



Universitat Autònoma de Barcelona

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  [http://cat.creativecommons.org/?page\\_id=184](http://cat.creativecommons.org/?page_id=184)

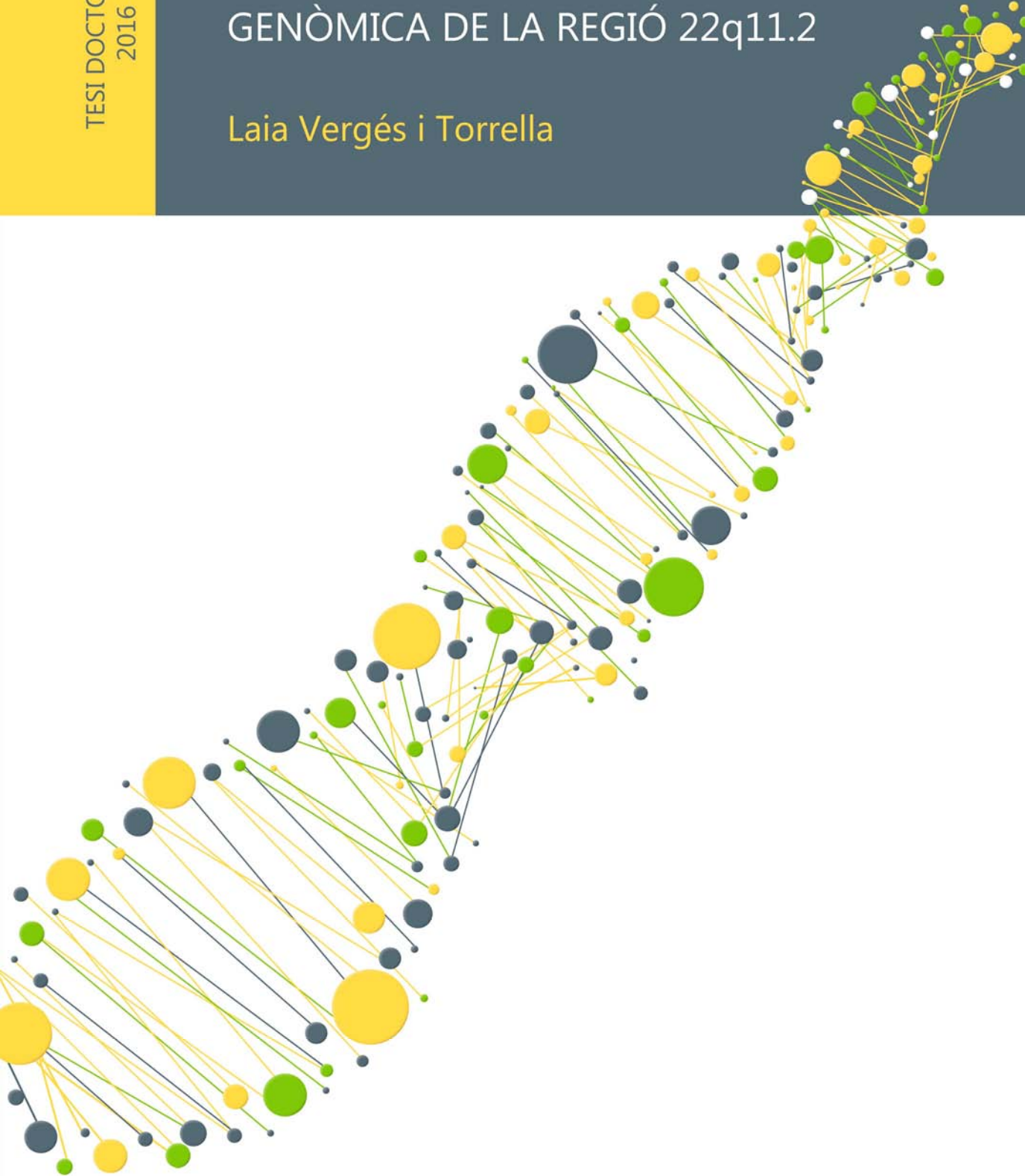
**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

TESI DOCTORAL  
2016

# AVALUACIÓ DELS FACTORS DE PREDISPOSICIÓ A LA INESTABILITAT GENÒMICA DE LA REGIÓ 22q11.2

Laia Vergés i Torrella





---

**Avaluació dels factors de  
predisposició a la inestabilitat  
genòmica de la regió 22q11.2**

---

Laia Vergés i Torrella

Memòria presentada per optar al Grau de Doctora en Biologia Cel·lular  
per la Universitat Autònoma de Barcelona

Tesi doctoral inscrita en el Departament de Biologia Cel·lular,  
Fisiologia i Immunologia, Facultat de Biociències, Universitat  
Autònoma de Barcelona

Directors:

Dra. Francesca Vidal Domínguez

Dr. Joan Blanco Rodríguez

Bellaterra, 2016





Universitat Autònoma de Barcelona

La **Dra. Francesca Vidal Domínguez**, catedràtica del Departament de Biologia Cel·lular, Fisiologia i Immunologia de la Universitat Autònoma de Barcelona,

i el **Dr. Joan Blanco Rodríguez**, professor agregat del Departament de Biologia Cel·lular, Fisiologia i Immunologia de la Universitat Autònoma de Barcelona,

### CERTIFIQUEN

Que Laia Vergés Torrella ha realitzat sota la seva direcció el treball d'investigació que s'exposa en la memòria titulada "**Avaluació dels factors de predisposició a la inestabilitat genòmica de la regió 22q11.2**" per optar al grau de Doctor per la Universitat Autònoma de Barcelona.

Que aquest treball s'ha dut a terme a la Unitat de Biologia Cel·lular del Departament de Biologia Cel·lular, Fisiologia i Immunologia de la Universitat Autònoma de Barcelona.

I perquè així consti, signen el present certificat.

Dra. Francesca Vidal Domínguez

Dr. Joan Blanco Rodríguez

Bellaterra, 17 de juny de 2016



La present tesi doctoral s'ha desenvolupat amb el finançament dels projectes:

- Projecte UAB CF-180034, Universitat Autònoma de Barcelona
- SAF 2010-22241, Ministerio de Ciencia e Innovación, Gobierno de España
- SGR2014-524, Agència de Gestió d'Ajuts Universitaris i de Recerca, Generalitat de Catalunya

Laià Vergés Torrella ha estat beneficiària d'una beca predoctoral per a Personal Investigador en Formació (PIF), concedida per la Universitat Autònoma de Barcelona (2012).





# PRÒLEG

Des dels primers simis com la *Pliobates cataloniae*, coneguda casualment amb el nom de Laia, fins a nosaltres, el DNA ha experimentat molts canvis. Tal com els modernistes feien en el trencadís, el DNA s'ha esmicolat, copiat i enganxat milions de vegades des de l'origen de la primera Laia fins a la que aquí escriu la seva tesi. De tant en tant, aquest procés ha generat modificacions de la informació genètica que si per fortuna han suposat una avantatge han estat preservades per la selecció natural i han contribuït a l'evolució. Aquest ball de bastons que han dansat els primats durant milions d'anys ha deixat estelles en el nostre genoma en forma de seqüències repetitives del DNA. Desencertadament, durant molt temps s'han catalogat aquestes repeticions com a DNA escombrària. Però en realitat, aquestes seqüències no són merament una empremta evolutiva sinó que constitueixen el motor que impulsa la formació de nous canvis en el genoma.

Per què els xerpes de l'Himàlaia estan més ben adaptats a les altes alçades? Algú podria respondre a aquesta pregunta tot dient que perquè han nascut a gran alçada. Però segurament, si haguessin nascut a Sabadell tampoc patirien el mal d'alçada com el que pateixo jo dalt d'un cim a 4000m. Darwin potser hagués suggerit que som espècies de pinsans diferents però s'hauria equivocat. I tampoc es tracta d'una qüestió de variants d'un gen perquè en aquest cas no estem parlant de si som verds o grocs com els pèsols de Mendel. La clau està en les variacions del número de còpies de certs gens causades per les reorganitzacions de seqüències repetitives del DNA que ha originat l'evolució.

Tot i que, en l'exemple dels tibetans aquesta variació ha sigut molt positiva per la seva adaptació, ja us podeu imaginar que la cosa no sempre va tan bé. Al igual que quan barreges dos gelats de gustos diferents el resultat no sempre és deliciós, és possible que et deixi indiferent o que et causi una bona indigestió. I quan el biòlegs traslladem la gelateria a la genètica aquesta indigestió assoleix el grau de malaltia.

Així doncs, tot el que ens ha fet evolucionar també ens està fent més fràgils? El DNA és com un vell disc de vinil, que com més sona més es ratlla, i això fa que de vegades s'encalli i la música soni enrogallada. La natura és enrevessadament irònica i ha creat en nosaltres un exemple perfecte del Yin i el Yang. Els humans podem parlar gràcies a l'àrea de Broca, escriure gràcies a l'articulació en cadira de muntar del dit polze, fins i tot, construïm avions i gratacels però a canvi tenim malalties que no tenen els babuïns del Serengeti. L'únic consol que ens queda és que, per sort, també hem dissenyat les eines necessàries per estudiar aquestes malalties. Aquesta tesi, doncs, parla de com es fan les ratllades del disc de vinil, de quines són les combinacions de gelats que menys ens agraden i de les estrelles del nostre genoma.

**Laia Vergés i Torrella**

# ÍNDEX DE CONTINGUTS



RESUM · ABSTRACT	1
<b>1. INTRODUCCIÓ</b>	<b>5</b>
1.1 Organització del genoma humà	7
1.2 Variabilitat del genoma humà	9
1.3 Mecanismes de formació de les reorganitzacions estructurals	10
Reorganitzacions recurrents	10
Reorganitzacions no recurrents	12
1.4 Estudi de la variabilitat estructural	14
Anàlisi genòmica	15
Anàlisi de regions específiques	17
1.5 Contribució de les reorganitzacions estructurals al genoma	18
Evolució	19
Polimorfismes	19
Trastorns genòmics	21
1.6 Risc de recurrència dels trastorns genòmics	22
1.7 Factors de predisposició dels trastorns genòmics	24
Variabilitat estructural dels LCR	25
Inversió de la regió crítica	26
Genotip de PRDM9	27
1.8 Regió cromosòmica 22q11.2	29
Reorganitzacions de la regió 22q11.2	29
Síndrome de DiGeorge/velocardiofacial	31
Risc de recurrència i factors de predisposició en la SDG/VCF	32

## 2. OBJECTIUS 35

## 3. MATERIAL I MÈTODES 39

3.1 Mostres biològiques i processament	41
Semen	41
Sang	41
3.2 Tècniques citogenètiques basades en la hibridació <i>in situ</i> fluorescent (FISH)	45
FISH en espermatozoides	45
FISH en fibres de cromatina	49
FISH en nuclis interfàsics	56
3.3 Tècniques moleculars basades en la reacció en cadena de la polimerasa (PCR)	60
Microsatèl·lits	60
Droplet Digital PCR (ddPCR)	61
Genotipatge	64

## 4. RESULTATS 67

4.1 Reorganitzacions en espermatozoides	69
Delecions i duplicacions de la regió 22q11.2 en pares SDG/VCF	69
Delecions i duplicacions de la regió 7q11.23, 15q11q13 i 22q11.2 en pares transmissors de la SDG/VCF	71
4.2 Característiques genètiques	76
Blocs de seqüències en tàndem dels LCR22-2 i LCR22-4	76
Seqüències paràlogues dels LCR22	77
Inversions 22q11.2	79
Variants al·lèliques de PRDM9	81

<b>5. DISCUSSIÓ</b>	<b>83</b>
5.1 NAHR en pares transmissors SDG/VCF	85
5.2 Factors de predisposició a la SDG/VCF	89
Arquitectura genòmica dels LCR	89
Inversions en heterozigosi de la regió crítica	97
Genotip de <i>PRDM9</i>	98
<b>6. CONCLUSIONS</b>	<b>103</b>
BIBLIOGRAFIA	107
ABREVIACIONS	122





# RESUM

El genoma humà està format en un 5% per low-copy repeats (LCR); segments de DNA d'entre 1 i 500 Kb que es repeteixen dues o més vegades al llarg del genoma i que comparteixen una homologia superior al 90%. L'aparellament no al·lèlic d'aquestes seqüències durant la meiosi promou la generació d'esdeveniments de recombinació homòloga no al·lèlica (NAHR) que donen lloc a reorganitzacions cromosòmiques heretables que originen un conjunt de malalties anomenades trastorns genòmics. La deleció de la regió 22q11.2, originada per la NAHR entre els LCR22, causa la síndrome de DiGeorge/velocardiofacial (SDG/VCF), un dels trastorns genòmics més freqüents en els humans.

L'objectiu principal d'aquesta tesi doctoral va ser l'avaluació de les característiques genètiques que modulen la susceptibilitat a la NAHR en individus transmissors de la SDG/VCF.

Es va caracteritzar la susceptibilitat a la NAHR en 8 homes amb descendència afectada per la SDG/VCF. Aquest estudi es va dur a terme a través de l'anàlisi de la freqüència d'espermatozoides portadors de delecions i duplicacions de la regió 22q11.2, i d'altres regions crítiques amb una arquitectura genòmica similar, mitjançant la hibridació in situ fluorescent (FISH) en nuclis descondensats d'espermatozoides. Els resultats obtinguts van demostrar que 2 parets transmissors de la SDG/VCF presentaven increments significatius d'espermatozoides portadors de la deleció 22q11.2 que es van atribuir a un increment de la NAHR intra-cromàtide focalitzada a la regió 22q11.2. En aquests individus, també es van analitzar els blocs de seqüències en tàndem dels LCR22-2 i LCR22-4 mitjançant la tècnica de la FISH en fibres de cromatina (fiber-FISH) obtingudes de leucòcits. Els resultats de l'anàlisi van mostrar variacions respecte la població control dels blocs de seqüències en tàndem corresponents als fòsmids L9 i K3. Aquestes variacions podrien constituir un factor de predisposició a la NAHR.



Per tal de validar els resultats obtinguts mitjançant fiber-FISH, es van analitzar les variacions de seqüències paràlogues en els LCR22 cobertes pels fòsmids L9 i K3 mitjançant la reacció en cadena de la polimerasa de tipus droplet digital (ddPCR). No es van observar diferències significatives entre les poblacions d'individus transmissors de la SDG/VCF i de controls analitzades. Tot i així, els resultats van suggerir una disminució de còpies del pseudogen de AK129567 en els individus transmissors.

D'altra banda, també es va analitzar la freqüència de la inversió 22q11.2 mitjançant la FISH interfàsica en limfòcits. No es va detectar cap haplotip d'inversió 22q11.2 en individus transmissors ni en controls. Així doncs, es va confirmar la inexistència d'aquest polimorfisme i es va excloure la seva implicació com a factor de predisposició per la SDG/VCF. Per últim, es va determinar el genotip de PRDM9 mitjançant PCR i seqüenciació Sanger. La comparació de les freqüències al·lèliques entre els individus transmissors i la població control no va mostrar diferències significatives. Tot i així, es va descriure un al·lel nou de PRDM9, que es va denominar L50, en un pare transmissor de la SDG/VCF que havia mostrat increments de delecions 22q11.2 en espermatozoides. Aquest resultat va suggerir que certs al·lells rars de PRDM9 podrien actuar com a factor de predisposició per la NAHR a la regió 22q11.2.

En conjunt, l'avaluació dels factors de predisposició a la inestabilitat genòmica de la regió 22q11.2 va reflectir que el risc a la NAHR en aquesta regió és complex i atribuïble a la confluència de diferents característiques genètiques.

# ABSTRACT

The 5% of the human genome is constituted by low copy repeats (LCRs). LCRs are DNA fragments from 1 to 500 kb in size, with at least two copies across the genome that share a high level of sequence identity (> 90%). The non-allelic alignment of these sequences during meiosis promotes non-allelic homologous recombination events (NAHR) that could lead to chromosome reorganizations which can be transmitted to the offspring. These reorganizations originate a group of diseases called genomic disorders. Deletions of the 22q11.2 region, caused by NAHR between LCR22, result in the DiGeorge/velocardiofacial syndrome (DGS/VCFS) which is one of the most frequent genomic disorder in humans.

The main objective of this thesis was to evaluate the genetic features that could modulate the NAHR susceptibility in DGS/VCFS transmitting individuals.

NAHR susceptibility was analysed in a total of eight men with DGS/VCFS affected progeny. This study was achieved by analysing the frequency of deletions and duplications of the 22q11.2 region, and other critical regions with a similar genomic architecture, using fluorescence in situ hybridization (FISH) in decondensed sperm nuclei. Results demonstrated that two transmitting fathers showed statistical significant increases of 22q11.2 deletions in sperm. This abnormality was attributed to an abnormal intra-chromatid NAHR activity focused on the 22q11.2 region. By applying FISH on chromatin fibers obtained from leucocytes (fiber-FISH), the blocks of tandem repeats within LCR22-2 and LCR22-4 were also evaluated for these individuals. With respect to the control population, SDG/VCF transmitting fathers showed copy number variations on tandem sequence blocks corresponding to the L9 and K3 fosmid sequences that could constitute a predisposing factor for NAHR.

In order to validate fiber-FISH data, variations of paralogous sequences covered by L9 and K3 fosmids were analysed by using droplet digital polymerase chain reaction

(ddPCR). Although differences between the population of DGS/VCFS transmitting parents and control individuals were not observed, the results suggested a decrease of the AK129567 pseudogene copy number in transmitting parents. The frequency of 22q11.2 inversions was also studied by applying interphase FISH in lymphocytes. The inversion haplotype was not detected in any transmitting or control individuals. Therefore, the absence of the 22q11.2 inversion polymorphism was confirmed and this genetic feature was excluded as a possible predisposing factor for DGS/VCFS. Finally, PRDM9 genotype was assessed by using PCR and Sanger sequencing. Allelic frequencies of control and DGS/VCFS transmitting individuals did not show differences. Nevertheless, we described a novel PRDM9 allele, L50, in a DGS/VCFS transmitting father with increased rates of 22q11.2 deletions in sperm. This observation suggested that certain rare alleles of PRDM9 might be a predisposing factor for NAHR at 22q11.2 region.

In summary, the evaluation of predisposing factors for the genomic instability at 22q11.2 showed that the NAHR risk in this region is complex and could be attributed to a confluence of different genetic features.



# INTRODUCCIÓ

- 1.1 ORGANITZACIÓ DEL GENOMA HUMÀ
- 1.2 VARIABILITAT DEL GENOMA HUMÀ
- 1.3 MECANISMES DE FORMACIÓ DE LES REORGANITZACIONS ESTRUCTURALS
- 1.4 ESTUDI DE LA VARIABILITAT ESTRUCTURAL
- 1.5 CONTRIBUCIÓ DE LES REORGANITZACIONS ESTRUCTURALS AL GENOMA
- 1.6 RISC DE RECURRÈNCIA DELS TRASTORNS GENÒMICS
- 1.7 FACTORS DE PREDISPOSICIÓ DELS TRASTORNS GENÒMICS
- 1.8 REGIÓ CROMOSÒMICA 22q11.2



## 1.1 ORGANITZACIÓ DEL GENOMA HUMÀ

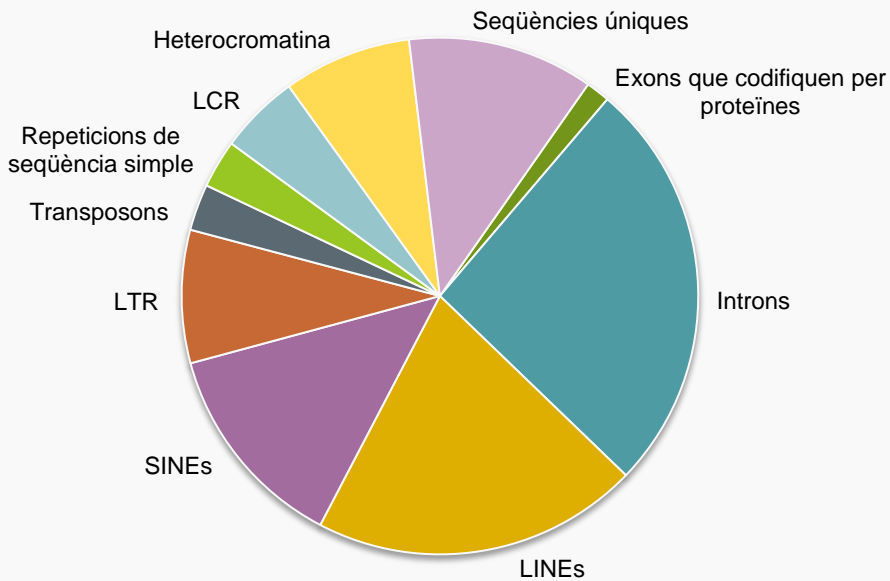
El genoma humà està format per aproximadament tres mil dos-cents milions de parells de nucleòtids. La publicació de la primera seqüenciació completa el febrer de l'any 2001, per part del consorci públic *Projecte Genoma Humà* i de la companyia privada *Celera Genomics*, va revolucionar el coneixement sobre la organització del genoma i va possibilitar la realització d'estudis fins aleshores inabastables (LANDER ET AL. 2001; VENTER ET AL. 2001).

La disponibilitat d'un mapa del DNA humà permetia abordar l'anàlisi amb detall dels seus components. Es determinà que un 30% del nostre genoma correspon a seqüències relacionades amb gens. Tot i així, tan sols un 1,5% conté exons que codifiquen proteïnes, mentre que la resta de seqüències gèniques corresponen principalment a introns. El 70% restant està constituït per DNA extragènic que correspon majoritàriament a seqüències repetitives (LANDER ET AL. 2001) (**Figura 1**).

Les seqüències repetitives es poden dividir en dues grans categories: les repeticions comunes i les duplicacions segmentals (SD de l'anglès *segmental duplication*), també anomenades *low-copy repeats* (LCR). Entre les repeticions comunes destaquen els elements transposables que són fragments de DNA o RNA capaços de reproduir-se i inserir-se en el genoma hoste. Els elements transposables es classifiquen en *Long Interspersed Nuclear Elements* (LINEs) (20,4%), *Short Interspersed Nuclear Elements* (SINEs) (13,1%), retrotransposons *long terminal repeat* (LTR) (8,3%) i transposons de DNA (2,9%). Les repeticions comunes també inclouen repeticions de seqüències simples inferiors a 10 pb (3%) (**Figura 1**).

El DNA extragènic també inclou seqüències de DNA que s'organitzen en forma d'heterocromatina (8%) i de seqüències úniques localitzades entre gens codificants (11,6%) (GREGORY 2005) (**Figura 1**).

Els LCR són segments de DNA d'entre 1 i 500 kb que es repeteixen dues o més vegades al llarg del genoma, en tàndem o en diverses localitzacions i de forma directa o invertida. Es caracteritzen per presentar un grau d'homologia superior al 90% i contenir



■ **Figura 1.** Tipus de seqüències que conformen el genoma humà. Adaptat de Gregory (2005).

gens, pseudogens, motius de recombinació i seqüències de DNA repetitiu com les seqüències *Alu* (BABCOCK ET AL. 2003; DITTWALD ET AL. 2013B; BAILEY I EICHLER 2006). En una primera aproximació a la detecció de LCR, Bailey et al. (2001) determinà un enriquiment d'aquestes seqüències en les regions pericentromèriques i subtelomèriques. Tanmateix, l'estimació del total de LCR en els primers assemblatges del genoma humà era complicada degut a què les seqüències amb una identitat superior al 98% es sobreposaven entre elles i, per tant, s'infravaloraven (BAILEY ET AL. 2001).

La primera estimació robusta de LCR en humans va ser aconseguida l'any 2002. El resultat fou que el genoma humà contenia un 5,2% de LCR de mida >1 kb i amb una identitat >90% (BAILEY ET AL. 2002). Posteriorment, s'ha re-calculat aquesta dada obtenint resultats similars en la construcció del genoma del 2009 (4,9% LCR>1 kb>90%) (HUMAN SEGMENTAL DUPLICATION DATABASE). D'altra banda, la distribució dels LCR entre els cromosomes humans no és homogènia. A part de l'enriquiment en regions pericentromèriques i subtelomèriques descrit per Bailey et al. 2001, s'ha determinat que hi ha cromosomes que presenten una concentració més elevada d'aquest tipus de repeticions. Concretament, els cromosomes 7, 9, 10, 15, 16, 17, 22, X i

Y són els que presenten una densitat més alta de LCR (>5%). Per contra, els cromosomes amb una densitat més baixa de LCR són el 3, 6, 8, 12 i 14 (<2%) (ZHANG ET AL. 2005).

## 1.2 VARIABILITAT DEL GENOMA HUMÀ

La seqüenciació del genoma humà també va permetre definir amb més precisió la variabilitat entre els genomes de diferents individus. D'acord amb la mida de la regió afectada podem distingir dos tipus de variacions: variacions puntuals i variacions estructurals.

Les variacions puntuals afecten a un sol nucleòtid per substitució o per inserció/deleció. En concret, les variacions d'un nucleòtid que s'observen en una freqüència mínima de l'1% de la població s'han anomenat SNP (de l'anglès *Single Nucleotide Polymorphisms*). S'estima que en el genoma humà hi ha aproximadament 10 milions de SNP (THE INTERNATIONAL HAPMAP CONSORTIUM 2005).

