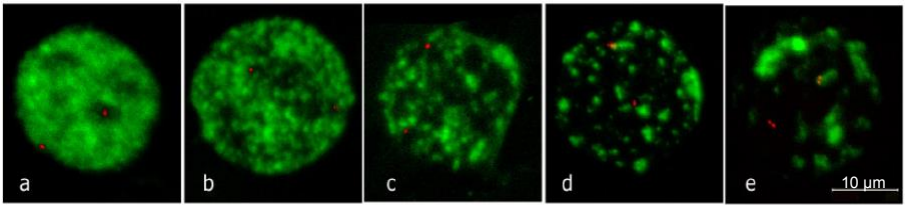


Figure 21. Five S-phase substages are shown with immunofluorescence anti-BrdU on interphasic nuclei, from the earliest (phase I, a) to the latest (V, e); red spots show FISH signals (Pelliccia et al., 2010).

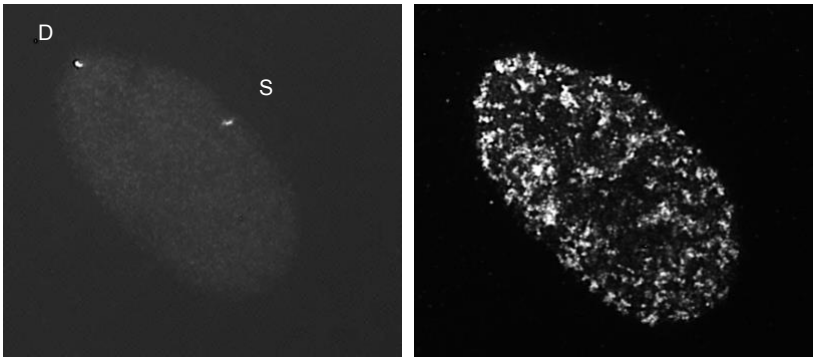
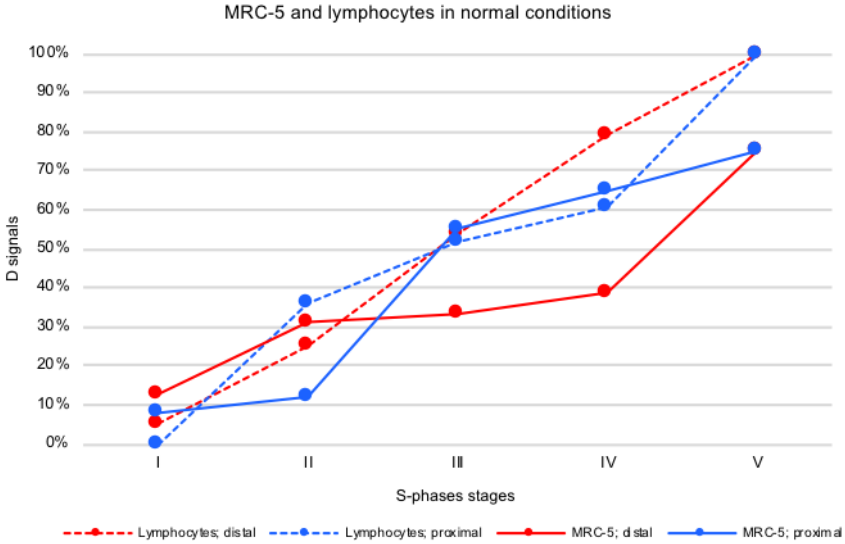


Figure 22. Double and Single spots on interphasic nuclei. Fluorescence *in situ* Hybridization's spots (left) are visible as double (D) and single (S) spots on asynchronous alleles; on the right, same nucleus is visualized with immunofluorescence against BrdU. A stage IV of S-phase is shown.

FISH was used to point out the replication timing of analyzed fragile regions; it makes possible the discrimination between non-replicated allele (S, single spot) from replicated one (D, double spot); the asynchronous alleles are visualized as double and single spot (DS) in each chromosome homologous (figure 22).
In each experimental protocol, two probes, specific for each fragile region, have been analyzed simultaneously with FISH experiments, the combination of FISH technique (S spot or D spot) with the BrdU-IF allows the recognition of the S phase in which each genomic region is replicating or not; for each probe at least 500 S-phase nuclei have been analyzed in each cell line and the signal obtained was associated with corresponding S-phase stage.
Replication timing pattern of CFSs in fibroblasts has been confronted, for the same region, in lymphocytes from peripheral blood, in which the same region is non-fragile. The data obtained from fibroblasts' fragile regions in lymphocytes were used for the replication timing analysis of both fibroblast tissues.
The experiments were conducted in duplicate with and without APH to analyze replication timing differences in normal and stressful condition in each putative CFS.

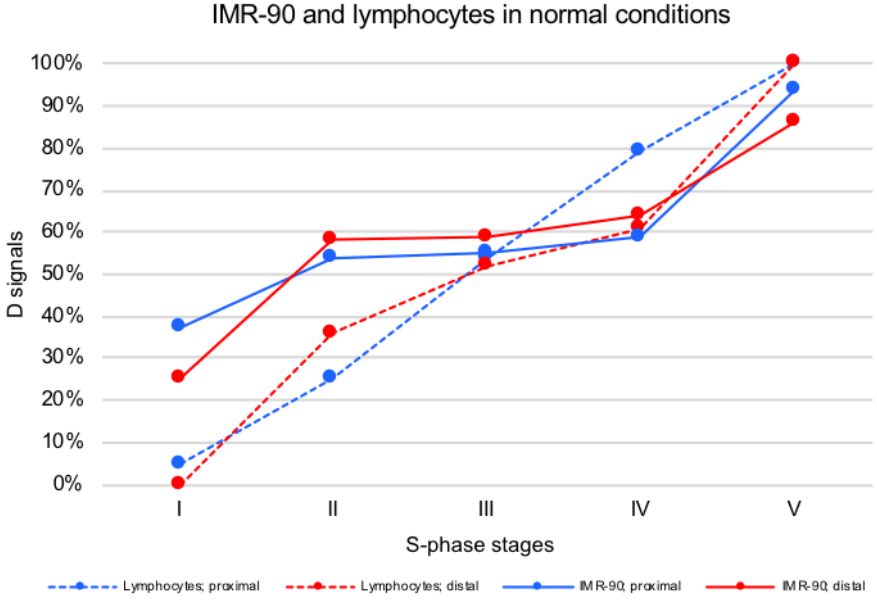
Replication Timing analysis of 1p31.1 Fragile Site.

- Normal conditions:



Starting with replication timing analysis of 1p31.3 fragile site, the two probes in MRC-5 fibroblasts and lymphocytes (proximal RP11-297N6; distal RP11-316C12) start early in phase I, with few differences in replicated alleles percentage (5% - 13% of D spots), apart from proximal probe in lymphocytes, in which no D spots can be found. Moving forward along replication timing the differences appear significative.
In lymphocytes, the replication follows a linear trend, with 52% of alleles already replicated in phase III in both regions until they arrive at the end of phase V with 100% of replicated alleles, indicating a normal replication timing, since they are non-fragile regions in lymphocytes.
In MRC-5, proximal region starts early in phase I (8% of D spot) and double alleles percentage rises until phase III in which 55% of alleles are already replicated, although replication slows down and arrives in phase V with 25% of non-replicated alleles. The replication trend in MRC-5 distal probes is even slower, it starts early in phase I (13% of

D spot), until a stall that lasts from phase II to phase IV is found, and replicated alleles do not increase (31%, 33% and 39% in phase II, III and IV respectively); the trend is confirmed from arrival in phase V with 75% of replicated alleles, the same percentage as proximal probe.



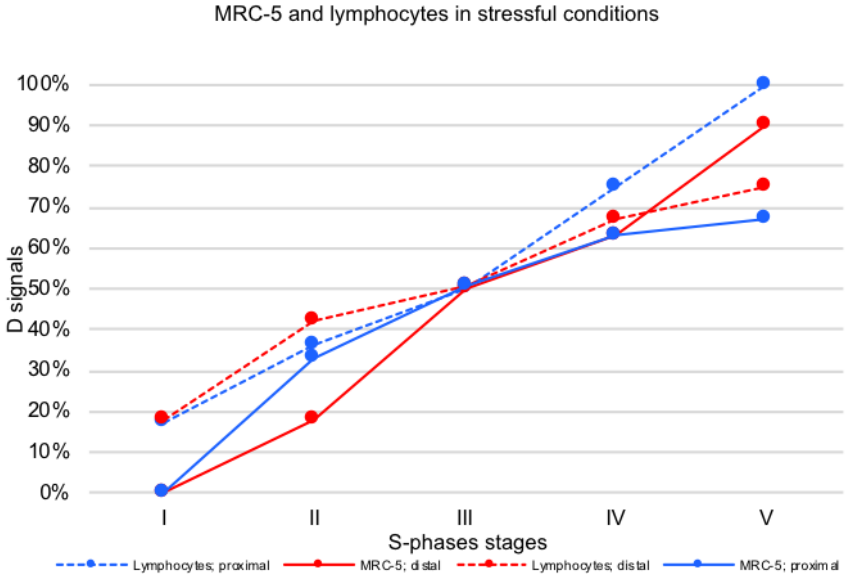
Analyzing the replication timing in IMR-90 fibroblast cell line in normal condition the trend is quite regular, apart differences in replicated alleles' timing. In fact, the IMR-90's probes start to be replicated earlier than lymphocytes' and MRC-5's ones; fibroblasts' proximal and distal probes have 38% and 25% of D-spots in phase I respectively, while the same in lymphocytes are later replicating, the distal one is not even replicated yet, while the proximal probe has 5% of D-spots in phase I.

Continuing the analysis, fibroblasts encounter a stall that lasts from phase II until phase IV and this stall could be explained with the fact that chromatin undergoes different conformational changes in order to let the replication proceed from earlier to later stages; they arrive at the end of the S-phase with almost all replicated alleles (94% form proximal and 86% for distal probes).

The remaining not replicated alleles are not expressed as breaks, since they could be repaired in G2 phase or even in mitosis.
Lymphocytes replicate all of their alleles at S-phase ending.

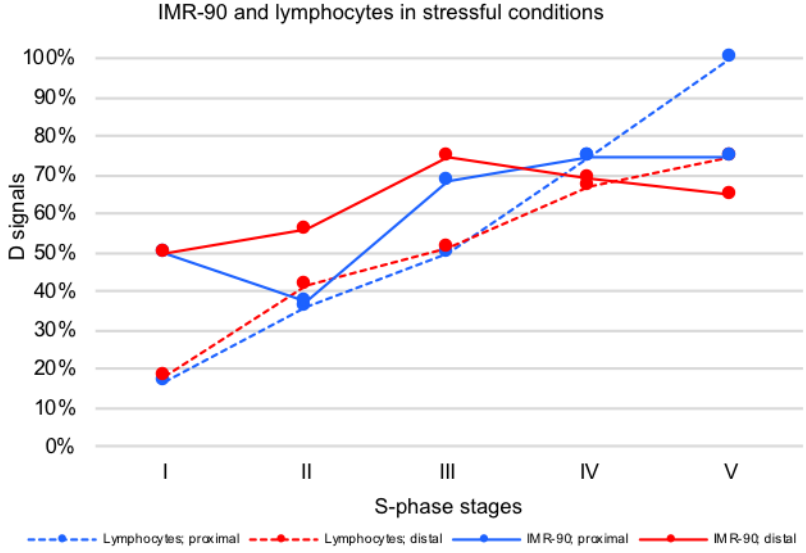
- Stressful conditions:

In stressful condition replication timing appears quite regular in both cell lines.



Probes in lymphocytes replicate early in phase I (18% of D spot) and the proximal probe arrives with totally replicated alleles; the distal probe shows a similar trend but it arrives in phase V with 25% of non-replicated alleles, indicating that this region, even if non-fragile in lymphocytes, shows replication's impairments.
In MRC-5, both regions start later than lymphocytes to be replicated, with no replicated alleles in phase I; an increase in replication is found only until phase III (51% of replicated alleles in both), then both proximal and distal regions arrive with few replicated alleles in phase V, with 67% and 75% of D spots respectively.
Analyzing the replication trend in both cell lines it is clear that there is an impediment in fork progression in phase III (55% of replicated

alleles in both cell lines in the two regions) probably given by structural characteristics, like nucleotide composition or sequences incline to form secondary structure, APH presence makes these characteristics more visible, indicating that this region is difficult to replicate even in tissue not expressing the fragile site.

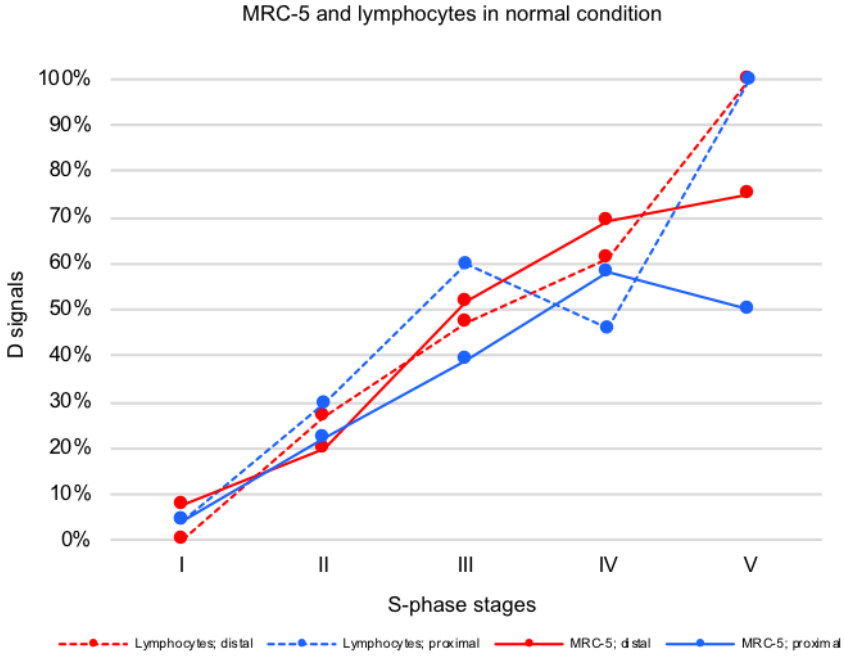


In stressful condition, replicative irregular trend reflects the region's fragile nature in both cell lines. Even in this case fragile regions replicate early, with 50% of alleles already replicated in phase I, even lymphocytes are early in starting the replication (17% of D-spots for proximal and 18% for distal). The trend is quite linear, the D-signals increase in both regions and in proximal reach 100%, while the distal one is more problematic arriving in phase V with 25% of non-replicated alleles.

In fibroblasts, distal probe increases replication timing but after phase III (75% of D-spots) replication's speed decreases arriving in phase V with 35% of alleles non-replicated; the proximal probe has a similar outcome, it decreases in phase II (38%) after which there is a sharp increase and a stasis from phase III (68% D-spots) until phase IV (75% D-spots) in which replicated alleles do not reache 100% (75% of D-spots in phase V). The stasis observed in the last phases could be explained with the fact that all replication origins have already been activated in earlier phases.

Replication timing analysis of 3q13.3 fragile site.

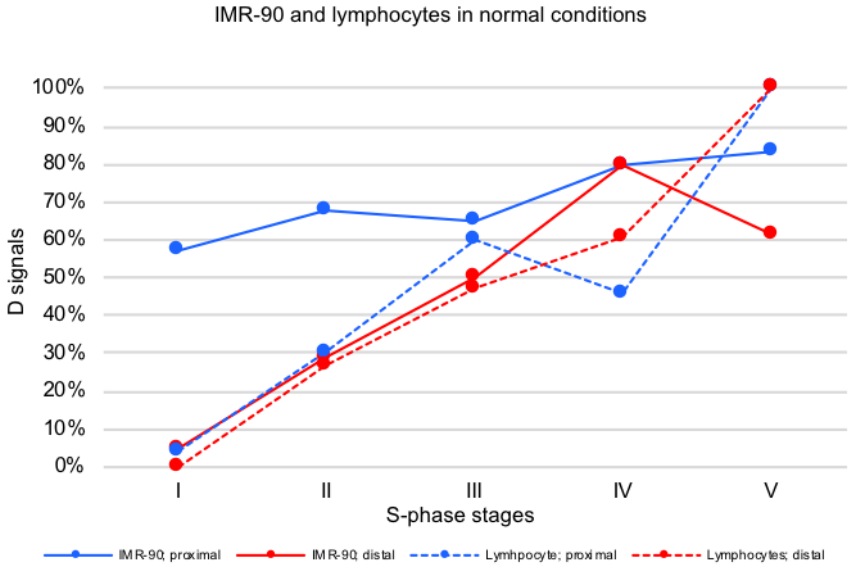
- Normal conditions



Replication timing of 3q13.3 in normal condition shows an almost linear trend. Both regions (proximal RP11-324H4 and distal RP11-305I9) in both cell lines start replication in phase I with 4% of replicated alleles in MRC-5 and lymphocytes' proximal regions, and 8% of D spots for MRC-5 distal region, apart from lymphocytes' distal probe (no D spots detectable in phase I). From phase I onward, replicated alleles increase until phase IV for both distal probes and MRC-5 proximal (58% and 61% for the first two, 69% for the latter); lymphocytes' proximal probe shows an irregular trend from phase III onward, but it ends the replication with 100% of replicated alleles, as in the distal one.