D'altra banda, les variacions estructurals poden afectar des de pocs nucleòtids fins a regions de l'ordre de megabases i inclouen deleccions, insercions, duplicacions, inversions i translocacions. Les variacions entre diferents individus que afecten el número de fragments genòmics idèntics amb una mida superior a 1kb s'han anomenat variacions en el número de còpies (CNV de l'anglès *Copy Number Variations*) (FEUK ET AL. 2006). Es calcula que la seva taxa d'aparició és quatre ordres de magnitud superior a la dels SNP (LUPSKI 2007).

Un dels factors claus de la generació de la variabilitat estructural és la formació de reorganitzacions en les quals es troben involucrades els LCR (MARQUES-BONET ET AL. 2008). De fet, les variacions en els propis LCR constitueixen una font important de variacions genòmiques (MARQUES-BONET ET AL. 2009B). Cada individu presenta un nombre variable de CNV i un 32% d'aquestes regions variables corresponen a variacions en el nombre de LCR (REDON ET AL. 2006).



## 1.3 MECANISMES DE FORMACIÓ DE LES REORGANITZACIONS ESTRUCTURALS

Les reorganitzacions estructurals poden produir-se de forma recurrent, amb una mida i uns punts de trencament fixes, o no recurrent, amb mides i punts de trencaments distints en cada esdeveniment. El principal mecanisme que origina les reorganitzacions recurrents és la recombinació homòloga no al·lèlica (NAHR de l'anglès *Non-Allelic Homologous Recombination*) (STANKIEWICZ I LUPSKI 2002; LIU ET AL. 2012). D'altra banda, les reorganitzacions no recurrents són causades majoritàriament per l'acció de la fusió d'extrems no homòlegs (NHEJ de l'anglès *Non-Homologous End-Joining*) i l'aturada de la forca de replicació i canvi de motlle (FoSTeS de l'anglès *Fork Stalling and Template Switching*) (GU ET AL. 2008; STANKIEWICZ I LUPSKI 2010).

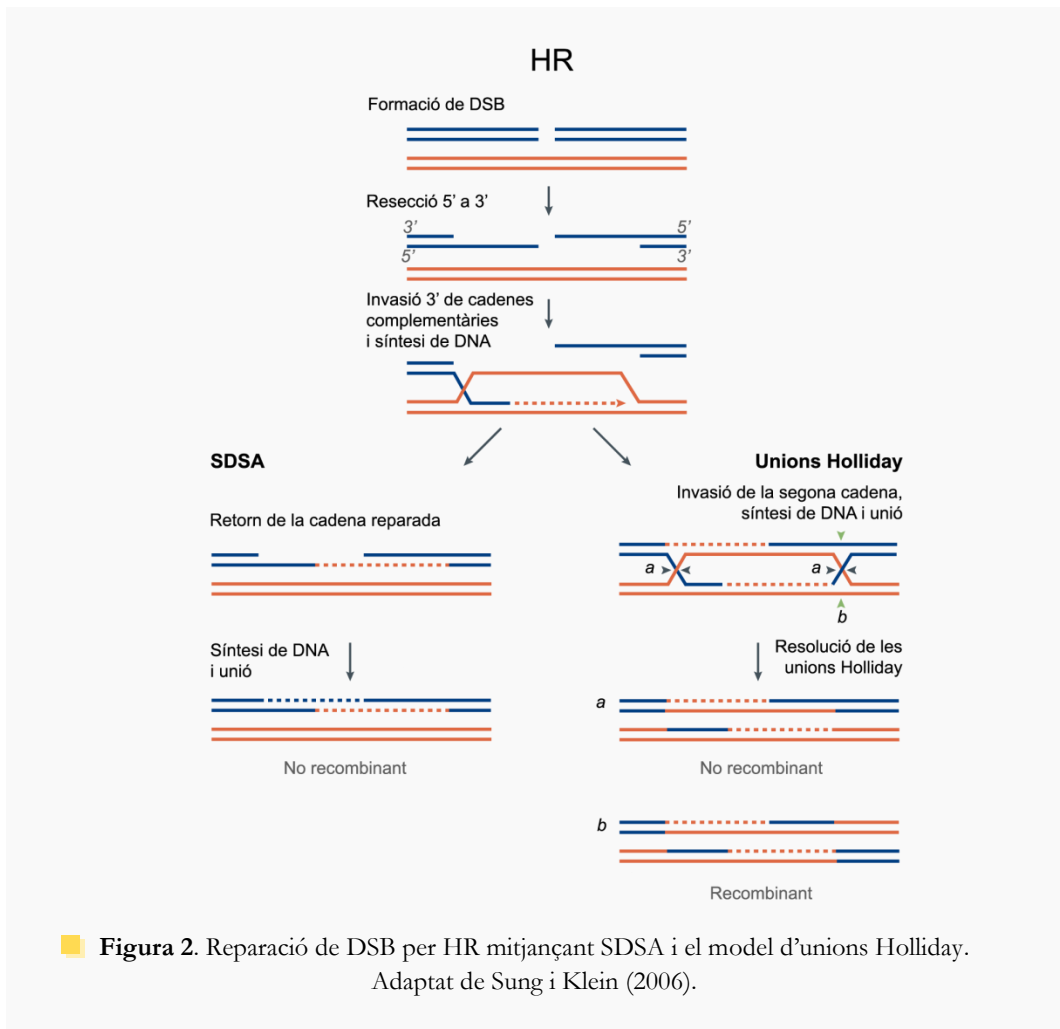
### REORGANITZACIONS RECURRENTS

La NAHR és el mecanisme majoritari per aquelles reorganitzacions que es produeixen de forma recurrent. Es considera que aquest mecanisme contribueix a la formació del 22% de les variacions estructurals presents en el genoma (KIDD ET AL. 2010). Els esdeveniments de NAHR que succeeixen durant la meiosi originen reorganitzacions estructurals que poden ser transmeses a la descendència (LIU ET AL. 2011; SASAKI ET AL. 2010). Mentre que quan la NAHR s'esdevé durant la mitosi dóna lloc a poblacions cel·lulars amb reorganitzacions estructurals, sovint relacionades amb processos cancerígens (DARAI-RAMQVIST ET AL. 2008).

El procés de la recombinació homòloga (HR de l'anglès *homologous recombination*) s'inicia amb la formació de trencaments de doble cadena del DNA (DSB de l'anglès *double-stranded breaks*). A continuació, es reconeixen els DSB i es genera una resecció dels extrems lliures 5'. En aquest punt, s'han descrit diferents vies basades en la homologia entre cadenes que actuen en la reparació del DNA. Els principals mecanismes que intervien en aquest procés són l'aparellament de cadena associat a síntesi (SDSA de l'anglès *Synthesi Dependent Strand Annealing*) i la formació de unions Holliday. En la SDSA un dels extrems lliures 3' envaeix una cadena complementària i es repara el DNA utilitzant aquesta cadena complementària com a motlle. Després de la reparació s'uneix

a la seva cadena original i actua com a motlle per reparar l'altra cadena. En canvi, en el model de formació de unions Holliday, els dos extrems 3' cerquen homologia en cadenes complementàries que serviran de motlle per la reparació del DNA. Com a resultat de l'entrellaçament de les cadenes complementàries es crea una estructura anomenada unió Holliday que dona nom a aquest model de reparació. La resolució de les SDSA sempre dona lloc a productes no recombinants ja que la cromàtide envaïda es manté intacta, mentre que en les unions Holliday es poden generar productes recombinants o no recombinants en funció de l'orientació de la resolució de l'entrecreuament de cadenes (**Figura 2**) (SUNG I KLEIN 2006).

Quan els DSB es produeixen en regions repetitives sinapsades de forma no al·lèlica o



en localitzacions pròximes a aquestes regions, la cerca d'homologia pot realitzar-se en còpies paràlogues i donar lloc a un tipus de reparació no al·lèlica. En aquesta situació, si la resolució dels DSB es produeix per la via de la unió Holliday i es resol per recombinació ens referim al procés esdevingut com a NAHR (PARKS ET AL. 2015).

Les seqüències recombinades han de presentar uns segments mínims complementaris (MEPS de l'anglès *Minimal Efficient Processing Segments*) amb una llargada de ~300 pb (REITER ET AL. 1998). Els LCR, d'acord amb la seva definició, presenten còpies paràlogues d'alta homologia i amb una longitud superior al MEPS per la qual cosa constitueixen regions sensibles a participar en aquest tipus de recombinació. Els processos de NAHR entre LCR originen diferents tipus de reorganitzacions estructurals en funció de l'orientació directa o indirecta dels LCR paràlegs implicats i de la seva localització en cromosomes no homòlegs, cromosomes homòlegs, cromàtides germanes o, fins i tot, dins una mateixa cromàtide (**Figura 3**) (OU ET AL. 2011; GU ET AL. 2008; DITTWALD ET AL. 2013A).

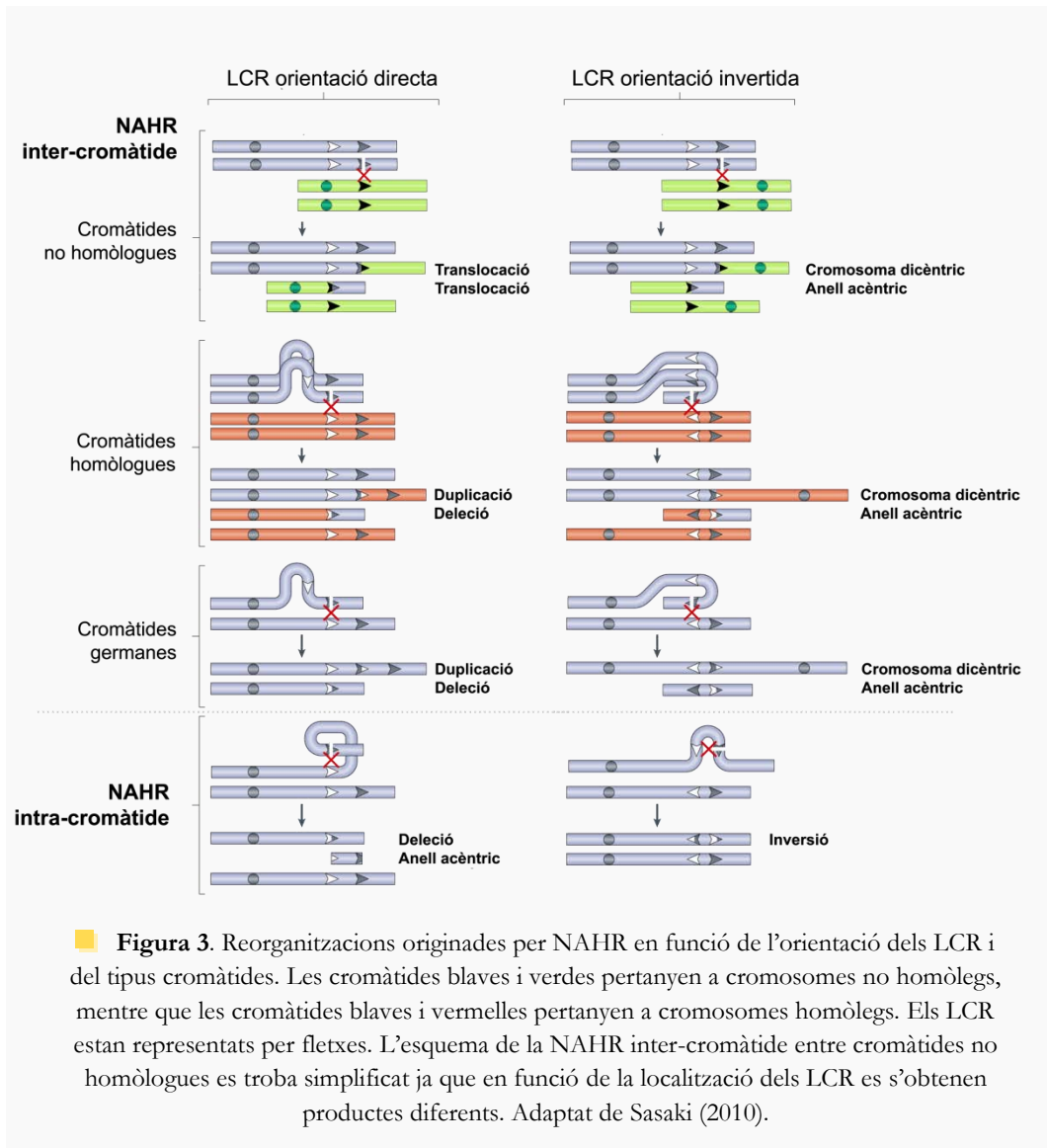
---

#### REORGANITZACIONS NO RECURRENTS

Les reorganitzacions no recurrents s'originen principalment per l'acció de la NHEJ i de mecanismes basats en errors durant la replicació com FoSTeS (**Figura 4**).

La NHEJ no requereix la presència de seqüències complementàries i, per tant, es considera independent de la presència de LCR. Tot i així, s'ha determinat que pot ser promoguda per microhomologies d'entre 5 – 25 pb (MMEJ de l'anglès *microhomology-mediated end joining*) (MCVEY I LEE 2008) i que els punts de trencament reparats per aquest mecanisme sovint coincideixen amb element repetitius com seqüències LINE, LTR o *Alu* (ZHANG ET AL. 2009).

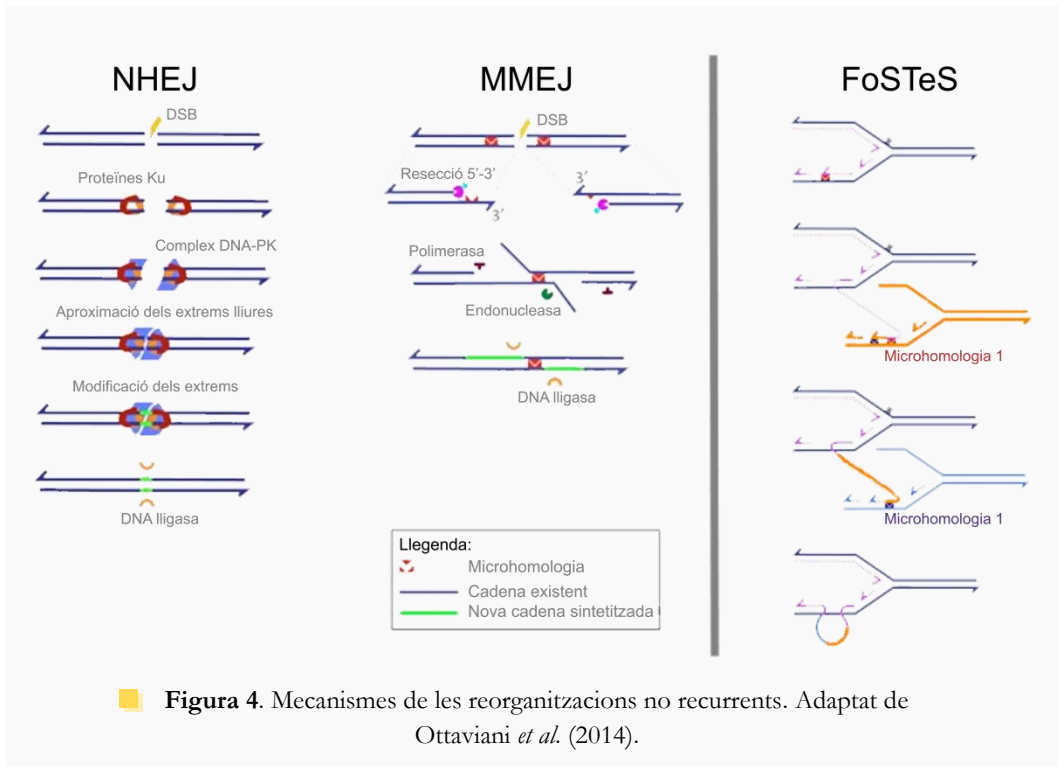
El procés de la NHEJ es pot dividir en quatre fases. Primerament es detecta la presència del DSB, a continuació es genera un complex de proteïnes pont entre els extrems lliures no complementaris, es modifiquen els extrems de manera que es puguin cohesionar (per addició o deleció de nucleòtids) i, finalment, s'enllacen. Aquest procés pot generar diferents reorganitzacions estructurals en funció de la localització original dels extrems units. Si els extrems units es situen en el mateix cromosoma es produiran



fonamentalment delecions, mentre que si es situen en cromosomes diferents es podran originar translocacions (STANKIEWICZ I LUPSKI 2010; DAVIS I CHEN 2013).

La FoSTeS també juga un paper rellevant en la generació de reorganitzacions no recurrents i, particularment, en les reorganitzacions complexes i la cromotripsia. Igual que la NHEJ, tot i ser un mecanisme independent de MEPS i LCR, la seva incidència també està influenciada per l'arquitectura genòmica local i per microhomologies. El

mecanisme de FoSTeS es fonamenta en errors que tenen lloc en l'etapa de replicació del DNA. Si durant la replicació del DNA es produeix una interrupció de l'activitat d'una forca, els extrems lliures poden unir-se a les cadenes d'una altra forca de replicació pròxima i amb certa homologia. La resolució d'aquests productes intermediaris generarà pèrdues o guanys de material genètic (STANKIEWICZ I LUPSKI 2010; LEE ET AL. 2007).



## 1.4 ESTUDI DE LA VARIABILITAT ESTRUCTURAL

El descobriment de les reorganitzacions estructurals està lligat a l'evolució de les tècniques d'anàlisi del genoma. Entre el conjunt de tècniques emprades en l'anàlisi de la variabilitat estructural podem diferenciar tècniques d'anàlisi global i tècniques dirigides a regions específiques (**Taula 1**).

■ **Taula 1.** Metodologies d'estudi de la variabilitat estructural del genoma.

		Translocacions	Inversions	CNV
<b>Anàlisi del genoma complet</b>				
Cariotip		● (>3 Mb)	● (>3 Mb)	● (>3 Mb)
M-FISH		● (>3 Mb)	●	● (>3 Mb)
CGH	Cromosomes metafàsics	●	●	● (>3 Mb)
	BAC	●	●	● (>50 kb)
	Oligonucleòtids	●	●	● (>35 kb)
SNP-array		●	●	●
Seqüenciació d'alt rendiment		●	●	●
<b>Anàlisi de regions específiques</b>				
PCR		●	●	● (>50 pb)
FISH	Cromosomes metafàsics	● (>3 Mb)	● (>3 Mb)	● (>3 Mb)
	Nuclis interfàsics	● (>50 kb)	● (>50 kb)	● (>50 kb)
	Fibres de cromatina	● (>10 kb)	● (>10 kb)	● (>10 kb)

● Aplicable; ● No aplicable per reorganitzacions equilibrades

CNV = *Copy Number Variation*; M-FISH = *Multiplex Fluorescence In situ Hybridization*; CGH = *Comparative Genome Hybridization*; BAC = *Bacterial Artificial Chromosome*; SNP-array = *Single Nucleotide Polymorphisms array*; PCR = *Polymerase Chain Reaction*; FISH = *Fluorescence In situ Hybridization*

## ANÀLISI GENÒMICA

L'anàlisi de la variabilitat genòmica es pot realitzar amb tècniques citogenètiques com el cariotip amb bandes G. L'ordenació dels cromosomes d'una cèl·lula en metafase permet detectar variacions estructurals com delecions, duplicacions, translocacions, insercions i inversions. Aquesta tècnica és eficient en la detecció de variacions d'una mida igual o superior a 3Mb que provoquin variacions rellevants de la morfologia cromosòmica o canvis en el patró de bandes (FEUK ET AL. 2006).

Les metodologies basades en la hibridació *in situ* fluorescent (FISH de l'anglès *Fluorescence In Situ Hybridization*) consisteixen en hibridar, mitjançant la complementarietat de bases, fragments de DNA marcats amb fluorescència sobre mostres problema per detectar seqüències concretes. Tot i que, en general, aquestes tècniques s'utilitzen en l'anàlisi de regions específiques, cal puntualitzar que existeix una varietat de FISH anomenada M-FISH (*Multiplex FISH*) que possibilita l'anàlisi de tot el

genoma mitjançant sondes de pintat cromosòmic (SARRATE ET AL. 2004). La M-FISH és especialment eficient en la detecció de reorganitzacions inter-cromosòmiques. El límit de detecció de reorganitzacions d'aquesta tècnica és similar al del cariotip amb 3 Mb (AZOFEIFA ET AL. 2000).

La tècnica de la hibridació genòmica comparada (CGH de l'anglès *Comparative Genome Hybridization*) consisteix en mesurar les diferències de fluorescència entre dos genomes distintivament marcats (un de referència i un de problema) per tal d'analitzar guanys i pèrdues de seqüències. Quan s'aplica sobre cromosomes metafàsics la resolució obtinguda és de 3 – 10 Mb. D'altra banda, la CGH en xips de DNA o *arrays* (*array-CGH*) utilitza fragments marcats d'un genoma problema que s'hibriden sobre fragments d'un genoma de referència marcat diferencialment. L'*array-CGH* constitueix una de les tècniques més robustes en l'anàlisi de CNV. Els fragments de referència poden ser BAC (de l'anglès *Bacterial Artificial Chromosome*) (50-150 kb) o oligonucleòtids (60-100 bp); com més petits són aquests fragments més precisió s'obté en l'anàlisi (FEUK ET AL. 2006). Així doncs, l'*array-CGH* permet la identificació de variacions del número de còpies amb un grau de resolució molt més elevat que el cariotip amb bandes G. El principal desavantatge d'aquesta tècnica és que no permet el diagnòstic de reorganitzacions equilibrades (COUGHLIN ET AL. 2012).

A nivell genòmic, també cal destacar l'ús de *SNP-arrays* en l'estudi dels CNV. Encara que, els *SNP-arrays* van ser dissenyats inicialment per la detecció de variacions d'un sol nucleòtid, a la pràctica també s'han aplicat amb èxit en l'anàlisi de guanys i pèrdues de seqüències (ATEN ET AL. 2008; REDON ET AL. 2006). A diferència de l'*array-CGH*, en els *SNP-arrays* la mostra problema, sense un marcatge diferencial, s'hibrida sobre un conjunt d'oligonucleòtids marcats amb fluorocroms. L'avaluació dels CNV es realitza comparant la intensitat de la fluorescència emesa per la hibridació de la mostra problema amb la valoració de la intensitat de fluorescència en mostres de controls. En l'aplicació dels *SNP-arrays* per l'anàlisi de la variabilitat estructural, les limitacions són les mateixes que per l'*array-CGH* (COUGHLIN ET AL. 2012).

Els mètodes de seqüenciació de segona generació han començat a substituir la tècnica de l'*array-CGH* com a mètode preferent per l'estudi de CNV a nivell de tot el genoma. Aquestes tècniques es basen en la generació de milers de reaccions de seqüenciació en

paral·lel i es denominen seqüenciacions d'alt rendiment (en anglès *High-throughput sequencing* o *Next generation sequencing*). La propagació de les tècniques de seqüenciació entre els estudis sobre la variabilitat estructural es deu a què, amb un cost econòmic raonable, en comparació a les tècniques de seqüenciació del genoma de primera generació (*Hierarchical shotgun sequencing* i *Whole-genome shotgun sequencing*), ofereixen una major resolució, una estimació més acurada de les variacions de còpies, més precisió en la detecció dels punts de trencament i una major capacitat d'identificació de nous CNV (ZARREI ET AL. 2015; PARKS ET AL. 2015; KIDD ET AL. 2008). A més a més, aquestes tècniques, a diferència de l'*array-CGH*, sí que permeten la detecció de variacions estructurals equilibrades. L'augment en la demanda de les tècniques de seqüenciació d'alt rendiment per l'anàlisi de CNV ha comportat també la instauració de múltiples mètodes computacionals i eines que faciliten aquest tipus d'estudis (ZHAO ET AL. 2013). No obstant, aquestes metodologies també mostren algunes limitacions per l'anàlisi de la variabilitat estructural. En primer lloc, algunes seqüències són difícils de seqüenciar degut a la interferència entre els senyals veritables amb els inespecífics. D'altra banda, hi ha mostres que presenten distribucions anormals de la cobertura de seqüenciació cosa que incrementa la detecció de falsos positius per l'estudi de CNV. I finalment en l'anàlisi de seqüències altament polimòrfiques o repetitives, com els LCR, es tendeix a subestimar la seva distribució (RETTNER ET AL. 2015; ALKAN ET AL. 2010).

---

## ANÀLISI DE REGIONS ESPECÍFIQUES

Per tal d'analitzar la variabilitat estructural de regions específiques del genoma, les tècniques més robustes són la FISH i les tècniques que es basen en la reacció en cadena de la polimerasa (PCR de l'anglès *Polymerase Chain Reaction*). La tècnica de *Southern Blot* també pot ser utilitzada per detectar variacions de regions específiques però, degut a la dificultat de l'anàlisi quantitatiu dels resultats obtinguts, ha tingut una aplicació minoritària en aquest context (ATEN ET AL. 2008). Les tècniques dirigides sovint s'utilitzen per validar els resultats de les tècniques que analitzen tot el genoma ja que analitzen un nombre limitat de seqüències en un mateix assaig.