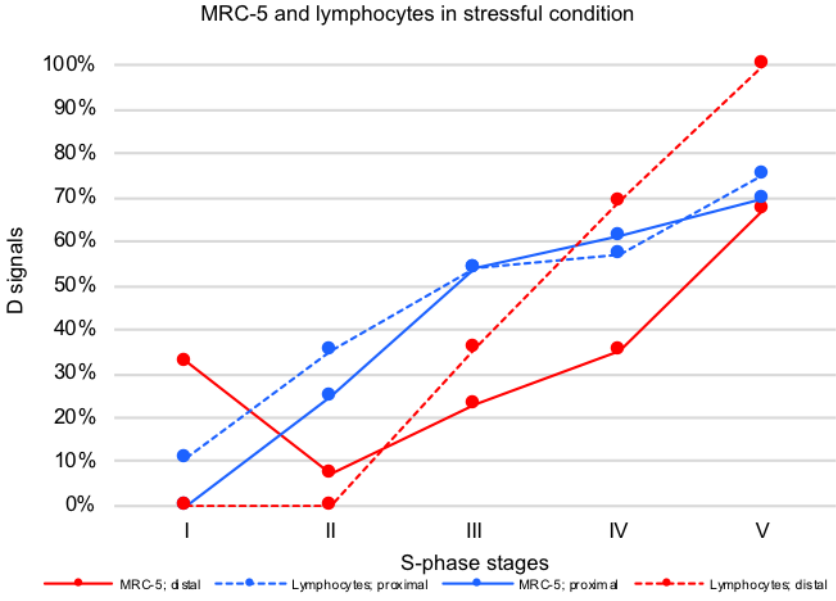
Both MRC-5 regions arrive in phase V with non-replicated alleles, 25% for the distal probe, and a 50% of alleles to be replicated in proximal one. MRC-5's trend reflects the nature of FSs fragility, in

fact they represent regions with structural peculiarities, even in non-stressful condition, as shown for MRC-5's replication.

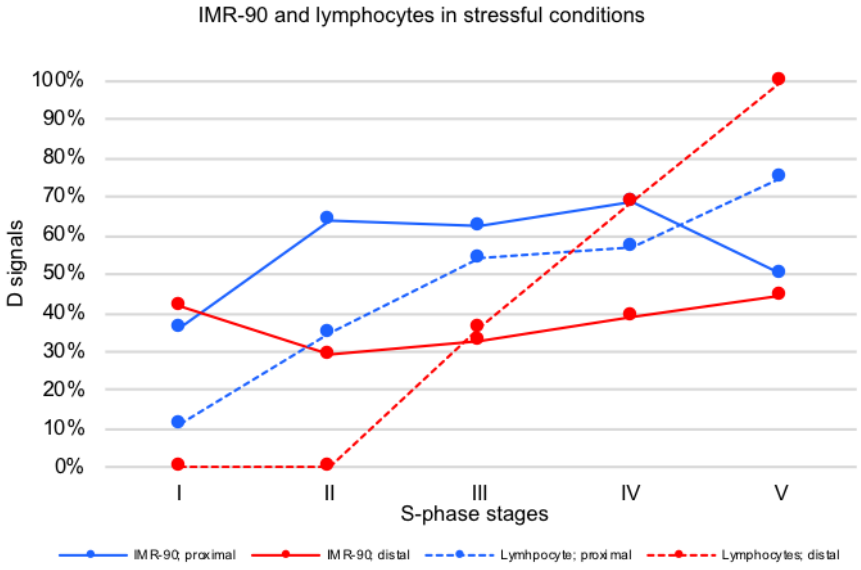


Replication timing of 3q13.3 fragile site is quite linear in IMR-90, apart from proximal probe, which is early replicating (57% of D-spots in phase I); the other regions proceed as a block in a linear trend mostly: they start from around 0-4% and increase linearly until phase III. From phase III onwards IMR-90 increase until phase IV (80% of D-spots) afterwards a decrease in phase V can be observed with only 61% of replicated alleles. The other fibroblasts' region replicates early with a regular trend which remains regular, ending with 83% of replicated alleles.
The lymphocytes' regions arrive in phase V with all of their alleles replicated.

- Stressful conditions



Replication trend in stressful condition is much more irregular.
The two proximal probes show a similar trend from phase I until phase III where a stall in replication can be visualized (54% of replicated alleles).
Distal probes show the most irregular trend. MRC-5's probe starts quite early in phase I with already 30% of D signals followed by a sharp decrease in phase II and a rising until phase V. The lymphocytes' one instead is a late replicating region in this tissue: it starts to be replicated only after phase II but it arrives at the end of S phase with 100% of replicated alleles.
Three probes out of four arrive in phase V with non-replicated alleles: lymphocytes' and MRC-5 proximal with 75% and 70% of replicated alleles, MRC-5 distal with 67% of D spots. Only lymphocytes' distal probe arrives at phase V with 100% of replicated alleles.
The typical trend showed by all regions, and in lymphocytes in particular in which these regions should be non-fragile, is concordant with previous FS's analysis results: these regions show difficult replication even in other tissue, suggesting structural impediment for efficient replicative fork progression.



This replication is quite peculiar in every region for both cell lines. Starting with fibroblasts, both regions are quite early in replication starting around 40% of replicated alleles in phase I, in phase II distal one decreases (30% of D-spots), after which the trend is linear until of S-phase's termination, in which non-replicated alleles are over 56%; proximal one, after a slight increase in phase II (64% of D-spots), encounters a stall which lasts until phase IV and the replicated alleles do not increase significantly (69%). After that, alleles decrease, and they end S-phase with over 50% of non-replicated alleles. The stasis observed in II-IV phases could again be explained with the fact that all replication origins have already been activated in previous phases, leaving regions not completely replicated in last phases.
The lymphocytes' replication is quite peculiar as shown in previous analysis for MRC-5.
The tables summarize the results obtained from replication timing analysis in all cell lines in both conditions for both analyzed fragile regions.

1p31.1 normal condition.

Lymphocytes	MRC-5	IMR-90

Canonical replication 100% in V	*Distal*: early but slow in II – IV, 75% in V *Proximal*: early and fast in II – III Slow in III – V (75%)	*Distal*: early but slow in II – IV (74%); 86% in V *Proximal*: early but slow in II – IV (65%); 94% in V

1p31.1 stressful condition.

Lymphocytes	**MRC-5**	**IMR-90**
Early start compared to normal condition *Distal*: almost regular but 75% in V *Proximal*: almost canonical 100% in V	Late start *Distal*: almost regular until V (75%) *Proximal*: canonical start until III; stasis until V (67%)	Very early (50%) *Distal*: slow progression in III (75%); decrease in V (65%) *Proximal*: decrease in II, fast progression in III, slows in V (75%)

3q13.3 normal condition.

Lymphocytes	**MRC-5**	**IMR-90**
Almost linear progression 100% in V	Linear progression until IV *Distal*: regular until IV (65%) but 75% in V *Proximal*: regular until IV (58%) but decrease in V (50%)	Peculiar trend for proximal *Distal*: regular until IV (80%) decrease in V (61%) *Proximal*: very early (57%); 89% in V

3q13.3 stressful conditions.