Les metodologies de FISH proporcionen una resolució de 3 Mb en hibridacions sobre cromosomes metafàsics i fins a ~50 kb en hibridacions sobre nuclis interfàsics. En



conjunt, les tècniques basades en la FISH possibiliten la detecció d'un gran nombre de reorganitzacions estructurals com translocacions, inversions, anells acèntrics, delecions i duplicacions, entre altres. La resolució de la FISH en nuclis interfàsics està limitada pel grau de condensació del DNA que s'ha estimat com a 100 vegades superior al DNA relaxat i que dificulta l'ordenació espacial de les seqüències marcades. Aquesta limitació ha pogut ser superada amb l'optimització de la FISH sobre fibres de cromatina estirades. La hibridació sobre fibres de cromatina facilita el mapatge de segments de DNA i augmenta la resolució fins a 10 – 20 kb (HENG ET AL. 2005). Aquesta darrera tipologia de FISH ha estat anomenada FISH d'alta resolució sobre fibres de cromatina (*fiber-FISH*).

Els mètodes basats en la PCR que faciliten la quantificació de les seqüències amplificades, com la PCR quantitativa en temps real (Q-PCR) i la ddPCR (de l'anglès *droplet digital PCR*), permeten l'anàlisi de CNV amb una resolució de fins a ~50 pb. La limitació en el nombre de seqüències analitzades simultàniament s'ha pogut millorar amb el disseny de tècniques basades en la PCR que realitzen anàlisis simultanis de múltiples seqüències com la QMPSF (de l'anglès *Quantitative Multiplex PCR of Short Fluorescent fragments*), la MAPH (de l'anglès *Multiplex Amplifiable Probe Hybridization*) o la MLPA (de l'anglès *Multiplex Ligation-dependent Probe Amplification*) (ATEN ET AL. 2008; FEUK ET AL. 2006). Tot i això, una limitació destacable és la dificultat d'anàlisi de reorganitzacions equilibrades.

## 1.5 CONTRIBUCIÓ DE LES REORGANITZACIONS ESTRUCTURALS AL GENOMA

Les reorganitzacions estructurals poden contribuir a la creació de nous fenotips a través del canvi del número de còpies de gens sensibles a la dosi, la interrupció de gens, la fusió de gens, els defectes de la regulació dels gens per canvis de posició o, fins i tot, el desemmascament de gens o mutacions recessives (EICHLER ET AL. 2007). Per aquest motiu, la generació de reorganitzacions estructurals mitjançant la NAHR entre LCR ha tingut un paper destacable en l'evolució del genoma humà, en l'increment de la variabilitat genètica entre individus i ha constituït l'origen d'un grup de malalties molt significatives en humans denominades trastorns genòmics.

## EVOLUCIÓ

Les variacions estructurals han esdevingut fonamentals en els processos d'especiació (PERRY ET AL. 2008). Diferents estudis sobre el genoma humà han posat de manifest una associació significativa entre la localització de regions d'inestabilitat cromosòmica (punts de trencament evolutius) i la presència de LCR (MARQUES-BONET ET AL. 2009A; BAILEY I EICHLER 2006; ARMENGOL ET AL. 2003). Fins i tot, s'ha determinat que 6 de les 9 inversions que distingeixen el cariotip humà del dels ximpanzés presenten punts de trencament en regions amb LCR (KEHRER-SAWATZKI I COOPER 2007).

Els LCR són una característica comuna en molts genomes, tot i així, en el genoma humà i en el d'altres primats se'n ha detectat un increment significatiu en comparació a altres espècies. Aquesta comparació ha evidenciat la rellevància dels LCR com a mecanisme evolutiu capaç de crear nous fenotips i d'introduir més variabilitat genètica. En els primats, els LCR amb un 90% d'identitat corresponen a duplicacions esdevingudes fa 35-40 milions d'anys, coincidint amb la divergència dels grans simis del Vell i del Nou món. Durant l'etapa en la qual es produí aquesta expansió dels LCR altres processos mutacionals s'alentiren, la qual cosa reforça el seu paper destacat com a mecanisme evolutiu (BAILEY I EICHLER 2006). Com a prova de l'acció d'aquest mecanisme, les regions amb LCR es correlacionen amb la localització de CNV i amb regions amb una alta densitat d'exons (REDON ET AL. 2006). A més a més, aquesta correlació s'intensifica en les regions que corresponen a gens relacionats amb la percepció sensorial i la resposta immunològica i que es consideren responsables de canvis adaptatius ràpids (COOPER ET AL. 2007).

## POLIMORFISMES

Els CNV contribueixen de forma rellevant en la generació de variabilitat genètica entre els humans així com en l'explicació de les diferències fenotípiques observades. Per tal d'aprofundir el coneixement d'aquest tipus de variacions entre les diferents poblacions humanes, l'any 2002 es va posar en marxa el projecte HapMap. Aquest projecte determinà els patrons comuns de la variabilitat genètica humana i establí mapes d'haplotips que han sigut emprats per la identificació de polimorfismes relacionats amb

salut, malalties i diferències en la resposta a fàrmacs o a factors ambientals (THE INTERNATIONAL HAPMAP CONSORTIUM 2005).

En l'estudi dels perfils genòmics de variació estructural entre la població europea, asiàtica i africana s'ha observat com els CNV es focalitzen en *loci* relacionats amb l'adaptació ambiental (resposta immunològica, metabolisme lipídic i espai extracel·lular) (ARMENGOL ET AL. 2009). Alguns d'aquests polimorfismes s'han proposat com a factors de susceptibilitat per diverses malalties entre les quals destaquen les que afecten al sistema immunitari (**Taula 2**). En aquest sentit, un dels exemples més rellevants és l'associació de l'increment de còpies del gen *CCL3L1* (OMIM 601395) amb una reducció significativa en la susceptibilitat a la infecció per VIH i a la progressió de la SIDA (GONZALEZ ET AL. 2005). Així mateix, increments del nombre de còpies del gen *FCGR3B* (OMIM 610665) del receptor d'immunoglobulines Fc s'han relacionat amb una davallada del risc de la nefropatia lúpica (AITMAN ET AL. 2006). D'altra banda, també s'han descrit alguns CNV que actuen

■ **Taula 2.** Polimorfismes de DNA relacionats amb susceptibilitat a malalties. Adaptada de Stankiewicz i Lupski (2010).

Malaltia	Gen / locus afectat (OMIM <sup>a</sup> )
Psoriasi	<i>DEFB</i> (615997)
Emfisema	5p15.33, 7q22.1
Infart agut de miocardi i atac de cor (fumadors)	<i>C4B</i> (120820)
Displàsia capil·lar alveolar	<i>FOXF1</i> (601089)
Hipertricosi	17q24.2-q24.3
Pancreatitis crònica	<i>PRSS1</i> (276000)
Glomerulonefritis lúpica	<i>FCGR3B</i> (610665)
Malaltia de Crohn	<i>HBD-2</i> (142000), <i>IRGM</i> (608212)
Hipercolesterolèmia familiar	<i>LDLR</i> (606945)
Arteriosclerosi	<i>LPA</i> (152200)
Hemofília A	<i>Factor VIII</i> (300841)
Susceptibilitat a la malària	<i>α-globina</i> (141800)
Lupus eritromatós sistèmic	<i>C4B</i> (120820), <i>FCGR3B</i> (610665)
Susceptibilitat VIH	<i>CCL3L1</i> (601395)

<sup>a</sup> Online Mendelian Inheritance in Man; <http://www.ncbi.nlm.nih.gov/omim>

com a factors de risc de malalties complexes com l'Alzheimer, el Parkinson, l'esquizofrènia, l'epilèpsia o l'autisme (STANKIEWICZ I LUPSKI 2010; MERIKANGAS ET AL. 2014).

Arran d'aquesta nova perspectiva sobre la variabilitat genètica poblacional, també hi ha hagut un creixent interès en la recerca de polimorfismes que expliquen diferències fenotípiques entre poblacions relacionades amb fenòmens adaptatius a l'entorn. Per exemple, un estudi en tibetans ha determinat una sèrie de gens amb CNV que podrien explicar els trets adaptatius d'aquesta població a les altes alçades com l'elevat pes al néixer i el to fosc de pell (ZHANG ET AL. 2012). També és destacable la relació observada entre les poblacions amb una elevada ingesta de midó en la seva dieta i l'elevat nombre de còpies del gen de l'amilasa salival *AMY1* (OMIM 104700). El nombre de còpies del gen *AMY1* es correlaciona directament amb la concentració enzimàtica de l'amilasa salival i, per tant, en les poblacions amb més ingesta de midó l'augment de còpies de *AMY1* resulta favorable per facilitar la digestió (PERRY ET AL. 2007).

---

## TRASTORNS GENÒMICS

Les reorganitzacions estructurals, com s'ha descrit en els apartats anteriors, poden ser aparentment innòcues, donar lloc a fenotips evolutivament avantatjosos o constituir un factor de risc per certes malalties, però també poden originar fenotips directament patològics. Les malalties produïdes per reorganitzacions estructurals en regions inestables del genoma flanquejades per LCR i que afecten a gens sensibles a dosi s'anomenen trastorns genòmics (STANKIEWICZ I LUPSKI 2002).

Majoritàriament aquestes reorganitzacions es produeixen per processos meiòtics de NAHR que, com s'ha explicat anteriorment (apartat 3.1), donen lloc a gàmetes portadores de reorganitzacions que podran ser transmeses a la descendència i causar els trastorns genòmics. Tot i que la majoria de reorganitzacions són de tipus recurrent, també podem distingir trastorns genòmics amb una mida i uns punts de trencament variables (GU ET AL. 2008). Alguns dels trastorns genòmics recurrents més freqüents es troben recollits en la **Taula 3**. Aquest projecte de tesi es centra en l'estudi de la síndrome de DiGeorge/velocardiofacial (SDG/VCF) que està causada per una deleció en la regió 22q11.2.

Degut a l'elevat nombre de trastorns genòmics descrits, s'han desenvolupat diverses bases de dades que recopilen informació de les reorganitzacions patològiques identificades. Per exemple, el 2004 es va constituir la base de dades DECHIPHER (de l'anglès *DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources*, <https://decipher.sanger.ac.uk/>) per tal de recopilar les reorganitzacions identificades en pacients procedents de diferents centres (europeus, americans i alguns asiàtics). La finalitat d'aquests projectes rau en determinar els efectes fenotípics dels CNV i discriminar els polimorfismes benignes i els patològics.

■ **Taula 3.** Trastorns genòmics recurrents causats per la NAHR entre LCR.

Trastorn genòmic	OMIM <sup>a</sup>	Anomalia	Regió	Mida (Mb) <sup>b</sup>
Síndrome de Sotos	117550	Deleció	5q35	2,2
Síndrome de Williams-Beuren	194050	Deleció	7q11.23	1,55
Neuropatia hereditària amb paràlisi sensible a la pressió	162500	Deleció	17p12	1,5
Charcot-Marie-Tooth tipus 1A	118220	Duplicació	17p12	1,5
Síndrome de Smith-Magenis	182290	Deleció	17p11.2	5
Síndrome de Potocki-Lupski	610883	Duplicació	17p11.2	5
Neurofibromatosi tipus 1	162200	Deleció	17p11.2	1,5
Síndrome de Prader Willi /Angelman	176270/ 105830	Deleció	15q11q13	4
Síndrome de Duplicació 15	608636	Duplicació	15q11q13	4
Síndrome de DiGeorge/velocardiofacial	188400/ 192430	Deleció	22q11.2	3
Síndrome de l'ull de gat	115470	Duplicació invertida	22q11.2	3
Ictiosi lligada al X	308100	Deleció	Xp22	1,9

<sup>a</sup> Online Mendelian Inheritance in Man; <http://www.ncbi.nlm.nih.gov/omim>

<sup>b</sup> Mida majoritària de l'anomalia.

## 1.6 RISC DE RECURRÈNCIA DELS TRASTORNS GENÒMICS

En general, en les famílies en les quals es diagnostica un descendent afectat per un trastorn genòmic, el risc de recurrència es considera del ~0,1% (RÖTHLISBERGER I KOTZOT 2007; CAMPBELL ET AL. 2014). Tanmateix, l'estudi del risc de recurrència pels trastorns genòmics és complex degut a diferents factors (RÖTHLISBERGER I KOTZOT

2007): la incidència d'aquestes malalties és molt baixa, és possible que no s'hagin publicat tots els casos de recurrència i les famílies amb un fill afecte poden plantejar-se no tenir més fills.

Els casos de recurrència descrits s'han atribuït a la presència de mosaïcismes germinals; tot i que, en la majoria de casos, aquesta hipòtesi no ha pogut ser verificada (**Taula 4**). En alguns casos la presència d'un mosaïcisme somàtic ha permès extrapolar amb major certesa la presència d'un mosaïcisme germinal (SÁNCHEZ ET AL. 2013; CAMPBELL ET AL. 2014; KOOLEN ET AL. 2012). Davant la confirmació d'un mosaïcisme somàtic, el risc de recurrència s'ha considerat superior al de la població general. Tanmateix, no s'ha pogut establir en quin grau es produeix aquest risc degut a la desconexença de la proporció de gàmetes portadores de reorganitzacions.

■ **Taula 4.** Casos de recurrència en trastorns genòmics.

Regió	Síndrome	Origen parental	Referència
5p12.2	Cri-du-Chat	Matern	Hajianpour et al. 1991
7q11.23	Williams-Beuren	Matern	Kara-Mostefa et al. 1999
7q11.23	Williams-Beuren	Patern	Scherer et al. 2005
15q11q13	Angelman	Matern	Kokkonen i Leisti 2000
15q11q13	Angelman	Matern	Sánchez et al. 2013
15q11q13	Prader-Willi	Patern	Fernandez-Novoa et al. 2001
17p11.2	Smith-Magenis	Matern	Campbell et al. 2014
17q21.31	Deleció 17q21.31	Matern	Koolen et al. 2012
22q11.2	DiGeorge/velocardiofacial	Matern	Hatchwell et al. 1998
22q11.2	DiGeorge/velocardiofacial	Matern	Sandrin-Garcia et al. 2002
22q11.2	DiGeorge/velocardiofacial Duplicació 22q11.2	Matern	Demaerel et al. 2016

Un complement als estudis de recurrència poblacional l'han aportat els estudis cromosòmics en gàmetes, els quals permeten determinar la freqüència en què un determinat individu està generant anomalies relacionades amb un determinat trastorn genòmic. No obstant, en els casos en els quals l'origen del trastorn genòmic és matern, aquest estudi és complex de realitzar degut a limitacions metodològiques, legals i

ètiques. En canvi, en els casos d'origen patern l'estudi de les gàmetes és assolible ja que les mostres d'espermatozoides són fàcils d'obtenir i permeten l'anàlisi simultani de milers de cèl·lules, cosa que facilita la detecció de reorganitzacions poc freqüents amb una elevada precisió.

En aquest sentit, resultats previs del nostre grup de recerca, obtinguts aplicant tècniques de FISH sobre nuclis descondensats d'espermatozoides, mostraren un increment significatiu de delecions i duplicacions de la regió 15q11q13 en un 60% dels pares transmissors de la síndrome de Prader-Willi (MOLINA ET AL. 2010). La baixa freqüència de delecions i duplicacions va descartar el mosaïcisme germinal i va fer plantejar un increment de la susceptibilitat a la NAHR en aquests individus com a causa de l'increment de reorganitzacions respecte la població control. En la mateixa població de pares Prader-Willi, s'analitzà la freqüència d'inversions de la regió 15q11q13 i es determinà que un 40% dels individus que havien presentat un increment de delecions i duplicacions també mostraven una freqüència incrementada d'inversions de la regió crítica (MOLINA ET AL. 2012A). Els resultats declinaren novament la possibilitat d'un mosaïcisme germinal de la inversió de la regió crítica i confirmaren la susceptibilitat a la NAHR en alguns individus.

Per tant, els estudis en espermatozoides han determinat que alguns pares transmissors de síndromes per delecions presenten una susceptibilitat major a partir esdeveniments de NAHR i, en conseqüència, un risc de recurrència incrementat respecte a la població general. L'increment de susceptibilitat a la NAHR en alguns individus podria estar causat per la presència de factors de predisposició que incentivarien la formació d'aparellaments no al·lèlics o que incrementarien la freqüència de DSB en les regions crítiques.

## 1.7 FACTORS DE PREDISPOSICIÓ DELS TRASTORNS GENÒMICS

S'ha demostrat que hi ha característiques genètiques que incrementen la susceptibilitat a la NAHR. Entre les característiques estudiades destaquen les variacions de l'arquitectura genòmica de la regió crítica. Aquestes variacions inclouen CNV en els LCR involucrats i inversions en heterozigosi de la regió crítica. D'altra banda, també s'ha proposat que

certes variants al·lèliques de la proteïna de recombinació meiótica PRDM9 confereixen una major susceptibilitat a la NAHR.

## VARIABILITAT ESTRUCTURAL DELS LCR

Alguns autors han descrit variacions estructurals en els LCR, implicats en trastorns genòmics, que han estat relacionades amb una major susceptibilitat a l'aparellament no al·lèlic d'aquestes seqüències i, per tant, amb la incentivació dels esdeveniments de NAHR. Malgrat l'existència de nombroses regions crítiques del genoma, tan sols s'han pogut determinar arquitectures genòmiques de risc pels LCR d'algunes regions.

La regió 17p11.2 conté repeticions de dos tipus de LCR denominats REPA i REPB que es disposen tan en orientació directa com indirecta de manera que confereixen una elevada inestabilitat genòmica (BARBOUTI ET AL. 2004). Carvalho i Lupski (2008) descobriren CNV d'aquests LCR que s'observen amb major freqüència en individus afectats per la síndrome de Smith-Magenis, originada per una deleció 17p11.2. D'acord amb els resultats que obtingueren, determinaren que la susceptibilitat de les múltiples reorganitzacions amb punts de trencament en la regió 17p11.2 es troba influïda per l'arquitectura genòmica dels LCR REPA i REPB (CARVALHO I LUPSKI 2008).

De forma similar, Antonacci et al. (2010) estudiaren les seqüències dels LCR implicats en la deleció 16p12.1 que dona lloc a una malaltia neurocognitiva. Observaren que els LCR flanquejants de la regió crítica mostraven dues possibles arquitectures genòmiques que divergien en la orientació i la repetició d'alguns blocs de LCR i que anomenaren S1 i S2. La configuració S2, amb blocs de LCR orientats directament en ambdós LCR flanquejant de la regió crítica, es trobà més representada en els pacients amb la deleció 16p12.1; mentre que la configuració S1 protegia contra aquesta reorganització. Per tant, la configuració S2 predisposa a les delecions 16p12.1 incrementant l'homologia entre els LCR flanquejants i facilitant modes d'aparellament que originen esdeveniments de NAHR (ANTONACCI ET AL. 2010).

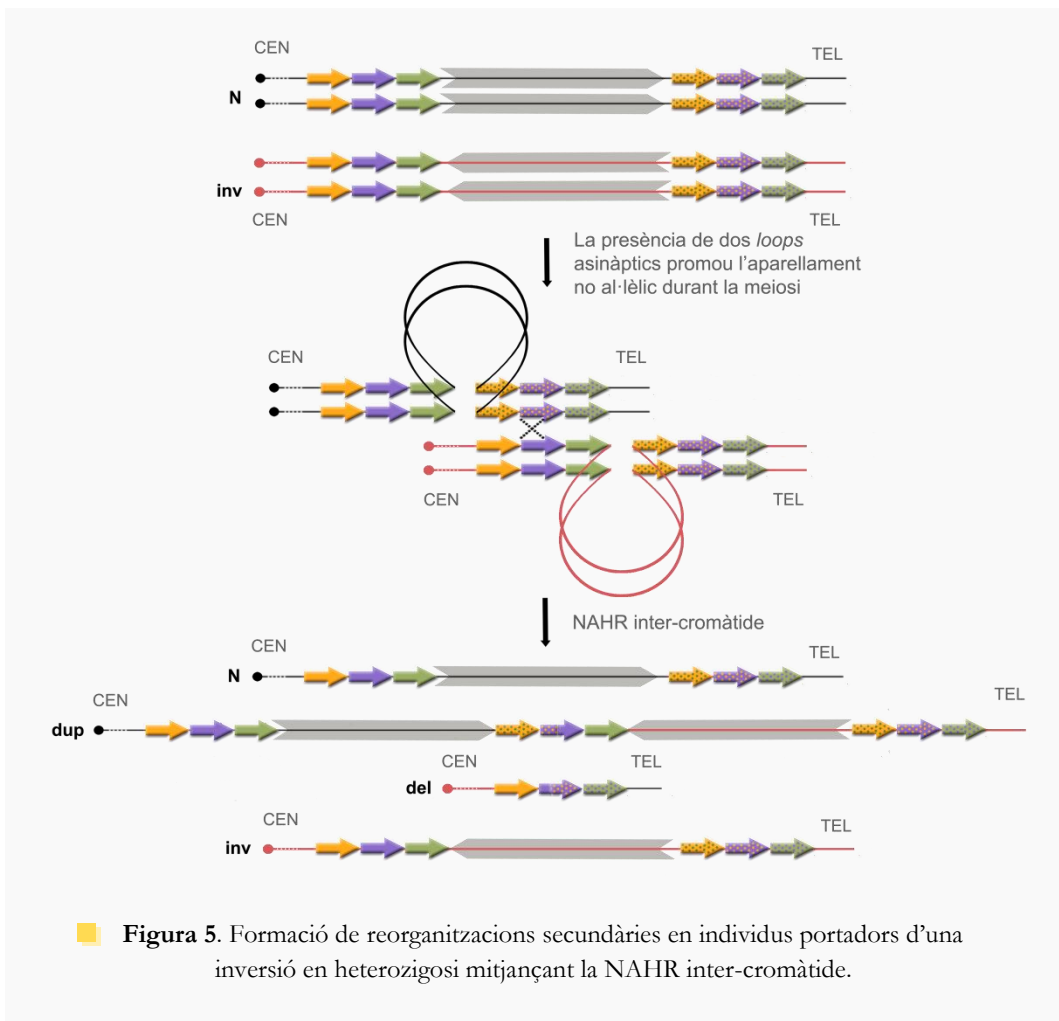
D'altra banda, Cuscó et al (2008) analitzaren les còpies de gens i pseudogens localitzades en els LCR de la regió 7q11.23. Observaren com en progenitors d'individus afectats per la síndrome de Williams-Beuren, causada per una deleció 7q11.23, hi havia un percentatge significativament més elevat de delecions i



duplicacions del gen i pseudogen *NCF1* (OMIM 608512). En conseqüència, es suggereix que aquests CNV es relacionen amb una probabilitat superior de NAHR i incrementen el risc de transmetre la síndrome de Williams-Beuren a la descendència dels individus portadors (CUSCÓ ET AL. 2008).

### INVERSIÓ DE LA REGIÓ CRÍTICA

S'han descrit inversions en individus amb descendència afectada com a factor de predisposició de reorganitzacions secundàries (**Figura 5**). Les inversions en heterozigosi de la regió crítica han estat descrites en les síndromes de Williams-Beuren



(SCHERER ET AL. 2005; HOBART ET AL. 2010; OSBORNE ET AL. 2001; BAYÉS ET AL. 2003), Angelman (GIMELLI ET AL. 2003), Smith-Magenis (YANG ET AL. 1997) i la deleció 17q21.31 (RAO ET AL. 2010), entre altres (ANTONACCI ET AL. 2009).

La presència d'una inversió en heterozigosi genera una regió asinàptica entre el cromosoma portador de la inversió i el seu homòleg normal que incrementa la inestabilitat genòmica. Aquesta situació facilita l'aparellament no al·lèlic dels LCR durant la meiosi i la generació d'esdeveniments de NAHR que originen reorganitzacions secundàries (**Figura 5**). La mida de la inversió, el contingut de possibles punts de trencament i la presència de regions repetitives flanquejants, com els LCR, determinen la susceptibilitat de la regió invertida per generar reorganitzacions (TORCASHEVA I BORODIN 2010).

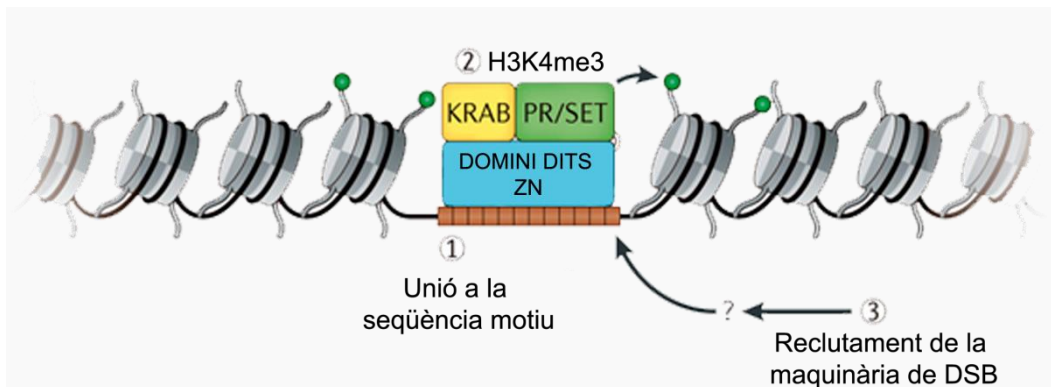
---

## GENOTIP DE *PRDM9*

Algunes proteïnes implicades en el procés de recombinació s'han associat amb la freqüència i l'estabilitat de la recombinació meiótica i, per tant, amb la susceptibilitat a patir un esdeveniment de NAHR (CHOWDHURY ET AL. 2009; PRATTO ET AL. 2014). *PRDM9* és una proteïna de dits de zinc que determina el patró de recombinació meiótica en el genoma de mamífers a través del reconeixement de seqüències motiu i del reclutament de la maquinària de generació DSB (**Figura 6**).