Lymphocytes	**MRC-5**	**IMR-90**
Early start compared to normal condition *Distal*: problematic beginning but 100% in V *Proximal*: 10% in I; regular trend and stasis in III - IV	Late start *Distal*: early start (30%), decrease in II and restart in V (70%) *Proximal*: late start and regular until III (54%) slows and 70% in V	Very early (50%) *Distal*: early start (40%) drop in II and 44% in V *Proximal*: early start (40%); "problematic" trend in V (50%)

Genomic Instability at 1p31.1 and 3q13.3 fragile regions.

In the last part of my work I searched, using online public databases, if both fragile regions were exposed to chromosomal instability in different type of tumors and diseases. Mitelman Database (https://cgap.nci.nih.gov/Chromosomes/Mitelman) of Chromosome Aberrations and Gene Fusions in Cancer relates chromosomal aberrations to tumor characteristics, based either on individual cases or associations, and has been used on 1p31.1 and 3q13.3 fragile regions; the results are shown in tables 1 and 2.
As data show (table 1 and 2), both fragile regions are under severe chromosomal instability that causes different rearrangements in both solid and hematopoietic tumors in different cell types.
In 3q13.3 region several genes are localized and two of them LSAMP and ARHGAP31 map at the boundaries of core region. Both of them are transcribed in both analyzed fibroblast cell lines and LSAMP is even a large gene showing a late replication timing, typical of large genes located in fragile regions; their transcription could explain the high frequency of expression in 3q13.3 region.
The most important genes located in this region are even under gene mutations involved in different kind of psychiatric diseases and many types of tumors, since most of them are tumor suppressor genes such TUSC7, LSAMP and IGSF11. The same can be said for NEGR1 large gene on chromosome 1; mutations in this gene are involved in psychiatric disorders and poor outcomes in neuroblastoma.
All these evidences seem to confirm the hypothesis that large genes localized within fragile regions are involved in chromosomal instability because most of them are transcribing, late replicating genes, most probably localized in regions with poor dormant origins unable to be activated upon stressful agents, moreover the genes located in these fragile regions are probably subjected to mutations themselves because of their localization in structural fragile regions.
The genes and the structural characteristics found in these fragile regions, will need further investigations to better understand the connection between replication, transcription and structural organization of the chromatin in these regions of instability.

Unbalanced chromosomal aberrations in solid tumors

Unbalanced chromosomal aberrations in hematopoietic tumors

Abnormality	Morphology
add(1)(p31)	Adenocarcinoma
add(1)(p31)	Neuroblastoma
add(1)(p31)	Squamous cell carcinoma
add(1)(p31)	Teratoma
del(1)(p12p31)	Mesothelioma
del(1)(p31)	Astrocytoma, grade III-IV/ Glioblastoma
del(1)(p31)	Ewing sarcoma
del(1)(p31)	Malignant melanoma
del(1)(p31)	Meningioma
del(1)(p31)	Neuroblastoma
del(1)(p31)	Squamous cell carcinoma
del(1)(p31)	Teratoma
der(1)del(1)(p31)del(1)(q21)	Adenocarcinoma

Abnormality	Morphology
add(1)(p31)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
add(1)(p31)	Diffuse large B-cell lymphoma
add(1)(p31)	Follicular lymphoma
add(1)(p31)	Multiple myeloma
del(1)(p11p31)	Mantle cell lymphoma
del(1)(p11p31)	Multiple myeloma
del(1)(p12p31)	Mantle cell lymphoma
del(1)(p12p31)	Multiple myeloma
del(1)(p13p31)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
del(1)(p13p31)	Mantle cell lymphoma
del(1)(p13p31)	Multiple myeloma
del(1)(p13p31)	Plasma cell leukemia
del(1)(p21p31)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
del(1)(p21p31)	Mantle cell lymphoma
del(1)(p21p31)	Multiple myeloma
del(1)(p22p31)	Mantle cell lymphoma
del(1)(p22p31)	Multiple myeloma
del(1)(p31)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
del(1)(p31)	Acute myeloblastic leukemia with maturation (FAB type M2)
del(1)(p31)	Multiple myeloma
del(1)(p31p36)	Multiple myeloma
der(11)t(1;11)(p31;q23)	Multiple myeloma

Balanced chromosomal aberrations in hematopoietic tumors

Abnormality	Morphology
t(1;3)(p31;q21)	Multiple myeloma

Table 1. The tables show the chromosomal rearrangements in 1p31.1 fragile region involved in different type of solid and hematopoietic tumors (adaptation from Mitelman database; https://cgap.nci.nih.gov/Chromosomes/Mitelman).

Unbalanced chromosomal aberrations in solid tumors

Abnormality	Morphology
add(3)(q13)	Adenocarcinoma
add(3)(q13)	Adenocarcinoma
add(3)(q13)	Adenocarcinoma
del(3)(q13)	Adenocarcinoma
del(3)(q13)	Adenocarcinoma
del(3)(q13)	Adenocarcinoma
del(3)(q13)	Adenocarcinoma
del(3)(q13)	Adenocarcinoma
del(3)(q13)	Basal cell carcinoma
del(3)(q13)	Mesothelioma
der(17)t(3;17)(q13;p11)	Mantle cell lymphoma
der(2)t(2;3)(q33;q13)	Adenocarcinoma
der(4)t(3;4)(q13;p15)	Adenocarcinoma
der(5)t(3;5)(q13;p13)	Adenocarcinoma
der(6)t(3;6)(q13;q16)	Adenocarcinoma

Unbalanced chromosomal aberrations in solid tumors

Abnormality	Morphology
add(3)(q13)	Diffuse large B-cell lymphoma
add(3)(q13)	Follicular lymphoma
del(3)(q13)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
del(3)(q13)	Adult T-cell lymphoma/leukemia (HTLV-1+)
del(3)(q13)	Chronic myeloid leukemia, t(9;22)
del(3)(q13)	Diffuse large B-cell lymphoma
del(3)(q13)	Multiple myeloma
del(3)(q13)	Myelodysplastic syndrome, NOS
del(3)(q13q13)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
del(3)(q13q13)	Acute lymphoblastic leukemia/lymphoblastic lymphoma
del(3)(q13q26)	Follicular lymphoma
del(3)(q13q27)	Acute myeloid leukemia, NOS
der(17)t(3;17)(q13;p11)	Mantle cell lymphoma

Balanced chromosomal aberrations in solid tumors

Abnormality	Morphology	Genes
inv(3)(q13q27)	Diffuse large B-cell lymphoma	BCL6+

Table 2. The tables show the chromosomal rearrangements in 3q13.3 fragile region involved in different type of solid and hematopoietic tumors (adaptation from Mitelman database; https://cgap.nci.nih.gov/Chromosomes/Mitelman).

CONCLUSIONS

In this work different aspects of CFSs were investigated.
The tissue specificity associated with CFSs was examined and confirmed: 1p31.1 and 3q13.3 are CFSs expressed in MRC-5 and IMR-90 fibroblast cell lines. The two fragile regions are among the most expressed breaks in both fibroblast cell lines and cannot be found in lymphocytes.
Their molecular characterization was performed to confirm the nature of these breakages as CFS, they both have a percentage of expression higher than 3% and high AT percentage, typical characteristics found in CFSs.

Sequence analysis shows the presence of many genes in both fragile regions and some of them are large genes, NEGR1 (886 kb) is located in the core region of 1p31.1, while LSAMP (1,33 Mb) is found at the boundary of 3q13.3 core region. It is known that large genes are late replicating and could promote chromosomal instability on CFS by R-loops formation; using online public database we found a correspondence in both fragile regions between large gene and late replication timing, in fact they both replicate late, and this evidence could explain the high frequency of expression observed in both CFSs. Moreover, in 3q13.3 region ARHGAP31 gene (1,26 kb) is an early replicating gene and could promote the fragility by causing replicative and transcriptional machineries collision.

Using online databases, characterization of repetitive and regulative elements was performed to search the causes of these CFSs instability; these elements are not different in percentage to those in non-fragile regions and in the rest of the genome, so they can be excluded, so far, from promoting fragility in 1p31.1 and 3q13.3.

Knowing that non-canonical replication timing is among the main causes for CFSs instability within genome, their replication analysis was performed as well.
The results show a typical trend of fragile regions, with alleles not completely replicated at the end of S-phase in both normal and

stressful conditions, moreover these regions, in a non-fragile background, displayed a non-canonical replication as showed from replication analysis in lymphocytes, suggesting the presence of structural peculiarities that impair replicative fork progression since these regions do not complete the replication of their alleles at the end of S-phase.
These results suggest a prominent role of replication in promoting fragility in these regions. Further analysis will be performed to confirm the role of replication on both fragile regions.