El gen de *PRDM9* (OMIM 609760) és altament polimòrfic amb desenes de variants al·lèliques descrites diferenciades pel número i la tipologia de dits de zinc de les seves seqüències. El número de repeticions varia entre 8 i 15 i s'han descrit 30 tipus de seqüències de dits de zinc que es distingeixen per canvis en alguns nucleòtids. L'al·lel A presenta 13 dits de zinc (ABCDDECFGHFIJ) i és el més abundant tan en la població europea com en l'africana amb una freqüència del 85,7% i del 50,4%, respectivament (BERG ET AL. 2010).

Els canvis de nucleòtids que s'introdueixen en les diferents variants al·lèliques de *PRDM9* canvien la seqüència consens reconeguda i, per tant, modifiquen el patró de distribució dels punts de trencament i recombinació del genoma (BAUDAT ET AL. 2010). En concret, l'al·lel A reconeix una seqüència consens degenerada de 13 nucleòtids, CCNCCNTNNCCNC, que s'ha detectat en aproximadament el 40% dels punts de



■ **Figura 6.** Mecanisme d'acció de PRDM9. PRDM9 s'uneix específicament a motius de DNA a través del domini de dits de zinc. A continuació, el domini PR/SET promou la trimetilació de la lisina 4 de les histones H3 dels nucleosomes adjacents (H3K4me3, esferes verdes). El domini KRAB (de l'anglès *Krippel-associated box*) interactua amb altres proteïnes. Aquests passos permeten el reclutament de la maquinària de formació dels DSB, incloent la proteïna de recombinació meiótica SPO11 a través d'un mecanisme desconegut. Adaptat de Baudat *et al.* (2013).

recombinació humans (MYERS ET AL. 2008). A partir del motiu de recombinació CCNCCNTNCCNC, s'han predit les seqüències consens reconegudes per altres al·lels de *PRDM9*. Alguns al·lels com B, L2, L3, L9, L10 i L24 reconeixen el mateix motiu de recombinació descrit per l'al·lel A. En canvi, en altres al·lels com el C, L1 i L20 les seqüències reconegudes mostren canvis en un o més nucleòtids (BERG ET AL. 2010).

A partir de la comparació de les freqüències al·lèliques en parets transmissors de trastorns genòmics i en individus control, alguns autors han determinat *PRDM9* com a un factor de predisposició per aquestes malalties. L'al·lel A s'ha descrit com a factor de risc per la síndrome de Charcot-Marie-Tooth tipus 1A, la neuropatia hereditària amb paràlisi sensible a la pressió i la síndrome de Williams-Beuren (BERG ET AL. 2010; BOREL ET AL. 2012). A més a més, l'anàlisi dels punts de trencament de les reorganitzacions causants de diferents trastorns genòmics demostra l'alta densitat de motius de recombinació de *PRDM9* en aquestes regions i dona suport a la consideració del genotip de *PRDM9* com a factor de predisposició (PRATTO ET AL. 2014).

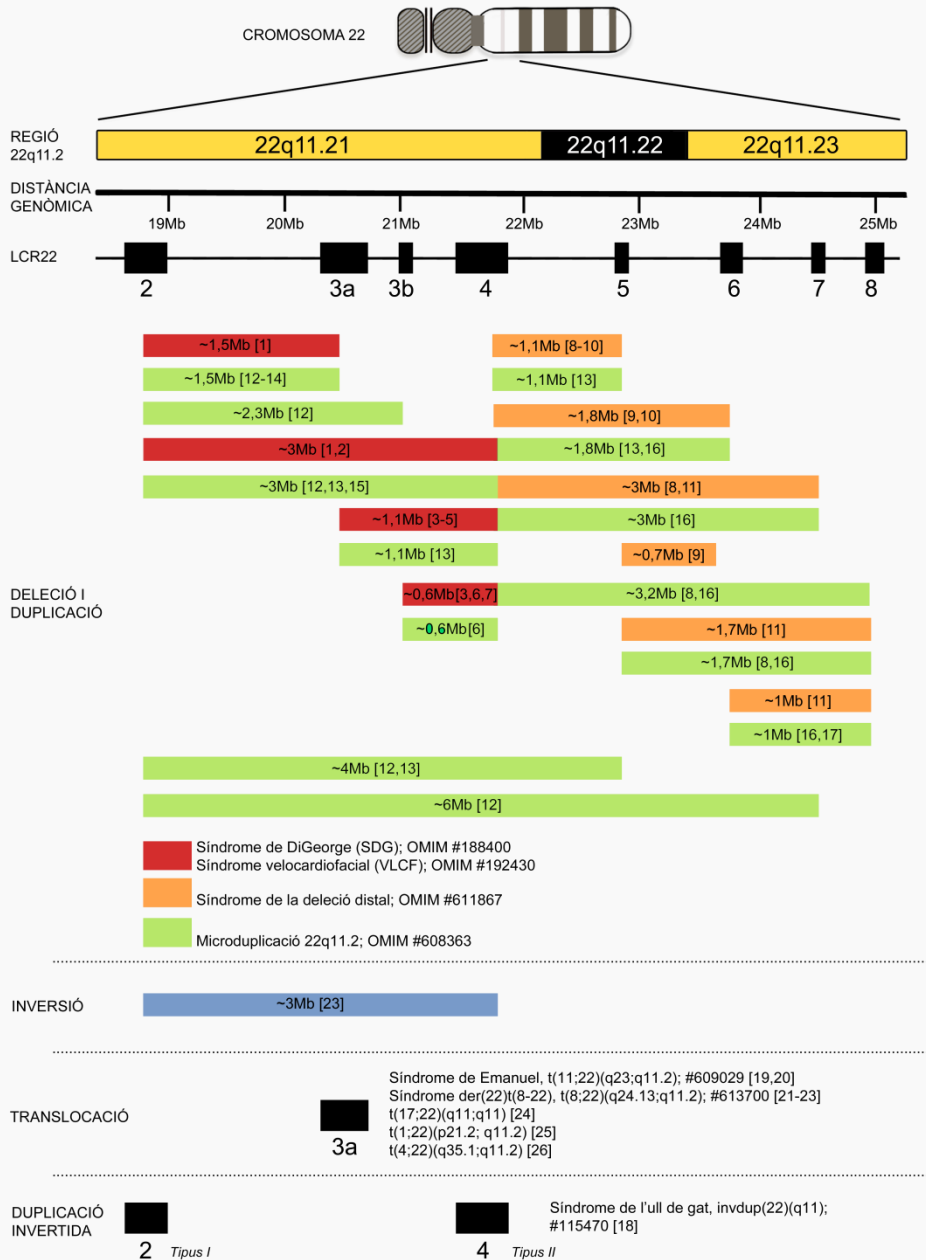
## 1.8 REGIÓ CROMOSÒMICA 22Q11.2

### REORGANITZACIONS DE LA REGIÓ 22Q11.2

La regió pericentromèrica del cromosoma 22 conté vuit LCR que comprenen aproximadament un 11% de la sub-banda 22q11.2. D'acord amb la seva posició respecte al centròmer, els LCR s'anomenen LCR22-2, LCR22-3a, LCR22-3b, LCR22-4, LCR22-5, LCR22-6, LCR22-7 i LCR22-8 (BABCOCK ET AL. 2007).

Degut a la seva complexitat, els LCR22 constitueixen un punt d'instabilitat genòmica que es reflexa en la seva implicació en un elevat nombre de reorganitzacions (**Figura 7**). El trastorn genòmic més freqüent és la síndrome de DiGeorge/velocardiofacial, que és deguda a una deleció de la regió 22q11.2. La deleció 22q11.2 constitueix una de les delecions més freqüents entre els humans amb una incidència estimada de 1:4000 nounats (BOTTO ET AL. 2003). En l'estudi de l'origen d'aquesta reorganització, diversos autors han apuntat un paper predominant de NAHR inter-cromosòmica (THOMAS ET AL. 2006; EDELMANN ET AL. 1999). Consistentment, s'ha observat que la freqüència de delecions i duplicacions de la regió 22q11.2 en espermatozoides és equivalent i que la majoria de les duplicacions són productes recíprocs a les delecions (MOLINA ET AL. 2011; OU ET AL. 2011). Aquestes dades, doncs, reforcen la NAHR inter-cromàtide o inter-cromosòmica com a mecanisme predominant de la generació de delecions i duplicacions 22q11.2. Tanmateix, la incidència dels trastorns genòmics que causen ambdues reorganitzacions no és equiparable ja que la síndrome de la duplicació 22q11.2 dona lloc a fenotips més lleus que la SDG/VCF. Aquest fet comporta que probablement la duplicació sigui infra-diagnosticada (ENSENAUER ET AL. 2003; PORTNOI 2009).

D'altra banda, a part de la Síndrome de DiGeorge/Velocardiofacial i les duplicacions complementàries, s'han descrit delecions i duplicacions que involucren als LCR22 distals (VERHAGEN ET AL. 2012; TAN ET AL. 2011; MIKHAIL ET AL. 2014), translocacions amb un punt de trencament en el LCR22-3a (KURAHASHI I EMANUEL 2001; SHERIDAN ET AL. 2010), inversions entre els LCR22-2 i LCR22-4 (GOTTER ET AL. 2007) i duplicacions invertides que afecten al LCR22-2 i al LCR22-4 (GENTILE ET AL. 2005). La



**Figura 7.** Reorganitzacions de la regió 22q11.2. [1] (CARLSON ET AL. 1997), [2] (EDELMAAN ET AL. 1999), [3] (VERHAGEN ET AL. 2012), [4] (GARCIA-MIÑAUER ET AL. 2002), [5] (RAUCH ET AL. 2005), [6] (FERNÁNDEZ ET AL. 2009), [7] (ANGELO ET AL. 2007), [8] (TAN ET AL. 2011), [9] (MIKHAIL ET AL. 2014), [10] (BEN-SHACHAR ET AL. 2008), [11] (JACKSON ET AL. 2007), [12] (ENSENAUER ET AL. 2003), [13] (OU ET AL. 2008), [14] (ALBERTI ET AL. 2007), [15] (EDELMAAN ET AL. 1999), [16] (COPPINGER ET AL. 2009), [17] (RIBEIRO-BICUDO ET AL. 2013), [18] (GENTILE ET AL. 2005), [19] (KURAHASHI I EMANUEL 2001), [20] (EDELMAAN ET AL. 2001), [21] (SHERIDAN ET AL. 2010), [22] (GIMELLI ET AL. 2009), [23] (GOTTER ET AL. 2007), [24] (KURAHASHI ET AL. 2003), [25] (GOTTER 2003), [26] (NIMMAKAYALU ET AL. 2003)

descripció d'una àmplia diversitat de reorganitzacions emfatitza l'elevada inestabilitat de la regió 22q11.2.

## SÍNDROME DE DIGEORGE/VELOCARDIOFACIAL

La deleció en hemizigosi de la regió 22q11.2 presenta un ampli ventall fenotípic. Les característiques clíniques atribuïdes a aquest genotip inclouen defectes cardiovasculars congènits, anomalies facials, anomalies palatals, deficiències immunitàries, disfunció paratiroidea amb hipocalcèmia i diversos trastorns neuropsiquiàtrics i discapacitats cognitives. L'expressió d'un fenotip que no és uniforme per la mateixa reorganització ha desembocat en l'aparició de diverses denominacions que descriuen la simptomatologia.

El 1965, el doctor Angelo DiGeorge va descriure un grup d'infants que presentaven una absència congènita del timus i de la glàndula paratiroide i va denominar aquesta síndrome com a síndrome de DiGeorge (OMIM 188400). D'altra banda, les anomalies facial, els defectes cardiovasculars i els retards en la parla s'englobaren com a síndrome velocardiofacial (OMIM 192430). La presència de diferents denominacions per una síndrome que correspon a una única causa ha creat molta confusió. Avui en dia, es recomana denominar comunament totes les síndromes que comparteixen la deleció del cromosoma 22 en hemizigosi com a síndrome de la deleció 22q11.2 (MCDONALD-MCGINN I SULLIVAN 2011). Tot i així, degut a l'ús freqüent de l'anotació DiGeorge/velocardiofacial en les publicacions científiques, al llarg d'aquest document farem referència al conjunt de síndromes que impliquen una deleció 22q11.2 com a síndrome de DiGeorge/velocardiofacial (SDG/VCF).

En més del 85% dels casos la deleció 22q11.2 està causada per esdeveniments de recombinació homòloga no al·lèlica que inclouen els LCR22-2 i LCR22-4 i que dona lloc a una deleció de 3 Mb (CARLSON ET AL. 1997). Al voltant del 7% de casos s'atribueixen a delecions menors de 1,5 Mb i flanquejades pels LCR22-2 i LCR22-3a. En la resta de casos es produeixen delecions minoritàries entre els LCR22 més distals (GARCIA-MIÑAUR ET AL. 2002; TAN ET AL. 2011; VERHAGEN ET AL. 2012).

La regió de 3 Mb delecionada conté 90 gens que inclouen 46 gens codificants per proteïnes, 7 micro-RNA, 10 RNA no codificants i 27 pseudogens (MCDONALD-

MCGINN ET AL. 2015). El gen més estudiat de la regió delecionada és *TBX1* (OMIM 602054). Es considera que l'haploinsuficiència d'aquest gen és responsable de cinc dels fenotips majoritaris observats en aquests pacients: anomalies facials, defectes cardiovasculars, hipoplàsia del timus, anomalies palatals i disfunció paratiroidea amb hipocalcèmia (YAGI ET AL. 2003).

Tanmateix, la variabilitat fenotípica observada entre els portadors de la delecio 22q11.2 suggereix la presència de gens modificadors del fenotip que podrien actuar en la via d'activitat dels gens delecionats en hemizigosi. Un dels fenotips amb penetrància incompleta més estudiats són els defectes cardiovasculars, que afecten a un 60-75% dels portadors de la delecio 22q11.2. S'ha demostrat que *TBX1* interacciona amb el gen *VEGF* (OMIM 192240), localitzat fora de la regió delecionada, i que certes variants al·lèliques del promotor d'aquest gen constitueixen un modificador del desenvolupament dels defectes cardiovasculars (STALMANS ET AL. 2003). D'altra banda, un altre gen d'interès de la regió delecionada és *DGCR8* (OMIM 609030) ja que codifica per la proteïna Pasha que participa en la biogènesi dels micro-RNA. Els micro-RNA regulen l'expressió gènica i, per tant, la interferència en la seva síntesi pot contribuir a l'expressió deficitària de gens fora de la regió 22q11.2 relacionats amb la simptomatologia associada a la SDG/VCF. S'ha demostrat que la desregulació de l'expressió dels micro-RNA es relaciona amb defectes neuronals, defectes cardíacs i anomalies de tiroïdes (SELLIER ET AL. 2014).

A més a més, la penetrància incompleta d'alguns fenotips també s'ha atribuït a la presència de CNV. S'ha suggerit que CNV en la regió 22q11.2 en gens localitzats tan fora com dins els LCR es relacionen amb els trastorns neuropsiquiàtrics diagnosticats en pacients SDG/VCF (HIROI ET AL. 2013). Recentment, s'han proposat CNV fora de la regió crítica 22q11.2 podrien contenir gens que modifiquen el risc de desenvolupar defectes cardíacs en alguns pacients (MLYNARSKI ET AL. 2016).

---

#### RISC DE RECURRENCIA I FACTORS DE PREDISPOSICIÓ EN LA SDG/VCF

El risc de recurrència per la SDG/VCF es considera pràcticament negligible, gairebé equivalent al risc de transmissió establert en la població general. Tanmateix, com en la resta de trastorns genòmics, es considera que podria ser més elevat en els casos de

mosaicismes germinals del progenitor transmissor (MCDONALD-MCGINN I ZACKAI 2008). Per la SDG/VCF s'han descrit alguns casos de recurrència relacionats amb un potencial mosaicisme parental (HATCHWELL ET AL. 1998; SANDRIN-GARCIA ET AL. 2002). Recentment, també s'ha publicat un cas de recurrència de reorganitzacions 22q11.2 en el qual un dels fills presentava la SDG/VCF mentre que l'altre mostrava una duplicació 22q11.2 complementària a la deleció del germà (DEMAEREL ET AL. 2016).

Pel que fa als factors de predisposició, es desconeix quines característiques genètiques podrien afavorir els esdeveniments de NAHR de forma específica en la regió 22q11.2. Malgrat que la variabilitat en aquesta regió ha estat poc estudiada, alguns autors han apuntat una possible relació entre la freqüència de reorganitzacions d'aquesta regió i variacions en l'arquitectura genòmica dels LCR22 (BAUMER 1998; TORRES-JUAN ET AL. 2007). D'altra banda, tot i que en altres trastorns genòmics s'han identificat inversions 22q11.2 en heterozigosi en el progenitors com a factor de predisposició a reorganitzacions, en la SDG/VCF no s'ha observat aquest polimorfisme (GEBHARDT ET AL. 2003). Per últim, el genotip de *PRDM9* també ha estat estudiat en progenitors SDG/VCF però no s'ha pogut establir una clara relació amb la susceptibilitat als esdeveniments de NAHR (BOREL ET AL. 2012; ALEMANY-SCHMIDT ET AL. 2014).








**OBJECTIUS**



L'objectiu principal del projecte de tesi és avaluar les característiques genètiques que modulen la susceptibilitat a la NAHR en individus transmissors de la SDG/VCF. Per tal d'assolir aquest objectiu principal es plantegen els següents objectius específics:

- I** Determinar la freqüència d'espermatozoides portadors de deleccions i duplicacions de la regió 22q11.2, i d'altres regions crítiques amb una arquitectura genòmica similar, en homes amb descendència afectada per la SDG/VCF.
- II** Caracteritzar l'arquitectura genòmica dels LCR22-2 i LCR22-4 en la població control i en individus transmissors de la SDG/VCF.
- III** Determinar la freqüència d'haplotips amb inversió 22q11.2 en la població control i en individus transmissors de la SDG/VCF.
- IV** Establir les freqüències al·lèliques del gen *PRDM9* en la població control i en individus transmissors de la SDG/VCF.
- V** Avaluar la rellevància dels factors de predisposició analitzats i la seva relació amb la susceptibilitat a la NAHR de la regió 22q11.2.





# MATERIAL I MÈTODES

- 3.1 MOSTRES BIOLÒGIQUES I  
PROCESSAMENT
- 3.2 TÈCNIQUES CITOGENÈTIQUES  
BASEDES EN LA HIBRIDACIÓ IN SITU  
FLUORESCENT (FISH)
- 3.3 TÈCNIQUES MOLECULARS BASEDES  
EN LA REACCIÓ EN CADENA DE LA  
POLIMERASA (PCR)



### 3.1 MOSTRES BIOLÒGIQUES I PROCESSAMENT

En el desenvolupament d'aquest projecte de tesi es van estudiar mostres biològiques procedents de 25 individus de famílies amb un membre afectat per la SDG/VCF ateses a l'Hospital Sant Joan de Déu de Barcelona: 9 homes i 8 dones amb descendència afectada i 8 nens afectats. També es van avaluar mostres biològiques de 28 individus sense descendència SDG/VCF.

Tots els individus van ser informats dels protocols d'estudi i van manifestar la voluntat de participar amb la signatura d'un consentiment informat.

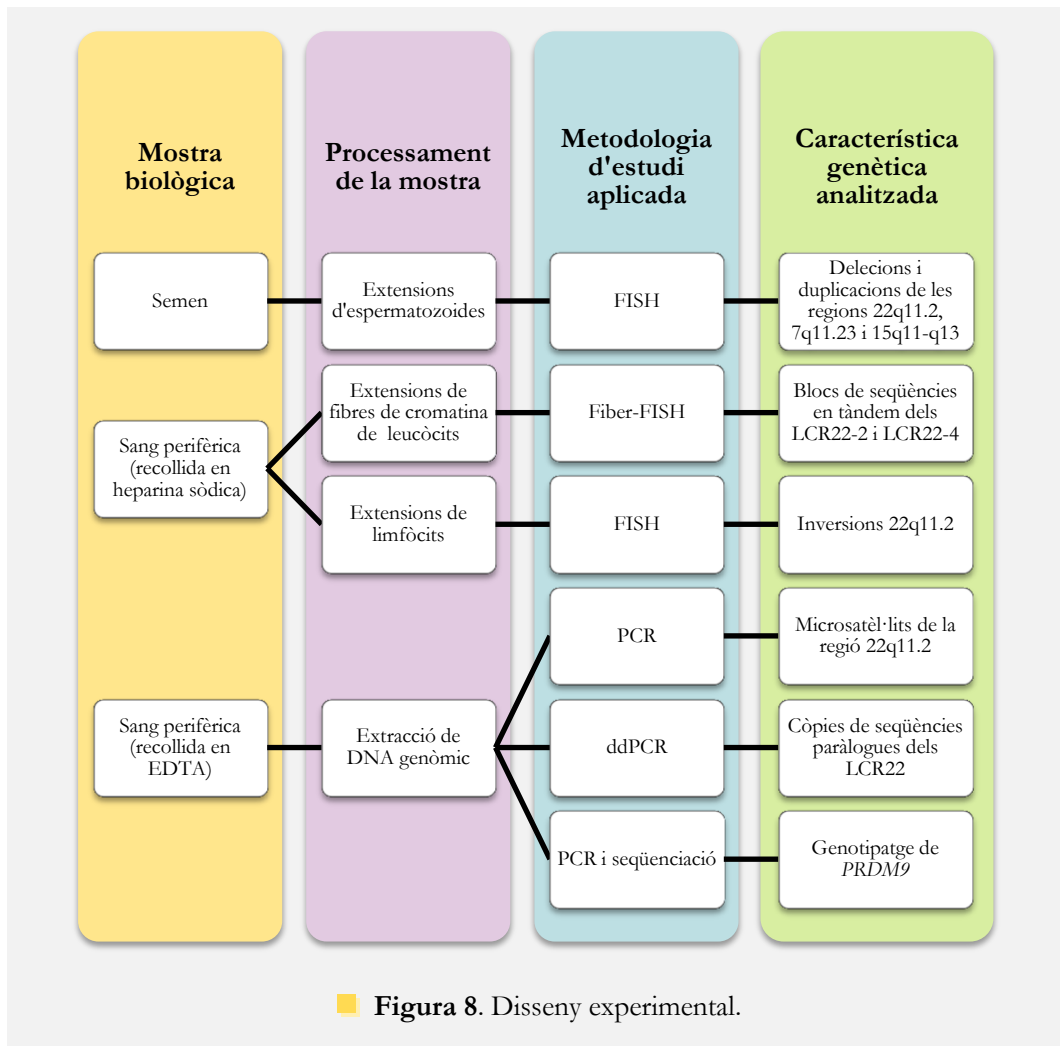
#### SEMEN

Les mostres de semen s'obtenen per masturbació en contenidors de polietilè i després d'un període mínim de tres dies d'abstinència. Un cop transferides a tubs de 10 ml, les mostres es centrifuguen a 1.000 g durant 5 minuts i es separa el plasma seminal de la fracció cel·lular mitjançant una pipeta Pasteur. La fracció cel·lular es tracta amb una solució hipotònica (KCl 0,075 M) prèviament escalfada a 37 °C i incorporada gota a gota. Es realitza una incubació a 37 °C durant 30 minuts i, després d'una altra centrifugació a 1.000 g durant 5 minuts, es decanta el sobrenedant i es resuspèn la mostra amb fixador de Carnoy (proporció 3:1 de metanol:àcid acètic) que s'afegeix gota a gota. Es repeteix tres vegades la centrifugació i el rentat amb Carnoy. Finalment, es preparen les extensions cel·lulars deixant caure una gota de la mostra sobre un portaobjectes desengreixat amb metanol. Les preparacions es deixen assecar a l'aire i s'emmagatzemen a -20 °C fins el moment del seu ús (**Figura 8**).

#### SANG

Les mostres de sang perifèrica es recullen en heparina sòdica 1% o en EDTA i es processen seguint diferents protocols en funció de l'objecte d'estudi: fibres de cromatina, nucli interfàsics i cromosomes metafàsics, o DNA genòmic (**Figura 8**).



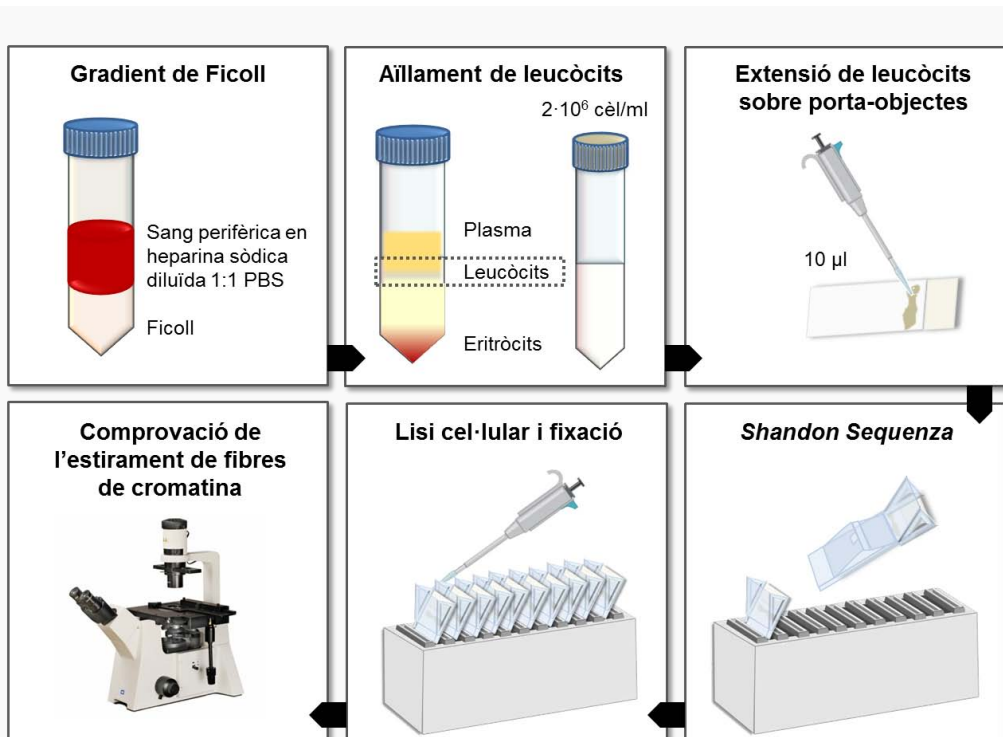


### Fibres de cromatina

Les fibres de cromatina s'obtenen a partir de 10 ml de sang perifèrica en heparina sòdica 1% d'acord amb el protocol descrit prèviament per Jefferson i Volpi (2010) (JEFFERSON I VOLPI 2010) (**Figura 9**). Breument, els leucòcits són aïllats de les mostres de sang perifèrica mitjançant un gradient de separació amb Ficoll-Paque (GE Healthcare). La sang es dilueix 1:1 en una solució 1x de tampó fosfat salí (PBS de l'anglès *phosphate buffered saline*) i s'afegeix gota a gota sobre un volum de Ficoll-Paque en una proporció de 3 volums de Ficoll-Paque per 4 volums de sang diluïda en 1x PBS. Després d'una

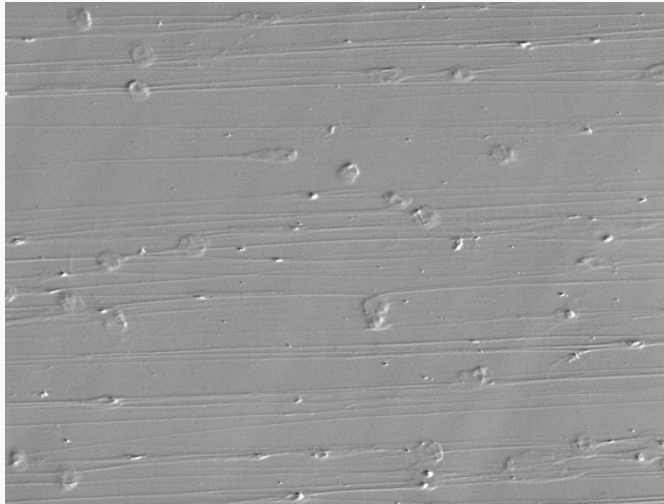
centrifugació a 400 g durant 30 minuts es recupera la interfase de leucòcits, que es distingeix entre el plasma sanguini i els eritròcits, amb l'ajuda d'una pipeta Pasteur. La fase leucocitària es centrifuga a 700 g durant 10 minuts i es resuspèn en 1x PBS. Aquest pas es repeteix un mínim de tres cops fins que el pellet queda net d'eritròcits. Després de realitzar un recompte cel·lular mitjançant una cambra de Neubauer, el pellet es resuspèn amb el volum de 1x PBS necessari per assolir una concentració final de  $2 \times 10^6$  cèl·lules/ml.

Un volum de 10  $\mu$ l de l'aïllament de leucòcits s'estenen sobre portaobjectes desengreixats amb metanol i s'assequen a 37 °C. A continuació, els portaobjectes es col·loquen en posició vertical en un suport *Shandon Sequenza* (Thermo scientific) i s'estiren les fibres de cromatina gràcies a l'aplicació de 150  $\mu$ l d'una solució de lisi de preparació diària (proporció 5:2 de 0,07 M NaOH:etanol). Finalment, les fibres de cromatina es fixen, amb els portaobjectes encara col·locats en el suport, aplicant 200  $\mu$ l



■ **Figura 9.** Metodologia d'estirament de fibres de cromatina basada en el protocol de Jefferson i Volpi (2010).

de metanol. Les preparacions s'assequen a temperatura ambient, es comprova l'extensió de les fibres de cromatina en un microscopi invertit de contrast de fases (**Figura 10**) i s'emmagatzemen a  $-20\text{ }^{\circ}\text{C}$  fins el moment del seu ús.



■ **Figura 10.** Fibres de cromatina al microscopi invertit de contrast de fases (200x). S'observa com les fibres de cromatina es disposen paral·lelament i es distingeixen algunes cèl·lules no lisades.

### Nuclis interfàsics i cromosomes metafàsics

Un volum de 0,5 ml de sang perifèrica en heparina sòdica 1% es cultiva en 5 ml de medi RPMI estàndard suplementat amb un 3,5% de fitohemaglutinina a  $37\text{ }^{\circ}\text{C}$  en 5%  $\text{CO}_2$  durant 72 hores. Passada la incubació, les cèl·lules es tracten amb 80  $\mu\text{l}$  de colcemid durant 25 minuts a  $37\text{ }^{\circ}\text{C}$ , es centrifuga el tub del cultiu a 800 g durant 8 minuts, es descarta el sobrenedant i es resuspèn gota a gota amb una solució hipotònica a  $37\text{ }^{\circ}\text{C}$  (KCl 0,075 M). La solució hipotònica es deixa actuar durant 20 minuts a  $37\text{ }^{\circ}\text{C}$  i, a continuació, es centrifuga de nou el cultiu a 800 g durant 8 minuts, es decanta el sobrenedant i es fixa amb Carnoy gota a gota. El pas del fixador es repeteix tres vegades. Quan el botó cel·lular està net es resuspèn en un volum baix de Carnoy i es realitzen les extensions sobre portaobjectes desengreixats en metanol. Les preparacions es deixen assecar a l'aire, es comprova la presència de metafases en un microscopi

invertit de contrast de fases i s'emmagatzemen a  $-20\text{ }^{\circ}\text{C}$  fins el moment del seu ús (**Figura 8**).

### DNA genòmic

S'extreuen 3 ml de sang perifèrica per cada individu en tubs amb EDTA i es realitza una extracció de DNA genòmic utilitzant el kit Gentra Puregene Blood (Qiagen) i seguint les instruccions proporcionades pels fabricants (Gentra Puregene Handbook, <https://www.qiagen.com/es/resources/>).

Un cop extret, el DNA es resuspèn en 250  $\mu\text{l}$  de solució d'hidratació i es deixa la mostra 16 hores a temperatura ambient. Posteriorment, es quantifica la concentració de DNA mitjançant l'espectrofotòmetre NanoDROP<sup>®</sup> 2000 (Thermo Fisher Scientific) a una longitud d'ona de 260 a 280 nm a partir de 1  $\mu\text{l}$  de mostra. El DNA s'emmagatzema a  $-20\text{ }^{\circ}\text{C}$  fins el moment del seu ús.

## 3.2 TÈCNiques CITOGÈNÈTIQUES BASADES EN LA HIBRIDACIÓ *IN SITU* FLUORESCENT (FISH)

### FISH EN ESPERMATOZOIDES

La FISH en nuclis descondensats d'espermatozoides es va aplicar en l'anàlisi de les freqüències de delecions i duplicacions de la regió 22q11.2 corresponent a la SDG/VCF, la regió 7q11.23 corresponent a la síndrome de Williams-Beuren i la regió 15q11q13 corresponent a la síndrome de Prader-Willi/Angelman (**Figura 8**). Per tal de determinar la susceptibilitat a la NAHR es van estudiar 10 individus control sense descendència afectada per cap trastorn genòmic (MOLINA ET AL. 2011) i 9 pares amb descendència afectada per la SDG/VCF.

L'objectiu de comparar la susceptibilitat a la NAHR entre aquestes tres regions era avaluar si els pares transmissors de la SDG/VCF presentaven una susceptibilitat generalitzada a la NAHR o si, per contra, la susceptibilitat a la NAHR es focalitzava únicament en la regió crítica 22q11.2.

## Selecció de sondes

---

Per analitzar cada regió crítica es van utilitzar diferents combinacions de sondes (**Figura 11**):

**Regió 7q11.23:** Sonda *locus* específica per la regió 7q11.23 (LSI *ELN* (OMIM 130160), *LIMK1* (OMIM 601329), D7S613, *Spectrum Orange*; Abbott Molecular), sonda *locus* específica del cromosoma 7 (LSI D7S486, *Spectrum Green*; Abbott Molecular) i sonda centromèrica del cromosoma 6 (CEP6, D6Z1, *Spectrum Aqua*; Abbot Molecular). Les sondes LSI D7S486 i CEP6 van ser utilitzades com a controls d'hibridació i ploïdia, respectivament.

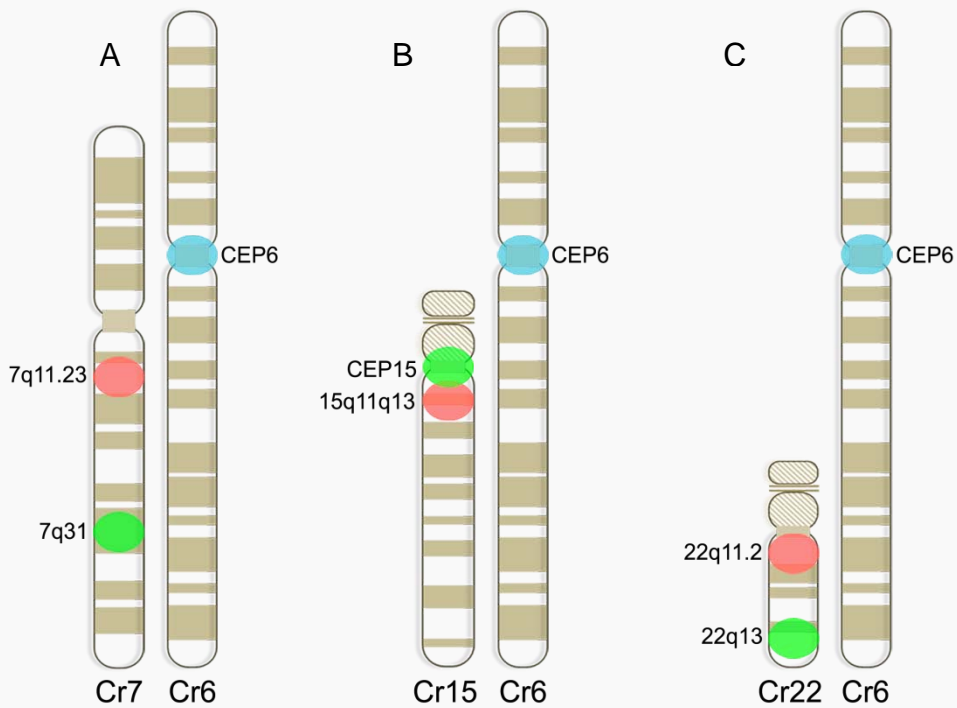
**Regió 15q11q13:** Sonda *locus* específica per la regió 15q11q13 (LSI D15S11, *Spectrum Orange*; Abbott Molecular), sonda centromèrica del cromosoma 15 (CEP15, D15Z1, *Spectrum Green*; Abbott Molecular) i sonda centromèrica del cromosoma 6 (CEP6, D6Z1, *Spectrum Aqua*; Abbot Molecular). Les sondes CEP15 i CEP6 van ser utilitzades com a controls d'hibridació i ploïdia, respectivament.

**Regió 22q11.2:** Sonda *locus* específica per la regió 22q11.2 (LSI *TUPLE1* (OMIM 600237), *Spectrum Orange*; Abbott Molecular), sonda *locus* específica del cromosoma 22 (LSI *ARSA* (OMIM 607574), *Spectrum Green*; Abbott Molecular) i sonda centromèrica del cromosoma 6 (CEP6, D6Z1, *Spectrum Aqua*; Abbot Molecular). Les sondes LSI *ARSA* i CEP6 van ser utilitzades com a controls d'hibridació i ploïdia, respectivament.

## Hibridació

---

Els portaobjectes d'espermatozoides es renten dues vegades en una solució 2x de citrat de sodi salí (SSC de l'anglès *saline-sodium citrate*) durant tres minuts, es deshidraten en una sèrie d'etanols (70%, 90% i 100% durant 2 minuts cadascun) i s'assequen a temperatura ambient. A continuació, els nuclis d'espermatozoides es descondensen submergint el portaobjectes en una solució que conté 5 mM 1,4-ditiotreitòl (DTT), 1% Tritó X-100 i 50 mM 2-amino-2-hidroximetil-1,3-propanodiol (Tris) a pH 7,4±0,2. El temps d'incubació oscil·la entre els 5 i els 10 minuts en funció de les mostres i les sondes



■ **Figura 11.** Combinacions de sondes per l'anàlisi de delecions i duplicacions de les regions crítiques de la síndrome de Williams-Beuren (A), la síndrome de Prader-Willi/Angelman (B) i la síndrome de DiGeorge/velocardiofacial (C).

aplicades. El DTT contribueix a descondensar el DNA dels espermatozoides a través del trencament dels ponts disulfur que s'estableixen entre les protamines. Després del pas de descondensació, es realitzen els mateixos rentats que en l'inici del protocol amb 2x SSC i una sèrie d'etanols de deshidratació.

Seguidament les mostres es desnaturalitzen submergint el portaobjectes en una solució de formamida 70%/2x SSC a 73 °C durant 2 minuts i deshidratats novament en una sèrie d'etanols (70%, 85% i 100% durant 1 minut en cadascun). Després d'assecar el portaobjectes a temperatura ambient, s'apliquen les sondes sobre la mostra amb el suport d'un cobreobjectes, es segella amb cola d'enquadernar i es deixa hibridar en una cambra humida a 37 °C durant un mínim de 16 hores.

Els rentats posthibridació es realitzen en 0,7x SSC amb 0,3% NP-40 a 73 °C durant 2 minuts seguit de 2x SSC amb 0,1% NP-40 a temperatura ambient durant 1 minut.

Finalment, els portaobjectes s'assequen a temperatura ambient i es contratenyeixen amb 4',6-diamino-2-fenilindol (DAPI, Abbott Molecular).

### Valoració microscòpica

---

Les hibridacions es van valorar en un microscopi d'epifluorescència Olympus BX60 equipat amb un filtre triple (*Aqua*, FITC i Cy3) i filtres específics per *Aqua*, FITC i Cy3. Per cada individu es van analitzar 10.000 nuclis d'espermatozoides aplicant els criteris de valoració descrits prèviament per Molina et al. (2010):

- Únicament s'analitzen el espermatozoides amb un contorn ben definit. Els espermatozoides solapats es descarten del recompte.
- D'acord amb el número i la distribució dels senyals observats s'assignen els següents genotips:

**Normal:** Espermatozoides amb un senyal corresponent a la sonda de la regió crítica específica, un senyal corresponent a la sonda de control d'hibridació i un senyal de la control de ploïdia.

**Deleció:** Espermatozoides amb absència del senyal corresponent a la sonda de la regió crítica específica, un senyal corresponent a la sonda de control d'hibridació i un senyal de la control de ploïdia.

**Duplicació:** Espermatozoides amb dos senyals corresponents a la sonda de la regió crítica específica, de igual mida i intensitat i separats l'un de l'altre per una distància superior al diàmetre de cada senyal únic, un senyal corresponent a la sonda de control d'hibridació i un senyal de la control de ploïdia.

### Anàlisi estadística

---

Les dades es van analitzar mitjançant el programa Graph Pad Prism 5.03. En les comparacions de les freqüències de delecions i duplicacions a nivell poblacional, el llindar de significança per l'anàlisi de les dades es va establir en  $p < 0,05$ . En canvi, per les anàlisis a nivell individual es va establir el llindar de significança en  $p < 0,01$  per tal de minimitzar l'error estadístic de tipus I.

Per tal de valorar la susceptibilitat a generar delecions i duplicacions de cada regió crítica, la freqüència mitjana de delecions, duplicacions i el sumatori de delecions i duplicacions (del+dup) es va comparar amb la mitjana determinada prèviament en controls mitjançant el test de Mann-Whitney (MOLINA ET AL. 2011). També es va avaluar si les freqüències de delecions, duplicacions i del+dup diferien significativament entre les tres regions estudiades mitjançant el test de Friedman. Amb l'objectiu de determinar la influència de l'edat les freqüències de delecions, duplicacions i del+dup de les tres regions analitzades es van contrastar amb l'edat mitjançant la correlació de Spearman.

A nivell individual, es va determinar la susceptibilitat a la NAHR en cada regió crítica comparant les freqüències de delecions, duplicacions i del+dup amb les freqüències basals observades en controls (MOLINA ET AL. 2011) utilitzant el test de la khi-quadrat.

D'altra banda, a fi de determinar la participació de la NAHR inter-cromàtide o intra-cromàtide en la generació de reorganitzacions, es va realitzar una correlació de Spearman entre les freqüències de delecions i duplicacions a nivell poblacional. Per tal d'analitzar si la mitjana de delecions era significativament diferent a les duplicacions es van aplicar el test de Wilcoxon i la khi-quadrat a nivell poblacional i individual, respectivament.

---

## FISH EN FIBRES DE CROMATINA

La *fiber*-FISH va ser la tècnica seleccionada per estudiar l'arquitectura genòmica dels LRC22-2 i LCR22-4 de forma individualitzada (**Figura 8**). Es van analitzar els blocs de seqüències en tàndem en 6 individus control sense descendència afectada per la SDG/VCF i 2 pares transmissors de la SDG/VCF que mostraven increments de delecions 22q11.2 en espermatozoides (veure apartat 4.1).

### Selecció de sondes

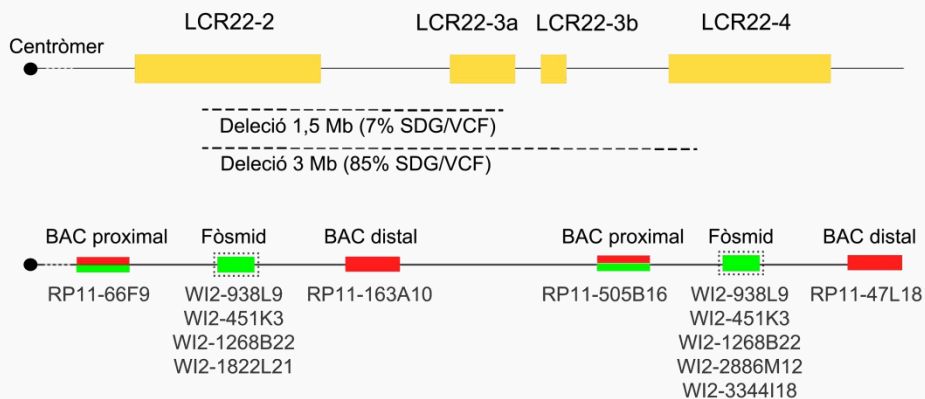
---

La caracterització de les variacions del número de còpies de seqüències contingudes en aquests LCR requereix l'ús simultani de tres sondes de DNA de tipus BAC i fòsmid. Els BAC i fòsmids són vectors de clonació inserits en plasmidis d'*Escherichia coli* que contenen seqüències d'interès de ~150 kb i ~40 kb, respectivament.



El disseny experimental per l'estudi dels LCR inclou la utilització de dos BAC de còpia única, que delimiten els extrems proximals i distals dels LCR, i una sèrie de fòsmids, estudiats en hibridacions independents en combinació amb els BAC, que permeten caracteritzar les seqüències internes dels LCR.

La selecció de les sondes per l'anàlisi dels LCR22-2 i LCR22-4 es va realitzar d'acord amb la versió d'assemblatge del genoma humà de Febrer de 2009 publicada a UCSC Genome Browser (<http://genome.ucsc.edu/>). Els BAC seleccionats per delimitar els extrems proximal i distal del LCR22-2 van ser RP11-66F9 (F9) i RP11-163A10 (A10) i del LCR22-4 RP11-505B16 (B16) i RP11-47L18 (L18). Per cada LCR es va seleccionar una combinació de fòsmids específica que cobria la majoria de les seqüències descrites en cada LCR (**Figura 12**). En concret, per LCR22-2 es van seleccionar els fòsmids WI2-938L9 (L9), WI2-451K3 (K3), WI2-1268B22 (B22) i WI2-1822L21 (L21) mentre que per LCR22-4 WI2-938L9 (L9), WI2-451K3 (K3), WI2-1268B22 (B22), WI2-2886M12 (M12) i WI2-3344I18 (I18) (**Taula 5**).



**Figura 12.** Selecció de sondes BAC i fòsmid per l'anàlisi de blocs de duplicacions segmentals dels LCR22-2 i LCR22-4.

Els BAC i fòsmids seleccionats van ser proveïts per *Source BioScience* (<http://www.sourcebioscience.com/>) i *BACPAC Resources Center*, C.H.O.R.I., USA (<http://bacpac.chori.org/>). L'especificitat i l'eficiència de totes les sondes BAC i fòsmid

va ser validada mitjançant la FISH en cromosomes metafàsics (**Figura 13**) i seguint el protocol d'hibridació de la FISH en nuclis interfàsics (veure apartat 2.3).

■ **Taula 5.** Característiques de les sondes d'hibridació BAC i fòsmid (UCSC Assembly Feb 2009).

Clon	Abreviatura	Referència UCSC	Mida (kb)	Gens UCSC	OMIM <sup>a</sup>
<b>BAC</b>					
RP11-66F9 <sup>1</sup>	F9	RP11-66F9	175,3	<i>GAB4, CECR7</i>	
RP11-163A10 <sup>2</sup>	A10	RP11-163A10	183,1	<i>TUPLE1</i>	600237
RP11-505B16 <sup>1</sup>	B16	RP11-505B16	171,3	<i>PI4KA</i>	600286
RP11-47L18 <sup>1</sup>	L18	RP11-47L18	186,6	<i>UBE2L3, PPIL2</i>	603721, 607588
<b>Fòsmid</b>					
WI2-938L91	L9	G248P80545F5	41,6	<i>USP18, AK129567</i>	607057
WI2-451K31	K3	G248P8704F2	44,3	<i>AK302545, GGT3P</i>	
WI2-1268B221	B22	G248P82259A11	41,9	<i>GGT3P</i>	
WI2-1822L211	L21	G248P86641F11	39,0	<i>DGCR6, PRODH</i>	601279, 606810
WI2-2886M121	M12	G248P88386G6	44,2	<i>RIMBP3C</i>	612701
WI2-3344I181	I18	G248P800126E9	45,5	<i>PI4KAP2</i>	

<sup>1</sup> BACPAC Resources Center, CHORI

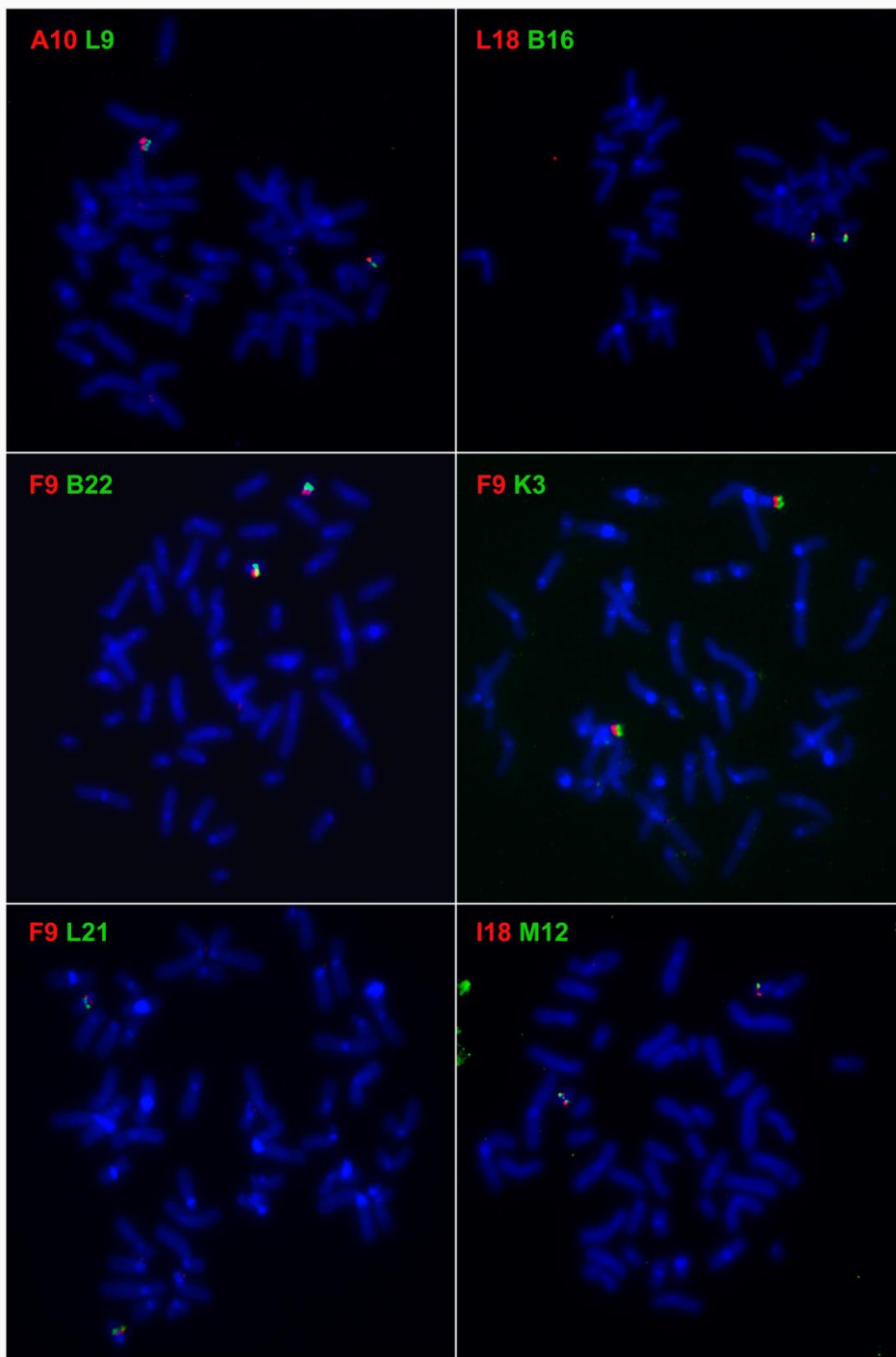
<sup>2</sup> Source BioScience

<sup>a</sup> Online Mendelian Inheritance in Man; <http://www.ncbi.nlm.nih.gov/omim>

### Obtenció i marcatge de sondes BAC i fòsmids

Els bacteris que presenten els constructes de les sondes seleccionades es cultiven en medi LB estàndard amb una concentració de cloramfenicol de 12,5 µg/ml durant 16 hores a 37 °C i amb una agitació de 250 rpm. La presència d'un gen de resistència a cloramfenicol en el plàsmid inserit garanteix el cultiu exclusiu dels bacteris que inclouen el plàsmid.