Finally, Mitelman database was used on 1p31.1 and 3q13.3 fragile regions to investigate if these regions were under chromosomal instability. The results showed a strong chromosomal instability in both regions analyzed, involved in solid and hematopoietic tumorigenesis.
Another evidence of the connection between CFSs and genomic instability is that most of the genes localized in both regions are mutated in different diseases that go from tumors in different cell types, since most of them are tumor-suppressor genes, to psychiatric disorders.

All these results suggest that both fragile regions are under a strong chromosomal instability and are involved in genesis of different type of disorders, further experiments will shed a new light on CFSs involvement in promoting genomic instability and diseases.

MATERIALS AND METHODS

Cell lines.

For this work three cell lines have been used: two fibroblast cell lines, IMR-90 and MRC-5, and lymphocytes from peripheral blood.

- *Fibroblast IMR-90*: from the lung epithelium of a 16-week female fetus, with normal karyotype. They are not immortalized; these cells can undergo trough 58 population doublings before entering replicative senescence. Their dividing potential, the virus susceptibility has extensively been studied, in fact they can be considered as an alternative to the WI-38 cell line.
- *Fibroblast MRC-5*: similar to the former cell line; derive from the lung epithelium of 14 weeks old male fetus, with a normal karyotype; the cells are capable of 42 to 46 population doublings before the senescence onset.
- *Peripheral blood lymphocytes*: at least 5 mL of blood from an adult normal individual are cultured with sodic heparin to avoid blood clotting.

Lymphocyte cell culture.

The lymphocytes have been obtained sampling 0.3 mL of peripheral blood of a male human healthy individual with a heparin prefilled syringe. Afterward 5 mL of specific suspension cell culture medium (RPMI 1640, Corning) supplemented with 10% of Fetal Bovine Serum (Gibco) and 1% of L - glutamine (Sigma) are added. Addition of 0.15 mL of phytohemagglutinin (PHA M - form, Gibco), allows T-cells activation. The cells are incubated at 37°C for 72 hours.
In the last 24 hours, 0.4 μM of aphidicolin (Sigma-Aldrich) are added to allows the expression of fragile sites in the last part of the S phase. Afterward, the addition of 10^{-4} M of colchicine (Sigma-Aldrich) for 2 hours, allows the depolymerization of microtubules and mitotic spindle disassembly, hence the visualization of the metaphase chromosomes is accomplished.
In order to study the replication timing, 10 μM bromodeoxyuridine are added (Sigma-Aldrich) in the last 20 minutes.

Fibroblasts cell culture.

The cells, that come from N_2 vapor phases, are rapidly thawed in a 37°C water bath for a maximum of 2 minutes and then 5 mL of 10% FBS - 1% L-glutamine medium (MEM 1X, Gibco) are gradually added. Finally, the cells are incubated at 37° C with 5% CO_2.
In the last 20 - 22 hours, 0.4 μM of aphidicolin are added for fragile sites induction; 10 μM of bromodeoxyuridine are added in the last 20 minutes for replication timing analysis studies. For metaphases spreads 10^{-4} M of colchicine (Sigma-Aldrich) is added for 4 hours.

Cytological samples from lymphocytes culture.

To harvest the cells, the tubes are centrifuged at 2000 rpm for 4 minutes; the addition of 0.0075 M of KCl hypotonic solution, for 7 minutes, swells the nuclei.
After a second centrifugation, Ibraimov solution is added (5% acetic acid - 3% methanol in distilled H_2O) as a pre-fixation before the third centrifugation, they are followed by two subsequent washes with – 20°C cold fixative (methanol – acetic acid 3:1) in order to dehydrate the chromosomes on the slide.
The last centrifugation removes the fixative, the pellet is air dried before the samples' preparation made by releasing few drops on an ice-cold glass slide.

Cytological samples from fibroblasts culture.

a) *cytocentrifuges samples preparation.*
The medium is gently removed from the plate, the hypotonic buffer addition (10 mM Tris-HCl [pH 7.4], 40 mM glycerol, 20 mM NaCl, 1.0 mM $CaCl_2$, 0.5 mM $MgCl_2$) for 15 minutes at 37°C allows detaching of mitotic cells; the cells are collected by pipetting hypotonic buffer several times.
The buffer is centrifuged at 1800 rpm for 8 minutes and fixed in -20°C cold fixative (methanol – acetic acid 3:1) for 20 minutes.

b) *Air drying samples preparation.*
The addition of 0.4 μM aphidicolin in the last 22 hours allows the induction of CFSs.

The medium is removed, after two subsequent washes in PBS (Phosphate Buffered Saline, Corning), the trypsin addition (0.05% trypsin, 0.53 mM EDTA, Gibco) for 8 minutes at 37°C, allows fibroblasts cells detachment.
The cells collected with 10% FBS - 1% L-glutamine (MEM 1X, Gibco) complete medium are centrifuged for 5 minutes at 1200 rpm. The nuclei are swelled with 0.075 M KCl hypotonic solution for 6 minutes before a 1200 rpm centrifuge for 5 minutes to remove solution. Two consequent 4 minutes centrifuges in -20°C cold fixative (methanol – acetic acid 3:1) at 1200 rpm leaves the pellet clean. The tubes are sets overnight at -20°C before slides preparation.
The cells are pelleted by 1200 rpm 5 minutes centrifuge and afterwards dried under a fume hood for 15 minutes before adding few drops of -20°C cold fixative.
Few drips of cells suspension are dropped on cold glass slides; the quality of the slides can be observed by phase-contrast microscopy.

R-banding with Chromomycin A3.

To discriminate chromosome and to identify the exact chromosomic region R-banding technique has been used.

The solution used for R-banding are:

- Buffer solution containing phosphate NaH_2PO_4 and Na_2HPO_4 0.14 M; pH 6.8.
- NaCl – Hepes Buffer 0.15 M (pH 7.0), 0.005 M Hepes.

Slides are treated for 10 minutes with phosphate buffer solution before Chromomycin A3 addition and leaved for 2 hours in humid chamber at room temperature.
The slides are briefly rinsed with NaCl – Hepes buffer and treated with methyl green solution for 15 minutes with no light exposure at room temperature; following two washes in NaCl-Hepes buffer, antifading solution is added (antifading:isopropilgallate, 1:300). The slides are stored 2-3 days at dark at 4°C before observation.

Fluorescent *in situ* Hybridization (FISH).

Probes selection
To circumscribe the fragile regions region, after a screening on www.genomedataviewer.com we chose a set of probes for each break: the *proximal* one, located between the centromere and the fragile region, and *distal* one between the fragile region and the telomere.

	1p31.3	3q13.3
Proximal	RP11-297N6	RP11-324H4
Distal	RP11- 316C12	RP11-305I9

Table 3 - Proximal and distal probes used for each putative CFS.

Bacterial cells culture and DNA extraction
The *E. coli* cells containing vectors are cultured for 16-20 hours in 50 mL tubes in 2x YT medium, (3.5% tryptone, 2.0% yeast extract, 0.5% NaCl), containing either 20 μg/mL chloramphenicol or 50 μg/ml kanamycin for BACs and PACs selection respectively[1].

The tubes are centrifuged at 4000 rpm for 7 minutes, the resuspension of pellet by GTE addition (50 mM Glucose, 25 mM Tris pH 8, 10 Mm EDTA).
The denaturation solution freshly made (0.2 N NaOH, 1% SDS), allows the lysis of bacterial cells before a centrifugation for 10 minutes at 14000 rpm; the supernatant containing BAC and genomic DNA is poured in new tubes. The addition of ammonium acetate 7.5 M and the subsequent centrifugation at 13000 rpm for 20 minutes, helps the precipitation of supercoiled genomic DNA and proteins, while the BAC, lighter, remains in solution. The supernatant containing BAC is then transferred in new tubes and the isopropanol addition promotes the BAC precipitation; after another centrifugation (14000 rpm for 20 minutes) the supernatant is discarded, the pellet is suspended in 70% ethanol to remove salts, and another centrifugation at 14000 for 5

[1] BACs and PACs have been kindly provided by Professor Mariano Rocchi (University of Bari).

minutes allows the removal of supernatant. The pellet containing the BAC DNA is suspended in TE buffer (10 mM Tris - HCl pH 8.1, 1 mM EDTA pH 8.0) containing RNase (final concentration 100 μg / mL) and later incubated for 30 minutes at 37°C.
The addition of 1/10 volume of Sodium Acetate and 3 volumes of Ethanol helps pelleting the BAC DNA, after 20 minutes of incubation at -20°C. A last centrifugation at 14000 rpm for 15 minutes removes RNA and salts; the pellet is suspended in an appropriate TE buffer volume, checked on 1% agarose gel and stored at 4°C.