Un cop finalitzat el temps d'incubació, els plàsmids es purifiquen mitjançant el kit QIAprep Spin Miniprep (Qiagen) adaptant les indicacions subministrades pel proveïdor. El cultiu es transfereix a tubs cònics de 10 ml i es centrifuga a 4.000 rpm durant 10 minuts. A continuació s'elimina el sobrenedant i es resuspèn el botó en 400 µl del *Buffer P1* amb una concentració de 10 mg/ml RNasa I. La solució es transfereix a un tub de microcentrífuga de 1,5 ml en el qual s'afegeixen 400 µl del *Buffer P2* (tampó de lisi) i es barreja per inversió entre 4 i 6 vegades. Seguidament s'incorporen 400 µl del *Buffer N3* (tampó de precipitació de restes cel·lulars i DNA bacterià) i es torna a invertir el tub



■ **Figura 13.** Eficiència i especificitat de la hibridació de les sondes BAC i fòsmid en extensions cromosòmiques. En cada imatge s'observa una metafase amb senyals de dues sondes marcades diferencialment que hibriden sobre el cromosoma 22.

entre 4 i 6 vegades. Es centrifuga el tub a 14.000 rpm durant 1 minut, es transfereix el sobrenedant a les columnes amb membrana de sílice proporcionades pel kit i es centrifuga a 14.000 rpm durant 1 minut. El DNA plasmídic adherit a la membrana de la columna es purifica mitjançant l'addició de 500 µl de *Buffer PB* a la columna, centrifugant a 14.000 rpm durant 1 minut, afegint 750 µl de *Buffer PE* i centrifugant dues vegades a 14.000 rpm durant 1 minut. Després de cada centrifugació és necessari eliminar el líquid que resta en el tub que conté la columna. Per últim, es transfereix la columna a un tub de microcentrífuga nou de 1,5 ml, s'afegeixen 50 µl de *Buffer EB* (tampó d'elució) prèviament escalfat a 65 °C i es deixa actuar 1 minut abans de centrifugar el tub amb la columna a 14.000 rpm durant 1 minut. El DNA extret és quantificat mitjançant l'espectrofotòmetre Nanodrop i emmagatzemat a -20 °C fins el moment del seu ús.

La incorporació de la fluorescència en les sondes de DNA es realitza mitjançant una reacció de *nick translation*. La reacció de *nick translation* es basa en la incorporació de nucleòtids marcats gràcies a l'acció de dos enzims: la DNasaI s'encarrega d'hidrolitzar DNA d'una de les dues cadenes introduint incisions, o *nicks*, i la DNA polimerasa I refà els fragments hidrolitzats incorporant els residus de desoxinucleòtids trifosfat (dNTP) complementaris a la cadena motlle (dUTP marcats, dATP, dGTP, dCTP). El marcatge de les sondes pot ser directe, si en la reacció s'incorporen dUTP conjugats amb un fluorocrom, o indirecte, si s'incorporen dUTP conjugats amb digoxigenina o biotina. En el cas que el marcatge es produeixi de forma indirecta, després de la hibridació de les sondes els dUTP incorporats conjugats amb digoxigenina es detecten aplicant un anticòs contra digoxigenina conjugat amb un fluorocrom. Mentre que si els dUTP estaven conjugats amb biotina la detecció es realitza amb estreptavidina o avidina conjugades amb un fluorocrom.

La reacció de *nick translation* es va realitzar mitjançant el kit de Abbott Molecular que permet marcar 1µg de DNA. La reacció es prepara seguint les indicacions de la casa comercial (<https://www.abbottmolecular.com/nick-translation-kit-preparing-the-reagents.html>) i es deixa actuar 16 hores a 15 °C. Passat aquest període, la reacció s'atura elevant la temperatura a 70 °C durant 10 minuts. Com a resultat de la *nick translation* s'obtenen fragments d'uns 200 parells de bases que es precipiten amb 125 µl etanol 100%, 12,5 µl

de DNA d'esperma de salmó 1 µg/ml (GE Healthcare) i 6,25 µl d'acetat de sodi 3M (pH 5,5) durant 3 hores a -80 °C. Les sondes centrifuguen a 14.000 rpm durant 30 minuts, es renten amb etanol al 70% i es centrifuguen de nou a 14.000 rpm 6 minuts. El pellet de sondes s'asseca a l'aire i es resuspèn amb 50 µl de solució tampó 1x TE (10 mM Tris, 1 mM EDTA, pH 7,0) per obtenir una concentració final de sonda de 20 ng/µl. Les sondes obtingudes s'emmagatzemen a -20 °C fins el moment del seu ús.

Tal i com es descriu en l'apartat anterior, per a la delimitació i anàlisi intern dels LCR és necessari hibridar simultàniament dos BAC i un fòsmid. El BAC proximal es va marcar doblement amb *Spectrum Orange*-dUTP (Abbot Molecular) i *Digoxigenin*-11-dUTP (Roche), el fòsmid amb *Digoxigenin*-11-dUTP i el BAC distal amb *Spectrum Orange*. Les sondes marcades amb digoxigenina es van detectar després de la seva hibridació mitjançant un anticòs antidigoxigenina conjugat amb FITC.

Les combinacions de sondes per la *fiber*-FISH es preparen barrejant 400 ng de cada sonda marcada amb 8 µg de DNA competidor Cot1 (Invitrogen) per sonda. La barreja de sondes s'evapora a 42 °C i el pellet resultant es resuspèn en 10 µl de tampó d'hibridació LSI/WCP (Abbot Molecular). Abans de la hibridació les sondes es desnaturalitzen a 75 °C durant 5 minuts i es deixen pre-hibridar a 37 °C durant 45 minuts.

## Hibridació

Els portaobjectes de fibres de cromatina es deshidraten en una sèrie d'etanols (70%, 90% i 100% durant 2 minuts cadascun) i s'assequen a temperatura ambient. Seguidament les fibres de cromatina es desnaturalitzen submergint el portaobjectes en una solució de formamida 70%/2x SSC a 70 °C durant 2 minuts, rentats en 2x SSC durant 5 minuts i deshidratades novament en una sèrie d'etanols (70%, 85% i 100% durant 1 minut en cadascun). Després d'assecar el portaobjectes a temperatura ambient, s'apliquen les sondes sobre la mostra, es protegeix amb un cobreobjectes, es segella amb cola d'enquadernar i es deixa hibridar en una cambra humida a 37 °C durant 48 hores.

Els rentats posthibridació es realitzen en 2x SSC a 42 °C durant 3 minuts seguit d'un rentat amb 1x PBS a temperatura ambient durant 5 minuts. Els portaobjectes s'incuben

amb 200 µl de solució de bloqueig (5% llet deshidratada en 1x PBS/0,05% Tween-20 (Sigma-Aldrich)) a 37 °C durant 30 minuts. Passada aquesta incubació els portaobjectes es transfereixen a 1x PBS a temperatura ambient durant 5 minuts i es procedeix a una segona incubació d'una hora a 37 °C amb 200 µl de solució de detecció (1% llet deshidratada en 1x PBS/0,05% Tween-20 (Sigma-Aldrich) amb 1 µg/ml de *fluorescein-conjugated antidigoxigenin* (Roche)). Finalment, els portaobjectes es renten tres vegades en 1x PBS durant un minut, s'assequen a temperatura ambient i es contratenyeixen amb DAPI.

### Valoració microscòpica

---

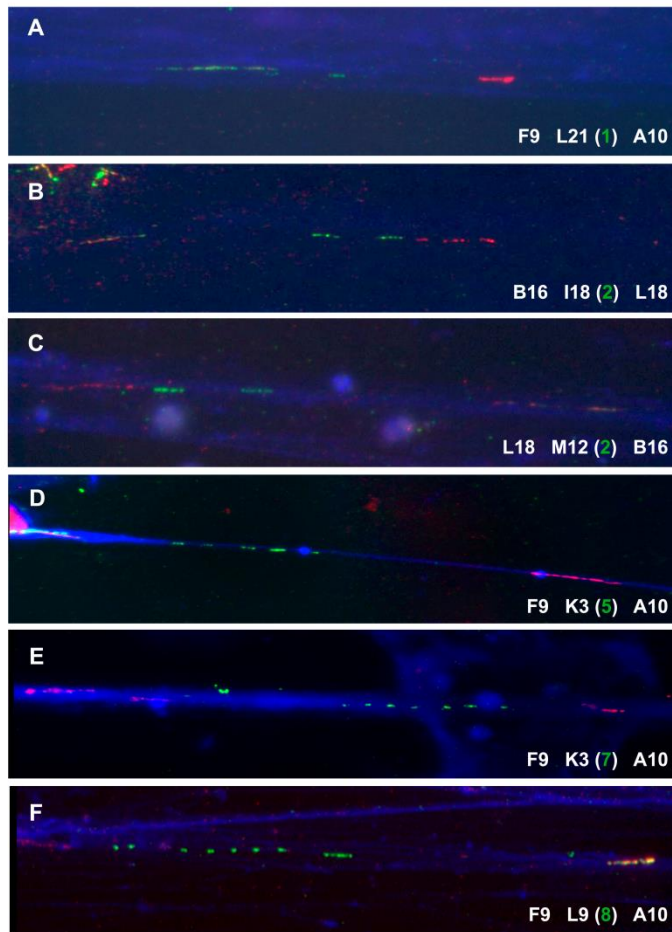
La visualització de les hibridacions es va realitzar en un microscopi d'epifluorescència Olympus BX60 equipat amb el sistema de captura d'imatges *Isis Imaging System* (Metasystems). L'anàlisi es va realitzar mitjançant un filtre dual (Cy3/FITC) i filtres específics per DAPI, FITC i Cy3 (**Figura 14**). Per cada combinació de sondes es va analitzar el número de còpies del fòsmid en qüestió en un mínim de 20 fibres informatives adaptant criteris de valoració descrits per Molina et al. (2012):

- Les fibres es consideren informatives quan s'observa en disposició lineal un senyal de cada BAC de posicionament, amb un marcatge diferenciat, que delimiten com a mínim un senyal corresponent al fòsmid (**Figura 15A**).
- Dos senyals o més del mateix fòsmid es consideren independents quan estan separats per una distància superior a la del segment més curt que conforma el conjunt de senyals (**Figura 15B**).
- Els senyals es consideren informatius independentment de la seva mida (**Figura 15C**).

### Anàlisi estadística

---

Les dades es van analitzar mitjançant el programa Graph Pad Prism 5.03. Els resultats obtinguts en controls es van sotmetre a un anàlisi de dades atípiques (Test de Grubbs). Posteriorment, es van establir rangs de normalitat dels blocs de seqüències en tàndem. En els dos pares SDG/VCF analitzats, tots els valors obtinguts fora el rang de

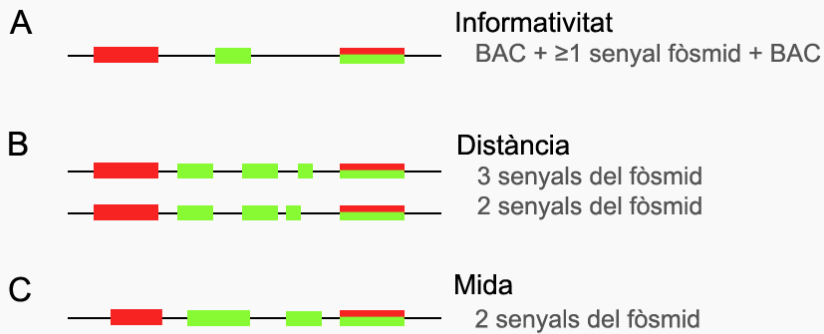


■ **Figura 14.** Anàlisi microscòpica dels LCR22-2 i LCR22-4 mitjançant *fiber*-FISH. En cada imatge s'observa una fibra de cromatina informativa, s'indica la combinació dels BAC i fòsmid hibridats i s'especifica entre parèntesis les còpies del fòsmid. **A, D, E i F** corresponen a l'anàlisi del LCR22-2, mentre que **B i C** corresponen al LCR22-4.

normalitat van ser considerats significativament diferents a la població control i es van atribuir com a CNV.

#### FISH EN NUCLIS INTERFÀSICS

La FISH en nuclis interfàsics es va emprar per la valoració de la presència d'inversions de la regió crítica 22q11.2 (**Figura 8**). La presència d'haplotips amb inversió 22q11.2 es va analitzar en 5 individus control (4 amb descendència afectada per la SDG/VCF però



■ **Figura 15.** Criteris de valoració per *fiber*-FISH.

no transmissors de la deleció i 1 sense descendència afectada) i 8 individus transmissors de la SGD/VCF.

### Selecció de sondes

El disseny experimental per la detecció de la inversió entre els LCR22-2 i LCR22-4 requeria la utilització de tres sondes de tipus BAC (**Figura 16**). Procedint de la mateixa manera que en la selecció de sondes per la FISH en fibres de cromatina, els BAC van ser seleccionats a la base de dades UCSC Genome Browser (Febrer 2009). En concret, per aquests assajos es van utilitzar els BAC RP11-66F9 (abbr. F9), RP11-505B16 (abbr. B16) i RP11-47L18 (abbr. L18) (**Taula 5**).

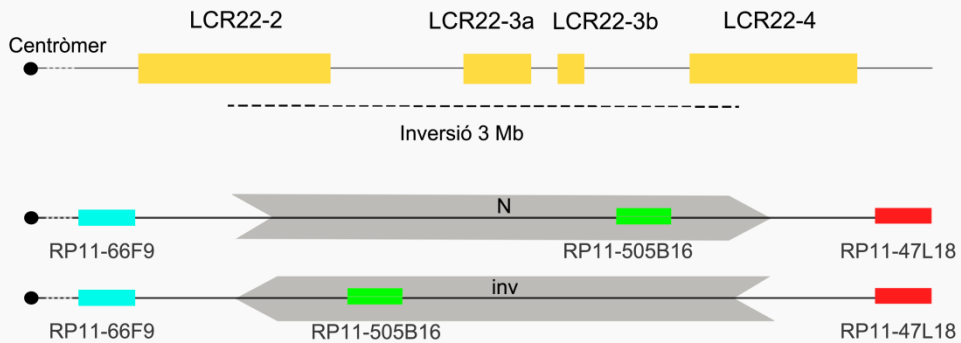
L'especificitat i l'eficiència de totes les sondes BAC i fòsmid es va validar mitjançant la FISH en cromosomes metafàsics i seguint el protocol d'hibridació de la FISH en nuclis interfàsics (**Figura 13**).

### Obtenció i marcatge de sondes BAC i fòsmids

Les sondes es van obtenir i marcar tal i com s'especifica en l'apartat anterior. Per l'anàlisi de les inversions es van marcar els BAC amb tres fluorocroms diferents: F9 amb *Spectrum Aqua*-dUTPs (Abbott Molecular), B16 amb *Spectrum Green*-dUTPs (Abbott Molecular) i L18 amb *Spectrum Orange*-dUTPs (Abbott Molecular).



Les combinacions de sondes es preparen barrejant 200 ng de cada sonda marcada amb 4 µg de DNA competidor Cot1 (Invitrogen) per sonda. La barreja de sondes s'evapora a 42 °C i el pellet resultant es resuspèn en 5 µl de tampó d'hibridació LSI/WCP (Abbot Molecular). Abans de la hibridació les sondes es desnaturalitzen a 75 °C durant 5 minuts i es deixen a 37 °C durant 45 minuts.



■ **Figura 16.** Selecció de sondes BAC per l'anàlisi d'inversions de la regió 22q11.2.

## Hibridació

Els portaobjectes es renten dues vegades en 2x SSC durant 3 minuts, es deshidraten en una sèrie d'etanols (70, 90, 100% durant 2 minuts cadascun) i s'assequen a temperatura ambient. Seguidament les mostres es desnaturalitzen submergint el portaobjectes en una solució de formamida 70%/2x SSC a 73 °C durant 2 minuts i deshidratades novament en una sèrie d'etanols (70, 85, 100% durant 1 minut en cadascun). Després d'assecar els portaobjectes a temperatura ambient, s'apliquen les sondes i es deixen hibridar en una cambra humida a 37 °C durant 24 hores.

Els rentats posthibridació es realitzen en SSC 1x amb 0,3% NP-40 a 73°C i 2x SSC amb 0,1% NP-40 a temperatura ambient durant 1 minut per cada rentat. Finalment, els portaobjectes s'assequen a temperatura ambient i es contratenyeixen amb DAPI.

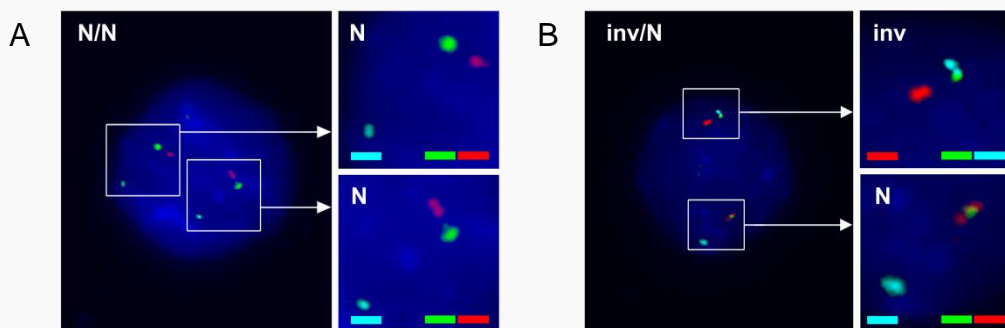
## Valoració microscòpica

La visualització de les hibridacions es va realitzar en un microscopi d'epifluorescència Olympus BX60 equipat amb el sistema de captura d'imatges *Isis Imaging System* (Metasystems). L'anàlisi es va realitzar mitjançant un filtre triple (*Aqua*, FITC i Cy3) i filtres específics per *Aqua*, FITC i Cy3.

En l'anàlisi d'inversions de la regió 22q11.2, es van analitzar un mínim de 100 haplotips informatius per cada mostra aplicant els següents criteris de valoració:

- Únicament s'analitzen els nuclis amb dos senyals de cada sonda distribuïts en dos haplotips diferenciats.
- Es consideren informatius aquells haplotips en els quals es pot distingir una distribució de dos senyals agrupats (molt pròxims o sobreposats) i un tercer senyal separat.
- El senyal distanciat es troba separat dels altres dos per almenys el doble de la distància que separa els dos senyals agrupats.

La distribució dels senyals de cada sonda permet distingir els haplotips normals (N) i els haplotips que mostren una inversió de 3Mb a la regió 22q11.2 (*inv*) (**Figura 17**).



■ **Figura 17.** Anàlisi microscòpica d'inversions 22q11.2 mitjançant FISH en nuclis interfàsics. El nucli de la figura **A** mostra dos haplotips N (N/N), mentre que el nucli de la figura **B** mostra un haplotip *inv* i un haplotip N (*inv*/N).

## Anàlisi estadística

Les dades es van analitzar mitjançant el programa Graph Pad Prism 5.03. Les freqüències dels haplotips normals i invertits van ser comparades entre la població

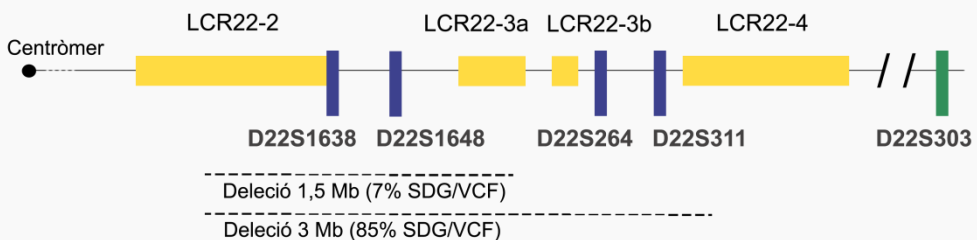
control i la SDG/VCF aplicant un test de Mann-Whitney. Les diferències es van considerar significatives quan  $p < 0,05$ .

### 3.3 TÈCNiques MOLECULARS BASADES EN LA REACCIÓ EN CADENA DE LA POLIMERASA (PCR)

#### MICROSATÈL·LITS

En les 8 famílies SDG/VCF de les quals es disposava de mostres de DNA genòmic de tots els membres (pare, mare i fill afectat), es va realitzar una anàlisi de microsatèl·lits per tal de determinar l'origen parental del cromosoma delecionat en els descendents afectats (**Figura 8**). Es van genotipar 4 microsatèl·lits interns de la regió crítica 22q11.2 (D22S1638, D22S1648, D22S264, D22S311) i un cinquè microsatèl·lit extern (D22S303) (**Figura 18**). Els encebadors *forward* (F) i *reverse* (R) dissenyats, el marcatge a 5' dels encebadors F, els parells de bases (pb) de l'amplicó de cada marcador, la seva posició cromosòmica i la temperatura d'hibridació optimitzada per cada parella es detallen a la **Taula 6**.

Els microsatèl·lits es van amplificar per PCR utilitzant encebadors marcats amb HEX o amb 6-FAM. En la preparació de la reacció de PCR es van barrejar 40-90 ng de DNA genòmic, 25 picomols de cada encebador (Roche), 2,5 mM de cada dNTP (Applied Biosystems), 25 mM MgCl<sub>2</sub> (Applied Biosystems), 10x PCR *Buffer II* (Applied



■ **Figura 18.** Posició dels microsatèl·lits amplificats per la determinació de l'origen parental de la SDG/VCF.

Biosystems) i 1 U de la polimerasa AmpliTaq Gold DNA (Applied Biosystems). L'amplificació es va dur a terme de la següent manera: 10 minuts a 94°C; 35 cicles de 30 segons a 94°C, 30 segons a una temperatura d'hibridació específica per cada encebador (**Taula 6**) i 35 segons a 72 °C; i finalment un pas d'extensió de 12 minuts a 72 °C.

Els productes de la reacció de PCR es van analitzar per electroforesi capil·lar en l'analitzador ABI3130 XL (Applied Biosystems) amb el suport del programa *Peak Scanner* versió 2.0 (Applied Biosystems).

■ **Taula 6.** Marcadors de microsatèl·lits utilitzats en la determinació de l'origen parental de la SDG/VCF.

Microsatèl·lit	Posició	Encebadors	Marcatge	Mida (pb)	Temperatura d'hibridació
D22S1638 <sup>a</sup>	22q11.21	F: GACAACAGCAAATTGCACATT R: TCACGCCACTACCCTCCAG	HEX	93	55°C
D22S1648 <sup>a</sup>	22q11.21	F: CAGATGCTTCAGGAGAAGTG R: AGTTGTCAGATGCCTAAGAGA	HEX	152	50°C
D22S264 <sup>a</sup>	22q11.21	F: ATTAACTCATAAAGGAGCCC R: CACCCACCCAGAGGTATTCC	HEX	190-198	56°C
D22S311 <sup>a</sup>	22q11.21	F: GCTAGTGTGAGATAACGAAGCC R: TTTTGTATTTTATAGTAGAGACGG	6-FAM	262	63°C
D22S303 <sup>b</sup>	22q11.22	F: AGGACCTCAGACTGGTCAGTC R: CTCCCATGAGAAGGTACTCTCC	6-FAM	220-233	56°C

<sup>a</sup> Microsatèl·lit intern a la regió crítica SDG/VCF.