Probes labeling through Nick Translation
Nucleotides are modified with biotin or digoxigenin introduction based on the revelation methods; indirect labeling methods and biotin/streptavidin or digoxigenin/alfa - digoxigenin systems are used in signal revelation.

50 μL reaction solution contains:

- 1 μg of DNA
- DNA polymerase I buffer reaction (0.5 M Tris-HCl [pH 7.8], 50 mM $MgCl_2$, 0.5 mM Bovine Serum Albumin (BSA));
- dATP, dCTP, dGTP non- labeled mixture (0.5 M);
- bio-16-dUTP or dig-16-dUTP labeled dTTP or labeled dTTP (0.5 M);
- β-mercaptoethanol (0.01 M);
- DNase I and DNA polymerase I (1:500);
- Ultrapure distilled H_2O in variable amount.

The solution is left at 15°C for 2 hours before checking DNA fragments on agarose gel. The reaction is blocked by EDTA 0.5 M addition (final concentration 10 μM).

Slides pretreatment
Slides are incubated in humid chamber at 37°C with RNase (100 μg/ml in 20x SSC buffer) for 1 hour to allow RNA degradation.
Slides are dehydrated with 70%, 90% and 100% ethanol washes for 5 minutes each and air dried. Afterward the slides are aged for 1 hour at 65°C and then moved at 80°C for exactly 2 minutes to promote DNA denaturation; the process is eased by 70% formamide solution in 20x

SSC buffer addition. The reaction is blocked with 70% -20°C cold ethanol wash; following dehydration with 90% and 100% ethanol, the slides are air dried before hybridization.

Probe pretreatment
200 ng of DNA are precipitated with a solution containing Cot-1 and Herring sperm DNA. 1/10 of sodium acetate 3M is added with 3 volumes of -20°C 100% ethanol and the probe incubated for 1 hour at -80°C or overnight at -20°C. The probe is centrifuged at 13000 rpm for 15 minutes, and 70% ethanol in equal quantity to the supernatant removed is added. Following 10 minutes centrifugation at 13000 rpm the pellet is air dried and suspended in hybridization solution containing 50% deionized formamide and dextran sulfate 10% in 20x SSC buffer.
The probe is denaturized at 80°C for 8 minutes, transferred on ice and incubated at 37°C for 15 minutes to allows the competitor DNA association; the probe rests on ice until hybridization.

In Situ hybridization
The slides are moved to 37°C and hybridization solution containing probe is added; slides are closed with rubber cement and leaved in humid chamber at 37°C for 18 hours minimum in order to let the hybridization between the target DNA on the slides and the probe.

Post-hybridization washes
After 18 hours the slides are washed 3 times in 1x SSC buffer at 60°C for 5 minutes each, the slides are moved in humid chamber at 37°C and treated with 3% BSA blocking solution in 0,1% Tween 20 and 4x SSC saline buffer.
Slides are incubated for 30 minutes at 37°C with 1% BSA, 0,1% Tween 20 in 1x SSC buffer along with antibodies specific for probes. Based on probes labeling the antibodies are chosen among anti-digoxigenin-rhodamine (1:300, Roche), FITC-anti-digoxigenin (1:20), Cy3-streptoavidin (1:300) or FITC- avidin (1:20). The FITC and Cy3 fluorophores anti-digoxigenin and avidin conjugated respectively allows the probes signal recognition on samples.
3 washes in 4x SSC and 0.1% Tween 20 at 42°C are made, the slides are closed with 1:300 DAPI: antifading solution (Vectashield).

The slides can be observed after 1 hour at 4°C.

Immunofluorescence anti-BrdU.

For replicative state of nuclei detection, after FISH experiments, the slides are treated with immunofluorescence against bromodeoxyuridine (BrdU). Thymine analogue BrdU, is incorporated only in active replication nuclei in S-phase and can be revealed with antibodies anti-BrdU.
The FISH treated slides are washed once with 4x SSC containing 0.1% Tween 20 before adding the primary antibody solution anti-bromodeoxyuridine (1:1000, Thermo-Fisher) in PBS 1x and 5% FBS buffer. After antibody incubation, slides are washed 3 times in 1x PBS, secondary antibody solution is added containing FITC anti - mouse IgG (1:100, Thermo-Fisher) in PBS 1x buffer and set in humid chamber for 1-hour minimum prior washing. The slide are washed 4 times in 1x PBS and closed with DAPI:antifading solution, 1:300 (Vectashield).
The slides set overnight at 4°C prior observation.

Microscope slides Observation.

Slides observation is done through fluorescence microscope connected to a CCD camera.
The chosen fluorophores, and the respective wavelengths are the following:

Fluorophores	**Excitation λ_{max}**	**Emission λ_{max}**
DAPI	365 nm	397 nm
FITC	494 nm	523 nm
Cy3	552 nm	565 nm
Rhodamine B	553 nm	627 nm
BrdU	350 nm	461 nm

The differences in the three excitation and emission of fluorophores wavelengths grants the observation of three probes at the same time: biotin for the first one, digoxygenin for the second and biotin/digoxygenin for the third one, the probes will appear as double spot signal on sister chromatids.

Photos are taken by RSImage software with three different filters, merged together and edited through Adobe Photoshop for probes' visualization and position on sister chromatid and interphasic nuclei.

Sequence analysis of Fragile regions.

BACs used in FISH experiment were chosen using Genome Data viewer online database (https://www.ncbi.nlm.nih.gov/genome/gdv/). AT percentage and repetitive elements composition were performed using Repeat Masker program (http://www.repeatmasker.org/), a program that screens DNA sequences for interspersed repeats and low complexity DNA sequences.
The replication timing analysis was performed using Replication domain database (https://www2.replicationdomain.com/), an online database resource for storing, sharing and visualizing DNA replication timing and transcription data, as well as other numerical epigenetic data types. Data is typically obtained from DNA microarrays or DNA sequencing.

Consulted Databases.

http://genome.ucsc.edu/cgi-bin/hgGateway
http://www.genecards.org/
http://genome-euro.ucsc.edu/index.html
http://www.ncbi.nlm.nih.gov/mapview/
http://www.repeatmasker.org/
http://replicationdomain.org/
https://www.ncbi.nlm.nih.gov/genome/gdv/
https://cgap.nci.nih.gov/Chromosomes/Mitelman

REFERENCES

Agnel Sfeir, Settapong T. Kosiyatrakul, Dirk Hockemeyer, Sheila L. MacRae, Jan Karlseder, Carl L. Schildkraut, and Titia de Lange, (2010). Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell*. 138(1): 90–103.

Arlt M. F. e Glover T. W., (2010). Inhibition of topoisomerase I prevents chromosome breakage at common fragile sites. *DNA Repair* (Amst). 9(6):678-89.

Arlt M.F., Xu B., Durkin S.G., Casper A.M., Kastan M.B., Glover T.W (2004). BRCA1 is required for common-fragile-site stability via its G2/M checkpoint function. *Mol. Cell. Biol*. 24(15):6701-9.

Arlt MF, Miller DE, Beer DG, Glover TW (2002). Molecular characterization of FRAXB and comparative common fragile site instability in cancer cells. *Genes Chromosomes Cancer*. 33(1):82-92.

Barlow J. H., et al., (2013). Identification of early replicating fragile sites tha contribute to genome instability. *Cell.* 152:1-13.

Barnes Ryan P., Hile Suzanne E., Lee Marietta Y., Eckert Kristin A. (2017). DNA polymerase seta and kappa exchange with the polymerase delta holoenzyme to complete common fragile synthesis. *DNA repair* 57: 1-11.

Branzei D., Foiani M. (2010). Maintaining genome stability at the replication fork. *Nat. Rev. Mol. Cell Biol.* 11(3):208-19.