<sup>b</sup> Microsatèl·lit extern a la regió crítica SDG/VCF.

## DROPLET DIGITAL PCR (DDPCR)

La droplet digital PCR (ddPCR) es basa en la realització de milers de reaccions de PCR simultànies individualitzades en gotes d'oli de manera que permet realitzar una quantificació total de les seqüències d'interès amb una alta precisió. Aquesta tècnica es va seleccionar per confirmar els CNV identificats mitjançant *fiber-FISH*.

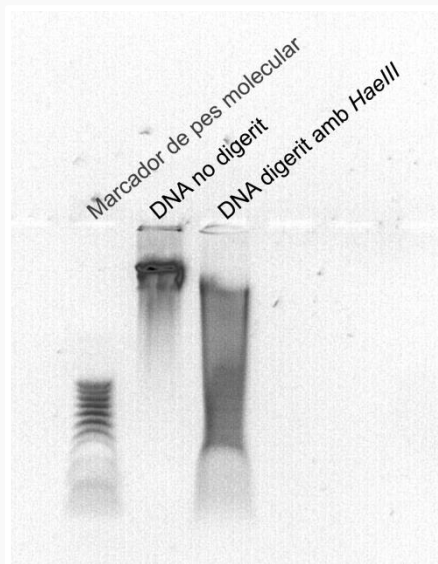
Es va analitzar el número de còpies de les seqüències cobertes pels fòsmids L9 i K3 en el conjunt de tots els LCR22 (**Figura 8**). L'elecció de les seqüències dels fòsmids L9 i K3 es va realitzar com a conseqüència dels CNV detectats en *fiber-FISH* (veure apartat 4.2).

Per l'anàlisi de les seqüències paràlogues del fòsmid L9 es van analitzar 21 mostres d'individus control, incloent individus amb descendència afectada per la SDG/VCF

però no transmissors i individus sense descendència afectada, i 8 mostres d'individus transmissors de la SDG/VCF. Per la valoració de les seqüències paràlogues del fòsmid K3 es van analitzar 6 mostres d'individus control, incloent individus amb descendència afectada per la SDG/VCF però no transmissors i individus sense descendència afectada, i 8 mostres d'individus transmissors de la SDG/VCF. L'ampliació del nombre de controls en l'anàlisi del fòsmid L9 es va realitzar per tal de caracteritzar d'una forma més acurada el número de seqüències del pseudogen d'aquest fòsmid en la població control (veure apartat 4.2).

### Digestió del DNA

La ddPCR requereix un fragmentació de la mostra de DNA genòmic per tal de què la inclusió de la mostra en els milers de gotes d'oli es dugui a terme satisfactòriament. La fragmentació òptima del DNA en el context del nostre estudi es va assolir gràcies l'acció de l'enzim de restricció *HaeIII*. Per la digestió del DNA es van incubar 2 µg de DNA genòmic amb 20 U de *HaeIII* (New England Biolabs) i 5 µl de la solució tampó NEB4, en un volum final de 25 µl ajustat amb aigua desionitzada, durant 1 hora a 37 °C (**Figura 19**). El DNA fragmentat es va diluir a una concentració final de 6 ng/µl.

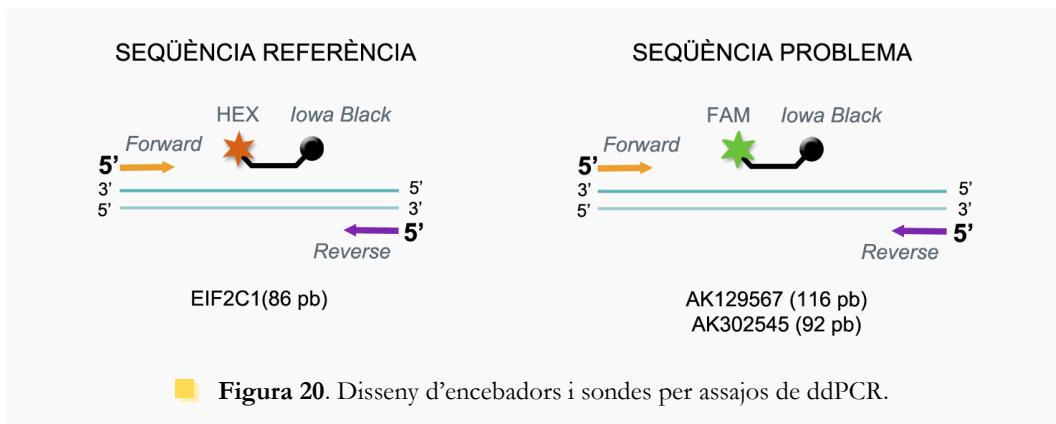


■ **Figura 19.** Verificació de la digestió del DNA.

## Disseny d'encebadors i sondes TaqMan

Per estudiar les seqüències de cobertes pels fòsmids L9 i K3 es va fer un disseny d'encebadors *forward* i *reverse* i de sondes de tipus TaqMan amb els extrems modificats amb un emissor de fluorescència a 5' i un desactivador de fluorescència 3'. Gràcies a l'activitat exonucleasa 5' – 3' de la Taq polimerasa es produeix l'escissió del fluorocrom de la sonda i al desvincular-se de la influència del desactivador emet fluorescència.

Per L9 es va generar un amplicó de 116 pb mitjançant l'encebador *forward* CAGTGGGACTCTCATCAAAC, l'encebador *reverse* AGGAGCGAGAAATAGAGTCC i la sonda FAM-AAAGTTGCCATCAGCCAGATGCCAG-Iowa Black®FQ (Bio-Rad). La seqüència d'aquest amplicó correspon específicament al gen *AK129567*. D'altra banda, per K3 es va generar un amplicó 92 pb mitjançant l'encebador *forward* ATTTTGCGGTGTCAGAAGGT, l'encebador *reverse* CCTCTCAGTGACTGTCTCCT i la sonda FAM-CGGGAGGCAAGCAGTCCCTGGTCC-Iowa Black®FQ (Bio-Rad). La seqüència d'aquest amplicó correspon específicament al gen *AK302545*. En ambdós casos es va aplicar simultàniament l'assaig de referència del gen *EIF2C1* (OMIM 606228) amb sonda marcada diferencialment amb HEX/Iowa Black®FQ (PrimePCR™ ddPCR™ Copy Number Assay *EIF2C1*, Bio-Rad) (**Figura 20**).



## Condicions d'amplificació

La reacció de ddPCR es va realitzar en un volum final de 20 µl per reacció. La quantitat de DNA es va optimitzar en 9 ng per reacció en un volum de 8 µl als quals s'afegeixen 10 µl Supermix PCR (Bio-Rad), 1 µl de la mescla d'encebadors i sonda per l'amplificació

de la seqüència d'interès i 1 µl de la mescla d'encebadors i sonda per l'amplificació de la seqüència de referència.

Les condicions d'amplificació van ser 10 minuts a 95 °C (1 cicle), 30 segons a 94 °C i 60 segons a 55,5 °C (40 cicles) i 10 minuts a 98 °C (1 cicle). Després de la PCR es va realitzar una lectura de la intensitat de cada gota de forma individualitzada en el lector de gotes de Bio-Rad per tal de discriminar el nombre total de gotes amb reacció positiva per cada amplicó. Cada assaig es va realitzar per duplicat de manera que per cada assaig es va obtenir un número mitjà de còpies.

### Anàlisi estadística

Les dades es van analitzar mitjançant el programa Graph Pad Prism 5.03. El resultat obtingut a través de la ddPCR es van processar a partir de la mitjana dels resultats de les diferents rèpliques de cada assaig. En individus no transmissors, es va determinar la presència de dades atípiques per l'anàlisi estadística (Test de Grubbs). Per tal de determinar si hi havia diferències entre les número de seqüències paràlogues dels LCR22 entre la població control i la població de pares transmissors de la SDG/VCF, la mitjana de còpies de cada fòsmid es va comparar entre les dues poblacions mitjançant un test de Mann-Whitney. Les diferències es van considerar significatives quan  $p < 0,05$ .

Els resultats obtinguts en controls van permetre establir rangs de normalitat i valorar a nivell individual si algun pare SDG/VCF presentava un CNV respecte els controls. Com en l'anàlisi dels resultats de blocs de seqüències en tàndem, tots els valors obtinguts fora el rang de normalitat van ser considerats significatius.

Per tal de valorar l'equivalència entre les dades de la *fiber*-FISH i la ddPCR, es va comparar el sumatori dels blocs de seqüències en tàndem dels LCR22-2 i LCR22-4 per cada individu amb les còpies de seqüències paràlogues dels LCR22 mitjançant un test de correlació de Spearman.

---

### GENOTIPATGE

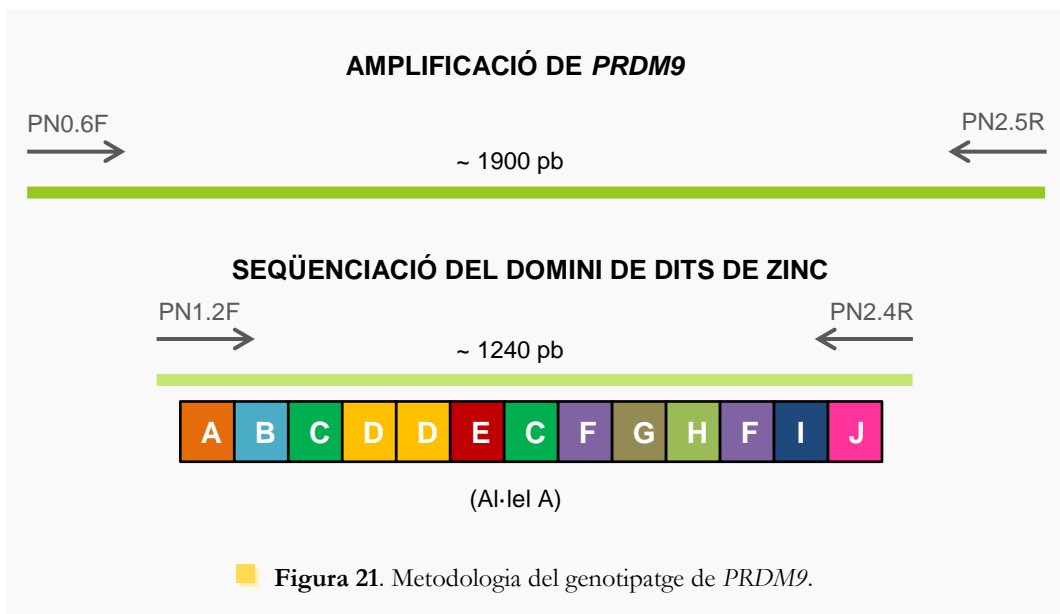
Per l'establiment dels diferents al·lels de *PRDM9* es va amplificar el domini de dits de zinc d'aquest gen format per entre 8 i 15 dits de zinc de 84 pb (**Figura 8**). L'al·lel A,

caracteritzat com el més freqüent en la població europea (BERG ET AL. 2010), està constituït per 13 dits de zinc (**Figura 21**). En total, es van genotipar 19 individus control (incloent individus sense descendència afectada per la SDG/VCF i individus amb descendència afectada però no transmissors) i 8 individus transmissors de la SDG/VCF.

L'anàlisi del genotip de *PRDM9* requereix la realització d'una PCR i una posterior seqüenciació sobre el producte amplificat. Els encebadors utilitzats en la PCR van ser PN0.6F i PN2.5R (PN0.6F: TGAGGTTACCTAGTCTGGCA, PN2.5R: ATAAGGGGTCAGCAGACTTC) (BERG ET AL. 2010). Les reaccions es preparen amb un volum final de 40 µl: 4 µl tampó 10x (MgCl<sub>2</sub> 20 mM), 3,2 µl dNTPs (2,5 mM cada un), 1 µl PN0.6F (10 mM), 1 µl PN2.5R (10 mM), 3% DMSO, 1,25 U BioTaq (Bioline), 100 ng DNA i l'ajustament corresponent d'aigua desionitzada.

Les amplificacions es van realitzar seguint les següents condicions: 3 minuts a 95 °C; 45 cicles de 30 segons a 95 °C, 30 segons a 62 °C i 105 segons a 72 °C. Els productes de la PCR es van visualitzar en un gel d'agarosa 1% per tal de detectar la presència d'individus amb al·lels heterozigots de longitud.

Els productes de la PCR es van purificar amb el kit *PCR DNA purification* (Thermo Scientific) i sotmesos a una seqüenciació bidireccional de Sanger mitjançant els





encebadors PN1.2F i PN2.4R (PN1.2F: TGAATCCAGGGAACACAGGC i PN2.4R: GCAAGTGTGTGGTGACCACA) (BERG ET AL. 2010). La lectura del DNA amplificat es va realitzar amb el seqüenciador BigDye Terminator v3.1 Cycle (Applied Biosystems) i l'analitzador de seqüències ABI 3130XL Genetic Analyzer (Applied Biosystems).

Per tal d'assignar l'al·lel de *PRDM9* corresponent a cada individu, les seqüències de dits de zinc es van alinear i comparar amb seqüències publicades (BAUDAT ET AL. 2010; BERG ET AL. 2010; HUSSIN ET AL. 2013; ALEMANY-SCHMIDT ET AL. 2014). A partir de les combinacions de dits de zinc es van poder definir els genotips corresponents.

La rellevància de les variants al·lèliques de *PRDM9* en el risc de transmetre la SDG/VCF es va determinar comparant les freqüències dels genotips dels individus control i els individus transmissors de la SDG/VCF. La freqüència de cada al·lel es va comparar entre les dues poblacions descrites mitjançant un test exacte de Fisher. Les diferències es van considerar significatives quan  $p < 0,05$ .



# RESULTATS

- 4.1 REORGANITZACIONS EN  
ESPERMATOZOIDES
- 4.2 CARACTERÍSTIQUES  
GENÈTIQUES



## 4.1 REORGANITZACIONS EN ESPERMATOZOIDES

### DELECCIONS I DUPLICACIONS DE LA REGIÓ 22q11.2 EN PARES SDG/VCF

Les freqüències de deleccions i duplicacions de la regió 22q11.2 van ser analitzades en un total de 90.776 espermatozoides de nou pares amb descendència afectada per la SDG/VCF. La freqüència mitjana de deleccions ( $\pm$ SEM) va ser de 0,28% ( $\pm$ 0,05) amb un rang de 0,13 a 0,57%. La freqüència mitjana de duplicacions ( $\pm$ SEM) va ser de 0,11% ( $\pm$ 0,02) amb un rang de 0,05 a 0,20% (**Taula 7**).

■ **Taula 7.** Resultats de FISH en espermatozoides de deleccions i duplicacions en homes amb descendència afectada per la SDG/VCF. Entre parèntesis s'indiquen els valors en %.

Cas	Edat	Normal	del 22q11.2	dup 22q11.2	del+dup	Altres <sup>a</sup>	Total
DG1	39	10000 (99,19)	24 (0,24)	11 (0,11)	35 (0,35)	47 (0,47)	10082
DG2	48	10000 (98,40)	57 (0,57)*.c	20 (0,20) <sup>d</sup>	77 (0,77)*	86 (0,86)	10163
DG3	-	9927 (99,09)	19 (0,19)	13 (0,13)	32 (0,32)	59 (0,59)	10018
DG4	43	10000 (99,30)	15 (0,15)	7 (0,07)	22 (0,22)	49 (0,49)	10071
DG5	36	10096 (98,87)	52 (0,52)*.c	12 (0,12) <sup>d</sup>	64 (0,64)*	51 (0,51)	10211
DG6	30	10000 (99,24)	28 (0,28) <sup>c</sup>	5 (0,05) <sup>d</sup>	33 (0,33)	43 (0,43)	10076
DG7	39	10000 (99,34)	18 (0,18)	7 (0,07)	25 (0,25)	41 (0,41)	10066
DG8	33	9948 (99,31)	13 (0,13)	10 (0,10)	23 (0,23)	46 (0,46)	10017
DG9	37	9930 (98,59)	25 (0,25)	16 (0,16)	41 (0,41)	31 (0,31)	10072
% $\pm$ SEM		99,04 $\pm$ 0,11	0,28 $\pm$ 0,05	0,11 $\pm$ 0,02	0,39 $\pm$ 0,06	0,50 $\pm$ 0,05	
% $\pm$ SEM <sup>b</sup>		99,08 $\pm$ 0,09	0,17 $\pm$ 0,02	0,12 $\pm$ 0,03	0,27 $\pm$ 0,05	0,64 $\pm$ 0,05	

<sup>a</sup> Disomies, diploïdies i nul·lisomies.

<sup>b</sup> Dades de la població control obtingudes per Molina et al. (2011).

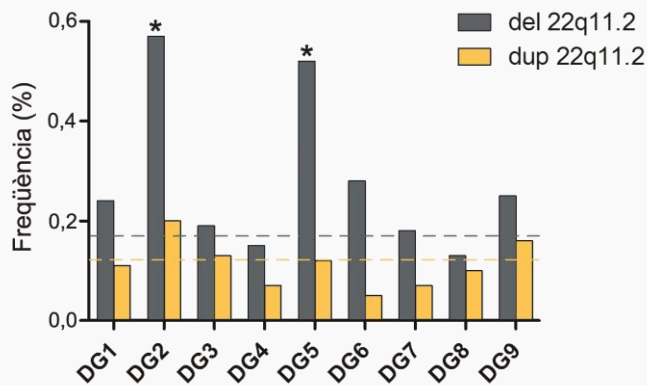
\* Incrementos significatius respecte la població control ( $p < 0,01$ ).

<sup>c,d</sup> Distorsió significativa de la ràtio 1:1 del:dup ( $p < 0,01$ ).

No es van observar diferències significatives entre la freqüència mitjana de deleccions, duplicacions i del+dup de la població de pares SDG/VCF i les dades de controls (Molina et al. 2011) (Test de Mann-Whitney;  $p = 0,111$ ;  $p = 1,000$  i  $p = 0,191$  respectivament) (**Taula 7**).

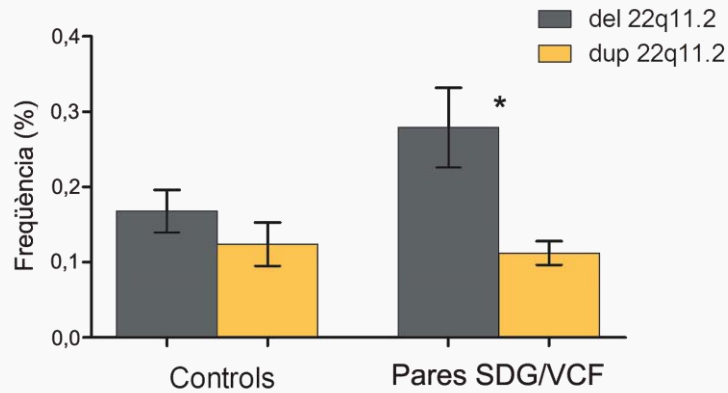
A nivell individual, es van detectar increments significatius de delecions i del+dup en dos dels nou pares SDG/VCF analitzats (Test de khi-quadrat; DG2 i DG5  $p < 0,01$ ) (**Figura 22, Taula 7**). Els risc de generar delecions en aquests individus en relació a la població control es va estimar en una oportunitat relativa de 2,03 (IC 95% = 1,63-2,54).

D'altra banda, no es va observar una correlació entre l'edat dels pares SDG/VCF i les freqüències de delecions (Correlació de Spearman;  $R = 0,084$ ;  $p = 0,840$ ), duplicacions (Correlació de Spearman;  $R = 0,380$ ;  $p = 0,360$ ) i del+dup (Correlació de Spearman;  $R = 0,144$ ;  $p = 0,752$ ).



**Figura 22.** Freqüències de delecions i duplicacions 22q11.2 en homes amb descendència afectada per la SDG/VCF. Les línies discontinües representen les mitjanes de delecions i duplicacions en controls. Els asteriscs indiquen els increments significatius de delecions respecte la població control.

Pel que fa a la participació dels mecanismes de la NAHR inter-cromàtide i intra-cromàtide en la generació d'anomalties, no es va observar una correlació significativa entre les freqüències de delecions i duplicacions en homes amb descendència afectada per la SDG/VCF (Correlació de Spearman;  $R = 0,477$ ;  $p = 0,194$ ). A més a més, es va detectar un increment significatiu de la freqüència mitjana de delecions en comparació a la freqüència mitjana de duplicacions (Test de Wilcoxon;  $p = 0,004$ ) (**Figura 23**). Les comparacions individuals van mostrar increments significatius de delecions 22q11.2 en tres dels nou casos analitzats (Test de khi-quadrat; DG2, DG5 i DG6  $p < 0,01$ ) (**Taula 7**). Els individus DG2 i DG5 també mostraven increments significatius de delecions i de del+dup 22q11.2 en comparació a les dades control.



■ **Figura 23.** Frequències mitjanes de delecions i duplicacions 22q11.2 en la població control (MOLINA ET AL. 2011) i en la població d'homes amb descendència afectada per la SDG/VCF. Les barres d'error representen l'error estàndard de la mitjana (SEM) i l'asterisc mostra diferències significatives entre la freqüència de delecions i duplicacions.

L'origen parental es va determinar en vuit de les nou famílies SDG/VCF. En sis famílies, l'origen de la deleción en els cromosoma 22 era patern; mentre que en els dos casos restants, la deleción tenia un origen matern (**Taula 8**). Cal destacar que en els pares SDG/VCF en els quals s'havia diagnosticat increments de delecions i del+dup 22q11.2 en espermatozoides, l'origen de la deleción que afectava a la descendència era patern.

#### DELECIONS I DUPLICACIONS DE LA REGIÓ 7q11.23, 15q11q13 I 22q11.2 EN PARES TRANSMISSORS DE LA SDG/VCF

Es van analitzar les freqüències de delecions i duplicacions de les regions crítiques 7q11.23 (**Taula 9**) i 15q11q13 (**Taula 10**) en la població formada pels 6 pares transmissors de la SDG/VCF en la qual s'havien analitzat les freqüències de delecions i duplicacions de la regió 22q11.2 (**Taula 7**).

En l'anàlisi de la susceptibilitat a la NAHR a nivell de població, no es van observar diferències significatives entre les freqüències de delecions, duplicacions i del+dup de cada regió crítica entre pares transmissors de la SDG/VCF i la població control (Molina et al. 2011) (Test de Mann Whitney;  $p > 0,05$ ).

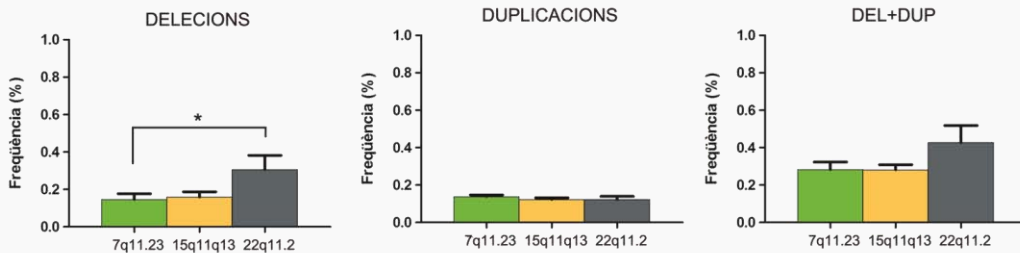
■ **Taula 8.** Resultats de l'avaluació de l'origen parental: pare (P), mare (M) i fill afectat (F). Els valors en negreta indiquen els microsatèl·lits informatius.