Byun T.S., Pacek M., Yee M.C., Walter J.C., Cimprich K.A., (2005). Functional uncoupling of MCM2-7 helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev*. 19(9):1040-52.

Cha R.S., Kleckner N., (2002). ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science.* 297(5581):602-6.

Cliby W.A., Roberts C.J., Cimprich K.A., Stringer C.M., Lamb J.R., Schreiber S.L., Friend S.H., (1998). Overexpression of a kinase-inactive ATR protein causes sensitivity to DNA-damaging agents and defects in cell cycle checkpoints. *EMBO J*. 1998 Jan 2;17(1):159-69.

de Lange T., (2005). Shelterin: the protein complex that shake and safeguards human telomeres. *Genes Dev* 19:2100-2110.

Debatisse M., Le Tallec B., Letessier A, Dutrillaux B, Brison O., (2012). Common fragile sites: mechanisms of instability revisited. *Trends Genet.* 28(1):22-32.

Denison SR, Callahan G, Becker NA, Phillips LA, Smith DI, (2003). Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. *Genes Chromosomes Cancer*. 38(1):40-52.

Dillon L. W., Burrow A. A., Wang Y. H., (2010). DNA instablity at chromosomal fragile sites in cancer. *Curr. Genomics*. 11(5):326-37.

Durkin, S.G. and Glover, T.W., (2007). Chromosome fragile sites. *Annu. Rev. Genet*. 41:169-92.

Focarelli, M.L. et al. (2006) SMC1 inhibition results in FRA3B expression but has no effect on its delayed replication. Mutat. Res. 595, 23–28.

Gacy A.M., Goellner G., Juranić N., Macura S., McMurray C.T., (1995). Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell*. 81(4):533-40.

Gaddini L., Pelliccia F., Limongi M.Z. and Rocchi A., 1995. Study of the relationships between fragile sites, chromosome breaks and sister chromatid exchanges. *Mutagenesis*, 11, 3: 257-260.

Gilbert D.M., (2007). Replication origin plasticity, Taylor-made: inhibition vs recruitment of origins under conditions of replication stress. *Chromosoma*. 116(4):341-7.

Gilson E. and Geli V., (2007). How telomeres are replicated. *Nat. Rev. Mol. Cell Biol*. 8(10):825-38.

Glover T. W., (1981). FUdR induction of the X chromosome fragile site: evidence for the mechanism of folic acid and thymidine inhibition. *Am. J. Hum. Genet*. 33(2):234-42.

Glover T.W., Berger C., Coyle J., Echo B., (1984). DNA polymerase a inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet*. 67(2):136-42.

Gómez M, and Antequera F., (2008). Overreplication of short DNA regions during S phase in human cells. *Genes Dev*. 22(3):375-85.

Gomez-Gonzalez B., Felipe-Abrio I., Aguilera A., (2009). The S-phase checkpoint in required to respond to R-loops accumulated in THO mutants. *Mol. Cell. Biol.* 29:5203-5213.

Gosden J.R., (1994). Chromosome analysis protocols. *Methods in molecular biology*. Vol 29, 1994.

Gu Y., Shen Y., Gibbs R.A., Nelson D.L., (1996). Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. *Nat. Genet*. 13(1):109-13.

Harvey J., Judge C., Wiener S., (1977). Familial X-linked mental retardation with an X chromosome abnormality. *J. Med. Genet*. 14(1):46-50.

Hayashi M, Katou Y, Itoh T, Tazumi A, Yamada Y, Takahashi T, Nakagawa T, Shirahige K, Masukata H., (2007). Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. *EMBO J.* 26(11):2821.

Hellman A,, Rahat A., Scherer S.W., Darvasi A., Tsui L.C., Kerem B., (2002). Replication delay along FRA7H, a common fragile site on human chromosome 7, leads to chromosomal instability. *Mol. Cell Biol*. 20(12):4420-7.

Helmrich A., Ballarino M., Tora L., (2011). Collisions between replication and transcription complexes cause common fragile sites instability at the longest human genes. *Mol. Cell.* 44:966-977.

Hewett D,R., Handt O., Hobson L.., Mangelsdorf M, Eyre H.J., Baker E., Sutherland G.R., Schuffenhauer S., Mao J.I., Richards R.I., (1998). FRA10B structure reveals common elements in repeat expansion and chromosomal fragile site genesis. *Mol Cell*. 1(6):773-81.

Horwitz SB, Chang CK, Grollman AP., (1971). Studies on Camptothecin. I. Effects of nucleic acid and protein synthesis. *Mol. Pharmacol*. 7(6):632-44.

Hosseini S.A., Horton S., Saldivar J.C., Miuma S., Stampfer M.R., Heerema N.A., Huebner K., (2013). Common chromosome fragile sites in human and murine epithelial cells and FHIT/FRA3B loss-induced global genome instability. *Genes Chromosomes Cancer*. 52(11):1017-29.

Howlett N. G., Taniguchi T., Durkin S.G., D'Andrea A.D., Glover T.W., (2005). The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. *Hum. Mol. Genet*. 14(5):693-701.

Huberman J.A. and Riggs A.D., (1966). Autoradiography of chromosomal DNA fibers from Chinese hamster cells. *Proc. Natl Acad. Sci.* 55,599-606.

Inoue H, Ishii H, Alder H, Snyder E, Druck T, Huebner K, Croce CM, (1997). Sequence of the FRA3B common fragile region: implications for the mechanism of FHIT deletion. *Proc. Natl. Acad. Sci.* USA. 94(26):14584-9.

Isrie M, Wuyts W, Van Esch H, Devriendt K., (2014). Isolated terminal limb reduction defects: extending the clinical spectrum of Adams-Oliver syndrome and ARHGAP31 mutations. Am J Med Genet A. 2014 Jun; 164A(6):1576-9.

Kalejta RF, Li X, Mesner LD, Dijkwel PA, Lin HB, Hamlin JL., (1998). Distal sequences, but not ori-beta/OBR-1, are essential for initiation of DNA replication in the Chinese hamster DHFR origin. *Mol Cell*. 2(6):797-806.

Kastan M. B., Bartek J., (2004). Cell-cycle checkpoints and cancer. *Nature.* 432(7015):316-323.

Le Beau M.M., Rassool F.V., Neilly M.E., Espinosa R. 3rd, Glover T.W., Smith D.I., McKeithan T.W., (1998). Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. *Hum. Mol. Genet*. 7(4):755-61.

Le Tallec B, Dutrillaux B, Lachages AM, Millot GA, Brison O, Debatisse M., (2011). Molecular profiling of common fragile sites in human fibroblasts. *Nat. Struct. Mol. Biol*. 18(12):1421-3.

Le Tallec B., Millot G.A., Blin M.E., Brison O., Dutrillaux B., Debatisse M., (2013). Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. *Cell Rep.* 4(3):420-8.

Letessier A., Millot G.A., Koundrioukoff S., Lachagès A.M., Vogt N., Hansen R.S., Malfoy B., Brison O., Debatisse M., (2011). Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*. 470(7332):120-3.

Limongi MZ, Pelliccia F, Rocchi A., (2003). Characterization of the human common fragile site FRA2G. *Genomics*. 81(2):93-7.

Lopes M., Cotta-Ramusino C., Pellicioli A., Liberi G., Plevani P., Muzi-Falconi M., Newlon C.S., Foiani M., (2001). The DNA replication checkpoint response stabilizes stalled replication forks. *Nature*. 412(6846):557-61.

Losada A. and Hirano T., (2005). Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev*. 19(11):1269-87.

Lukusa T, Fryns JP., (2008). Human chromosome fragiliy. *Biochimica et Biophysica Acta*. 1779(1): 3–16.

Magenis R.E., Hecht F., Lovrien E.W., (1970). Heritable fragile site on chromosome 16: probable localization of haptoglobin locus in man. *Science*. 170(3953):85-7.

Martin JP, Bell J., (1943). A PEDIGREE OF MENTAL DEFECT SHOWING SEX-LINKAGE. *J Neurol Psychiatry*. 6(3-4):154-7.

McAvoy S., Ganapathiraju S.C., Ducharme-Smith A.L, et al.,(2007).Non-random inactivation of large common fragile site genes in different cancers. *Cytogenet. Genome Res.* 118:260-269.

McMurray C. T. (2010). Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* 11(11):786-99.

Méchali M, (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat. Rev. Mol. Cell. Biol*. 11(10):728-38.

Mesner L.D., Li X, Dijkwel PA, Hamlin JL., (2003). The dihydrofolate reductase origin of replication does not contain any nonredundant genetic elements required for origin activity. *Mol. Cell. Biol*. 23(3):804-14.

Mimori K, Druck T, Inoue H, Alder H, Berk L, Mori M, Huebner K, Croce CM (1999). Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. *Proc Natl Acad Sci U S A*. 96(13):7456-61.

Mishmar D., Mandel-Gutfroind Y., Margalit H., Kerem B., (1999). Common fragile sites: G-band characteristics within an R-band. *Am. J. Hum. Genet.* 64:908–910.

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui LC, Rosenthal A, Kerem B (1998). Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. *Proc Natl Acad Sci U S A*. 95(14):8141-6.

Mona Yekezare, Belen Gomez-Gonzalez and John F. X. Diffley, (2010). Controlling DNA replication origins in response to DNA damage – inhibit globally, activate locally. *Journal of Cell Science*. 126, 1297–1306.

Morelli C, Karayianni E, Magnanini C, Mungall AJ, Thorland E, Negrini M, Smith DI, Barbanti-Brodano G., (2002). Cloning and characterization of the common fragile site FRA6F harboring a replicative senescence gene and frequently deleted in human tumors. *Oncogene*. 21(47):7266-76.

Mortusewicz O., Herr P., Helleday T., (2013). Early replication fragile sites: where replication-transcription collisions cause genetic instability. *EMBO*. J. 32(4)493-495.

Mrasek K., Schoder C., Teichmann A.C., Behr K., Franze B., Wilhelm K., Blaurock N., Claussen U., Liehr T., Weise A., (2010). Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones. *Int. J. Oncol*. 36(4):929-40.

Murano I., Kuwano A., Kajii T. (1989). Fibroblast-specific common fragile sites induced by aphidicolin. *Hum. Genet*. 83(1):45-8.

Murano I., Kuwano A., Kajii T., (1989). Cell type-dependent difference in the distribution and frequency of aphidicolin-induced fragile sites: T and B lymphocytes and bone marrow cells. *Hum. Genet*. 84(1):71-4.

Musio A., Montagna C., Mariani T., Tilenni M., Focarelli M.L., Brait L., Indino E., Benedetti P.A., Chessa L., Albertini A., Ried T., Vezzoni P., (2005). SMC1 involvement in fragile site expression. *Hum. Mol. Genet*. 14(4):525-33.

Nakamura A, Sedelnikova OA, Redon C, Pilch DR, Sinogeeva NI, Shroff R, Lichten M, Bonner WM. (2006). Techniques for gamma-H2AX detection. *Methods Enzymol*. 409:236-50

O'Connor PM, Nieves-Neira W, Kerrigan D, Bertrand R, Goldman J, Kohn KW, Pommier Y., (1999). S-phase population analysis does not correlate with the cytotoxicity of camptothecin and 10,11-methylenedioxycamptothecin in human colon carcinoma HT-29 cells. *Cancer Commun*. 3(8):233-40.

Oestergaard VH & Lisby M, (2017). Transcription-replication conflicts at chromosomal fragile sites—consequences in M phase and beyond. *Chromosoma* (2017) 126:213–222.

Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K (1996). The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell*. 84(4):587-97.

Pacek M., Tutter A.V., Kubota Y., Takisawa H., Walter J.C. (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell*. 21(4):581-7.

Palakodeti A., Han Y., Jiang Y., Le Beau M.M., (2004). The role of late/slow replication of the FRA16D in common fragile site induction. *Genes Chromosomes Cancer*. 39(1):71-6.

Pasic I, Shlien A, Durbin AD, Stavropoulos DJ, Baskin B, Ray PN, Novokmet A, Malkin D, (2010). Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. *Cancer Res*. 70(1):160-71.

Patil M., Pabla N., Dong Z., (2013). "Checkpoint kinase 1 in DNA damage response and cell cycle regulation". *Cell. Mol. Life Sci*. 70 (21): 4009–21.

Pelliccia F., Rocchi A., (1992). The effect of caffeine on DAPI-inducible fragile sites. *Mutation Research Letters*. 282(1):43-48.

Pincheira J., Bravo M., Navarrete M.H., Marcelain K., López-Sáez J.F., de la Torre C. (2001). Ataxia telangiectasia: G2 checkpoint and chromosomal damage in proliferating lymphocytes. *Mutagenesis*. 2001 Sep;16(5):419-22.

Ried K, Finnis M, Hobson L, Mangelsdorf M, Dayan S, Nancarrow JK, Woollatt E, Kremmidiotis G, Gardner A, Venter D, Baker E, Richards RI (2000). Common chromosomal fragile site FRA16D sequence: identification of the FOR gene spanning FRA16D and homozygous deletions and translocation breakpoints in cancer cells. Hum Mol Genet. 9(11):1651-63.

Sahar E., Latt S.A., (1978). Enhancement of banding patterns in human metaphase chromosomes by energy transfer. *Proc. Natl. Acad. Sci. USA.* 75(11):5650-5654.

Schwartz M., Zlotorynski E., Kerem B. (2006). The molecular basis of common and rare fragile sites. *Cancer Lett.* Jan 28;232(1):13-26.

Seo J., Kim K., Chang D. Y., Kang H.B., Shin E. C., Kwon J., Choi J. K., 2013. Genome-wide reorganization of histone H2AX toward

particular fragile sites on cell activation. *Nuecleic Acids Res.* 42 (2): 1016-1025.

Sutherland G. R., Jacky P. B., Baker E. G. (1984). Heritable fragile sites on human chromosomes. XI. Factors affecting expression of fragile sites at 10q25, 16q22, and 17p12. *Am. J. Hum. Genet.* 36(1):110-22.

Sutherland G.R. (1979). Heritable fragile sites on human chromosomes I. Factors affecting expression in lymphocyte culture. *Am. J. Hum. Genet*. 31(2):125-35.

Szekeres T., et al. 1997. The enzyme ribonucleotide reductase: target for antitumor and anti-HIV therapy. *Crit. Rev. Clin. Lab. Sci.* 34(6): 503-528.

Taniguchi T. and D'andrea A.D., (2002). The Fanconi anemia protein, FANCE, promotes the nuclear accumulation of FANCC. *Blood*. 100(7):2457-62.

Tercero J.A. and Diffey J.F.X. (2001). Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature*. 412(6846):553-7.

Tuduri S, Crabbé L, Conti C, Tourrière H, Holtgreve-Grez H, Jauch A, Pantesco V, De Vos J, Thomas A, Theillet C, Pommier Y, Tazi J, Coquelle A, Pasero P., (2009). Topoisomerase I suppresses genomic

instability by preventing interference between replication and transcription. *Nat. Cel.l Biol*. 11(11):1315-24.

Verkerk A.J., Pieretti M., Sutcliffe J.S., Fu Y.H., Kuhl D.P., Pizzuti A., Reiner O., Richards S., Victoria M.F., Zhang F.P., et al. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*. 65(5):905-14.

Walter J., Newport J., et al., 2000. Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Mol. Cell*. 5(4):617-27.

Wang J.C., (2002). Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol*. 3(6):430–40.

Wright WE, Tesmer VM, Liao ML, Shay JW, (1999). Normal human telomeres are not late replicating. *Exp. Cell. Res.* 251:492-499.

Yu S., Mangelsdorf M., Hewett D., Hobson L., Baker E., Eyre H.J., Lapsys N., Le Paslier D., Doggett N.A., Sutherland G.R., Richards R.I., (1997). Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. *Cell*. 88(3):367-74.

Yunis JJ, Soreng AL., (1984). Constitutive fragile sites and cancer. *Science. 226(4679):1199-204.*

Zlotorynski E., et al. (2003). Molecular basis for expression of common and rare fragile sites. *Mol. Cell. Biol. 23:7143-51*.

LIST OF PUBLICATIONS

Capitano F., Gargiuli C., Angerilli A., Maccaroni K., Pelliccia F., Mele A., Camilloni G., (2016). RNA polymerase I transcription is modulated by spatial learning in different brain regions.
J. Neurochem. 136(4):706-716.

Pelliccia F., Genovesi ML., Maccaroni K., (2015). Fanconi anaemia, chromosome instability, DNA replication and fragile sites.
Eur. J Hum Genet, 23, suppl.1, 451.