Cas	D22S1638	D22S1648	D22S264	D22S311	D22S303	Origen parental	Mida deleció
DG1P	114/116	174/174	185/195	<b>251/251</b>	220/222		
DG1M	116/116	174/174	195/195	<b>250/256</b>	222/223		
DG1F	116	174	195	<b>256</b>	220/223	Paternal	3 Mb
DG2P	116/116	170/172	<b>195/203</b>	<b>250/250</b>	212/222		
DG2M	116/116	172/172	<b>201/201</b>	<b>250/257</b>	212/222		
DG2F	116	172	<b>201</b>	<b>257</b>	222/222	Paternal	3 Mb
DG3P	<b>99/105</b>	174/178	193/199	<b>256/256</b>	220/223		
DG3M	<b>116/116</b>	174/174	193/199	<b>256/259</b>	222/224		
DG3F	<b>116</b>	174	193	<b>259</b>	220/224	Paternal	3 Mb
DG5P	116/120	174/174	<b>195/201</b>	250/256	224/225		
DG5M	116/118	174/174	<b>188/203</b>	256/256	224/226		
DG5F	-	174	<b>203</b>	256	224/226	Paternal	3 Mb
DG6P	<b>105/116</b>	<b>174/174</b>	<b>187/191</b>	250/250	222/222		
DG6M	<b>109/109</b>	<b>173/178</b>	<b>185/193</b>	250/254	222/222		
DG6F	<b>116</b>	<b>174</b>	<b>191</b>	250	222/222	Maternal	3 Mb
DG7P	<b>109/120</b>	174/174	185/199	<b>250/250</b>	222/224		
DG7M	<b>116/114</b>	174/174	199/205	<b>254/257</b>	212/224		
DG7F	<b>116</b>	174	199	<b>254</b>	212/224	Paternal	3 Mb
DG8P	<b>114/114</b>	174/174	187/199	252/256	213/223		
DG8M	<b>109/116</b>	174/174	199/201	250/256	223/225		
DG8F	<b>116</b>	174	199	256	223/223	Paternal	NA
DG9P	<b>116/116</b>	<b>174/174</b>	<b>199/201</b>	<b>250/256</b>	212/222		
DG9M	<b>111/122</b>	<b>170/171</b>	<b>187/191</b>	<b>250/254</b>	222/222		
DG9F	<b>116</b>	<b>174</b>	<b>201</b>	<b>256</b>	212/222	Maternal	3 Mb

NA = no avaluable

D'altra banda, la comparació de les mitjanes de delecions, duplicacions i del+dup entre les tres regions estudiades va demostrar diferències entre les mitjanes de delecions de les tres regions crítiques (Test de Friedman;  $p = 0,008$ ). Amb l'objectiu de determinar entre

quines regions es produïen aquestes diferències es van realitzar comparacions dos a dos. Es van observar diferències entre les freqüències mitjanes de delecions de les regions 7q11.23 i 22q11.2 (Test de Wilcoxon;  $p = 0,031$ ) (**Figura 24**).



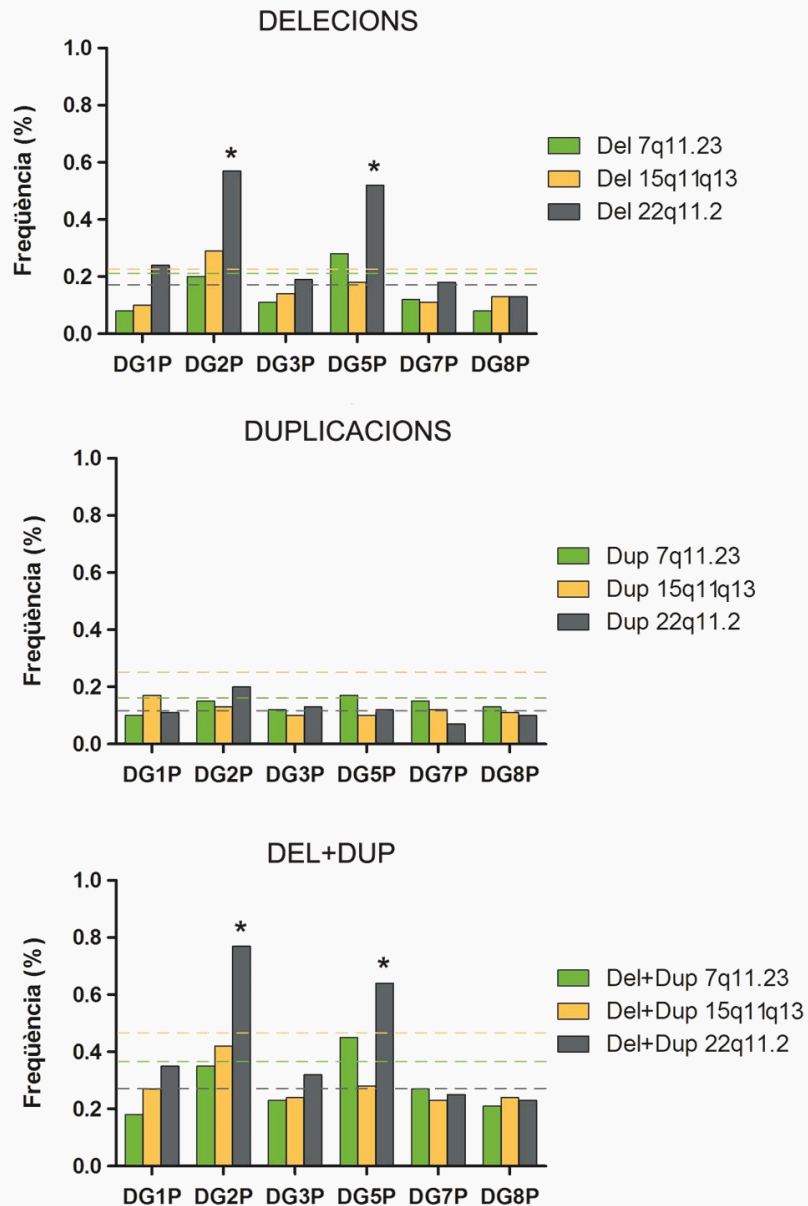
**Figura 24.** Freqüències mitjanes de delecions, duplicacions i del+dup de les regions 7q11.23, 15q11q13 i 22q11.2 en parets transmissors de la SDG/VCF. Les barres d'error mostren l'error estàndard de la mitjana (SEM). L'asterisc identifica diferències significatives en la mitjana de reorganitzacions entre les regions crítiques.

A nivell individual, els parets que havien mostrat un increment significatiu de delecions i del+dup per la regió crítica 22q11.2 respecte la població control, no van obtenir increments significatius de reorganitzacions per les regions 7q11.23 i 15q11q13 (Test de khi-quadrat,  $p > 0,01$ ). D'altra banda, els parets que no havien mostrat un increment de reorganitzacions per la regió crítica 22q11.2 tampoc van mostrar cap increment en les altres regions analitzades (**Figura 25**).

En l'anàlisi dels mecanismes de la NAHR que originen les reorganitzacions de cada regió crítica, a nivell poblacional es va observar com les freqüències de delecions i duplicacions de la regió 7q11.23 presentaven una correlació positiva (Correlació de Spearman;  $R = 0,882$ ;  $p = 0,033$ ). En canvi, no es va observar aquesta correlació per les regions 15q11q13 (Correlació de Spearman;  $R = -0,406$ ;  $p = 0,419$ ) i 22q11.2 (Correlació de Spearman;  $R = 0,771$ ;  $p = 0,103$ ). Les mitjanes de delecions i duplicacions no van mostrar diferències significatives en les regions 7q11.23 i 15q11q13 (Test de Wilcoxon;  $p = 1,000$  i  $p = 0,313$ ; respectivament), mentre que sí que es van observar diferències significatives entre la mitjana de delecions i duplicacions de la regió 22q11.2 (Test de Wilcoxon;  $p = 0,031$ ). A part dels dos parets prèviament identificats amb diferències significatives entre la proporció de delecions i duplicacions 22q11.2, a nivell individual



no es van observar altres distorsions de la proporció en la regió 7q11.23 i 15q11q13 (Test de khi-quadrat;  $p > 0,05$ ) (Taula 9 i 10).



■ **Figura 25.** Freqüències de delecions, duplicacions i del+dup de les tres regions analitzades. L'asterisc identifica els individus amb increments significatius respecte a la població control. La línia discontinua representa la freqüència mitjana descrita en la població control (MOLINA ET AL. 2011).

■ **Taula 9.** Resultats de FISH en espermatozoides de deleccions i duplicacions 7q11.23 en pares transmissors de la SDG/VCF. Entre parèntesis s'indiquen els valors en %.

Cas	Normal	del 7q11.23	dup 7q11.23	del+dup	Altres <sup>a</sup>	Total
DG1	10035 (99,46)	8 (0,08)	10 (0,10)	18 (0,18)	46 (0,46)	10089
DG2	10049 (99,20)	20 (0,20)	15 (0,15)	35 (0,35)	61 (0,60)	10130
DG3	10111 (99,46)	11 (0,11)	12 (0,12)	23 (0,23)	46 (0,45)	10168
DG5	10321 (99,27)	28 (0,28)	14 (0,13)	42 (0,41)	48 (0,46)	10397
DG7	10112 (99,19)	12 (0,12)	15 (0,15)	27 (0,27)	42 (0,41)	10166
DG8	10110 (99,56)	8 (0,08)	13 (0,13)	21 (0,21)	37 (0,36)	10155
% ± SEM	99,36 ± 0,06	0,15 ± 0,03	0,13 ± 0,01	0,28 ± 0,04	0,46 ± 0,04	
% ± SEM <sup>b</sup>	98,80 ± 0,23	0,21 ± 0,03	0,17 ± 0,02	0,37 ± 0,04	0,70 ± 0,01	

<sup>a</sup> Disomies, diploidies i nul·lisomies.

<sup>b</sup> Dades de la població control obtingudes per in Molina et al. (2011).

■ **Taula 10.** Resultats de FISH en espermatozoides de deleccions i duplicacions 15q11q13 en pares transmissors de la SDG/VCF. Entre parèntesis s'indiquen els valors en %.

Cas	Normal	del 15q11q13	dup 15q11q13	del+dup	Altres <sup>a</sup>	Total
DG1	9939 (99,13)	10 (0,10)	17 (0,17)	27 (0,27)	60 (0,60)	10026
DG2	9970 (99,27)	29 (0,29)	13 (0,13)	42 (0,42)	31 (0,31)	10043
DG3	9980 (98,83)	14 (0,14)	10 (0,10)	24 (0,24)	94 (0,93)	10098
DG5	10048 (99,38)	18 (0,18)	10 (0,10)	28 (0,28)	35 (0,35)	10111
DG7	10050 (99,09)	11 (0,11)	12 (0,12)	23 (0,23)	69 (0,68)	10142
DG8	9942 (99,21)	13 (0,13)	11 (0,11)	24 (0,24)	55 (0,55)	10021
% ± SEM	99,15 ± 0,08	0,16 ± 0,03	0,12 ± 0,01	0,28 ± 0,03	0,57 ± 0,09	
% ± SEM <sup>b</sup>	99,17 ± 0,07	0,22 ± 0,03	0,25 ± 0,04	0,47 ± 0,07	0,38 ± 0,07	

<sup>a</sup> Disomies, diploidies i nul·lisomies.

<sup>b</sup> Dades de la població control obtingudes per in Molina et al. (2011).

## 4.2 CARACTERÍSTIQUES GENÈTIQUES

### BLOCS DE SEQÜÈNCIES EN TÀNDEM DELS LCR22-2 I LCR22-4

Es van establir rangs de normalitat de blocs de seqüències en tàndem a partir dels valors observats en sis individus control. Pel LCR22-2, els rangs de normalitat van ser de 2 – 6 còpies per L9, 3 – 9 còpies per K3, 1 – 4 còpies per B22 i 1 – 2 còpies per L21. Pel LCR22-4 els rangs de normalitat van ser de 1 – 2 còpies per L9, 1 – 3 còpies per K3, 1 – 2 còpies per B22, 2 – 2 còpies per M12 i 2 – 3 còpies per I18 (**Taula 11**).

■ **Taula 11.** Blocs de seqüències en tàndem dels LCR22-2 i LCR22-4. Els dos haplotips establerts per cada individu es separen per una contrabarra. Els números en negreta indiquen els valors fora del rang de normalitat.

Mostres	LCR22-2				LCR22-4				
	K3	L9	B22	L21	K3	L9	B22	M12	I18
Controls									
C1	5 \ 7	5 \ 6	2 \ 3	1 \ 2	2 \ 3	2 \ 2	1 \ 2	2 \ 2	2 \ 2
C2	7 \ 9	5 \ 10*	3 \ 5*	1 \ 2	2 \ 2	2 \ 2	1 \ 2	2 \ 2	2 \ 3
C3	3 \ 5	3 \ 5	1 \ 2	1 \ 2	1 \ 2	1 \ 2	1 \ 1	2 \ 2	2 \ 2
C4	3 \ 5	3 \ 5	2 \ 3	1 \ 1	2 \ 2	2 \ 2	1 \ 1	2 \ 2	2 \ 2
C5	3 \ 3	2 \ 5	2 \ 4	1 \ 2	1 \ 2	1 \ 2	2 \ 2	2 \ 2	2 \ 2
C6	3 \ 5	2 \ 4	1 \ 2	1 \ 1	1 \ 2	2 \ 2	1 \ 2	2 \ 2	2 \ 2
Rang de normalitat <sup>a</sup>	3-9	2-6	1-4	1-2	1-3	1-2	1-2	2-2	2-3
Pares SDG/VCF <sup>b</sup>									
DG2P	<b>2 \ 2</b>	3 \ 6	2 \ 3	1 \ 2	2 \ 2	2 \ <b>3</b>	1 \ 2	2 \ 2	2 \ 2
DG5P	5 \ 8	<b>8 \ 8</b>	2 \ 3	1 \ 2	2 \ 2	2 \ <b>3</b>	1 \ 2	2 \ 2	2 \ 2

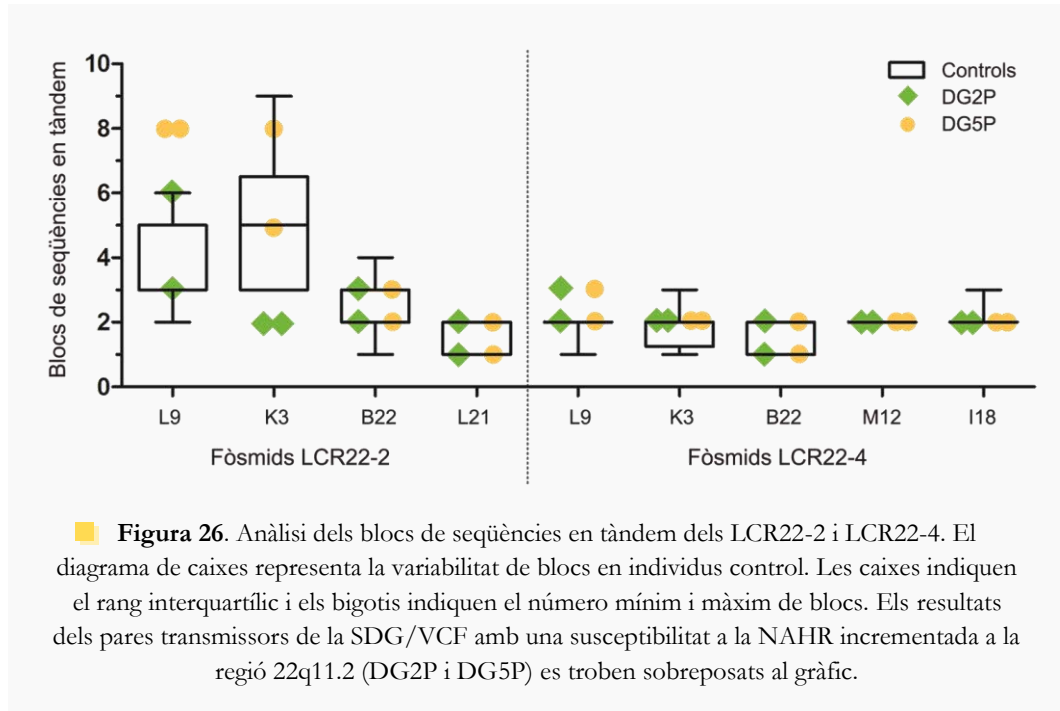
\* Dada atípica (Test de Grubbs).

<sup>a</sup> Rangs establerts a partir dels valors mínims i màxims en individus control exclouent les dades atípiques.

<sup>b</sup> Pares transmissors de la SDG/VCF amb increments de delecions 22q11.2 en espermatozoides.

D'altra banda, es van analitzar dos pares transmissors de la SDG/VCF amb una susceptibilitat a la NAHR provada (DG2P i DG5P) (veure apartat 4.1). En ambdós casos es van detectar números de blocs fora dels rangs de normalitat. L'individu DG2P presentava dues còpies de blocs K3 a LCR22-2 en els dos haplotips i tres còpies de blocs L9 a LCR22-4 en un dels haplotips. Així mateix, l'individu DG5P presentava vuit

còpies de blocs L9 a LCR22-2 en ambdós haplotips i tres còpies de blocs L9 a LCR22-4 en un haplotip (Taula 11; Figura 26).



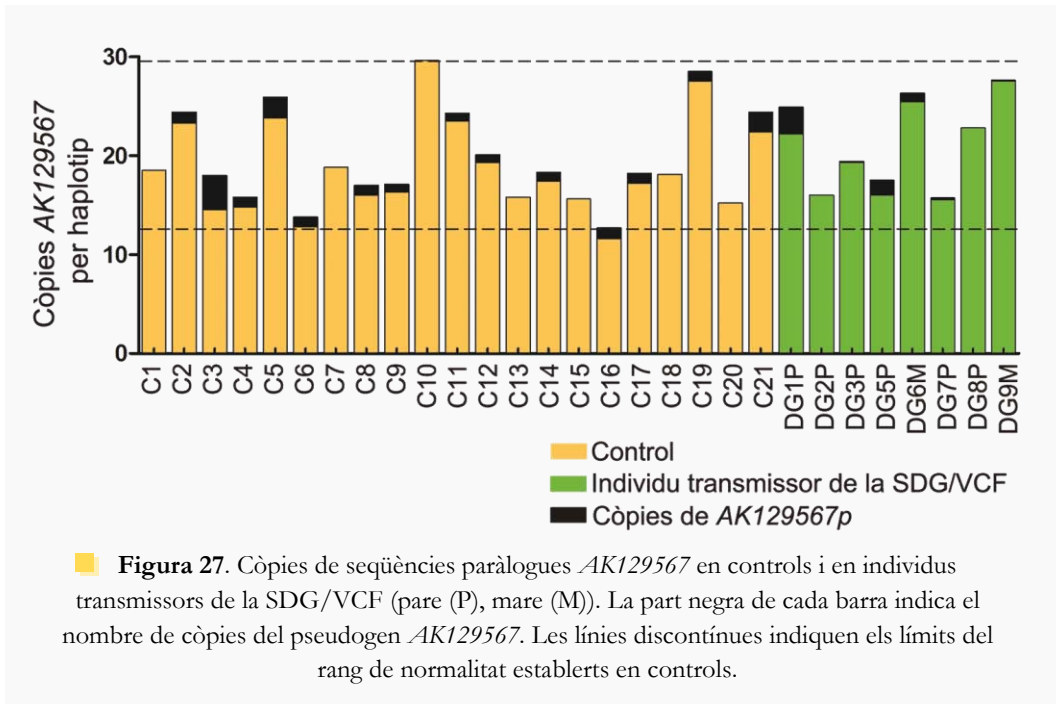
■ **Figura 26.** Anàlisi dels blocs de seqüències en tàndem dels LCR22-2 i LCR22-4. El diagrama de caixes representa la variabilitat de blocs en individus control. Les caixes indiquen el rang interquartílic i els bigotis indiquen el número mínim i màxim de blocs. Els resultats dels pares transmissors de la SDG/VCF amb una susceptibilitat a la NAHR incrementada a la regió 22q11.2 (DG2P i DG5P) es troben sobreposats al gràfic.

## SEQÜÈNCIES PARÀLOGUES DELS LCR22

En les mostres prèviament analitzades per *fiber*-FISH es va analitzar el número de còpies dels gens *AK129567* i *AK302545* coberts pels fòsmids L9 i K3, respectivament. Els resultats obtinguts per la tècnica de ddPCR no van correlacionar amb els de la *fiber*-FISH (Correlació de Spearman;  $r = 0,146$  i  $0,386$ ;  $p = 0,703$  i  $0,360$ ; respectivament).

Els rangs de normalitat de còpies per *AK129567* va ser de  $12,7 - 29,6$  ( $n=21$ ) (Figura 27), mentre que per *AK302545* el rang van ser de  $10,6 - 21$  ( $n=6$ ) (Figura 28).

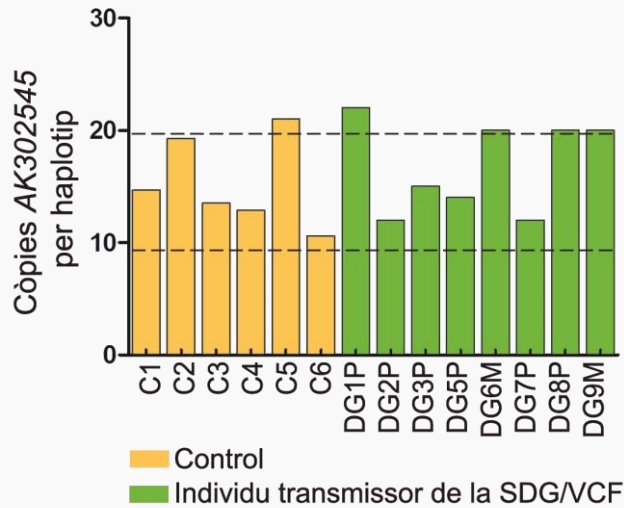
La mitjana de còpies de *AK129567* en controls ( $n=21$ ) i en individus transmissors de la SDG/VCF ( $n=8$ ) va ser de  $19,54$  i  $21,28$  respectivament; mentre que la mitjana de còpies de *AK302545* va ser de  $15,33$  en controls ( $n=6$ ) i de  $17,46$  en individus transmissors de la SDG/VCF ( $n=8$ ). No es van observar diferències significatives entre les mitjanes de còpies entre els dos grups (Test de Mann Whitney;  $p = 0,420$  i  $0,477$  respectivament).



A nivell individual, únicament el pare DG1 va obtenir còpies de *AK302545* per sobre del rang de normalitat observat en controls (22 còpies) (**Figura 27**). Els individus en els quals s'havien detectat CNV mitjançant *fiber*-FISH no es van observar diferències significatives.

D'altra banda, es van detectar dos tipus d'amplicons per *AK129567* que es diferenciaven per la intensitat de fluorescència. D'acord amb la relació inversament proporcional entre la fluorescència emesa i la mida dels amplicons descrita per Laurie et al. (2013), els amplicons amb una intensitat de fluorescència inferior tenen una longitud major. Aquests amplicons més grans es van atribuir a seqüències d'un possible pseudogen de *AK129567* (*AK129567p*) no descrit en les bases de dades (Servei tècnic de BioRad; comunicació personal) (**Figura 27**).

En un 66,67% dels individus control (n=21) es va observar la presència de dues o més còpies per genoma de *AK129567p*. Mentre que en individus transmissors de la SDG/VCF (n=8) tan sols es van detectar còpies de *AK129567p* en un 37,5% dels casos (**Figura 29**). No obstant, les diferències observades en la freqüència de *AK129567p*

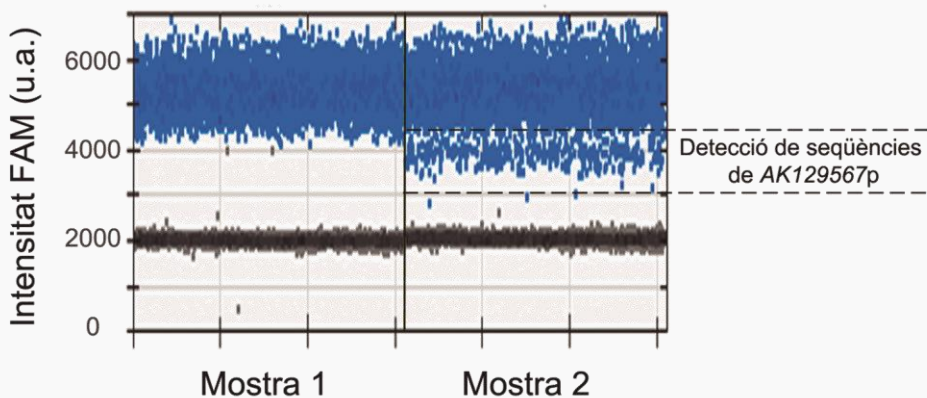


**Figura 28.** Còpies de seqüències paràlogues *AK302545* en controls i en individus transmissors de la SDG/VCF (pare (P), mare (M)). Les línies discontinües indiquen els límits del rang de normalitat establerts en controls.

entre individus control i transmissors de la SDG/VCF no van assolir diferències estadísticament significatives (Test de Fisher;  $p = 0,218$ ).

## INVERSIONS 22q11.2

Les inversions 22q11.2 es van avaluar en cinc individus control i en una població de vuit parets transmissors de la SDG/VCF. En total es van analitzar 1354 haplotips



**Figura 29.** Quantificació de *AK129567* mitjançant ddPCR. En la mostra 1 s'observa un sol tipus d'amplicó mentre que en la mostra 2 s'observa un segon tipus d'amplicó amb una intensitat de fluorescència menor corresponent a *AK129567p*.

informatius, 512 en individus control i 842 en individus transmissors de la SDG/VCF (**Taula 12**).

La freqüència mitjana d'haplotips normals va ser del 96,08% en individus control (rang de 94,12% a 98,06%) i de 96,92% en parens transmissors de la SDG/VCF (rang de 94,55% a 99,02%). La freqüència mitjana de l'haplotip amb la inversió va ser de 1,57% en controls (rang de 0,96% a 2,94%) i de 1,56% en transmissors de la SDG/VCF (rang de 0% a 2,91%). També es va detectar un percentatge negligible d'haplotips informatius amb una associació de sondes inesperada que va ser atribuïda a errors d'hibridació (**Taula 12**).

■ **Taula 12.** Anàlisi d'inversions 22q11.2 en controls i individus transmissors de la SDG/VCF (pare (P), mare (M)). Entre parèntesis s'indiquen els valors en %.

Mostres	Total nuclis	Haplotips informatius	Haplotips			Genotip
			N	inv	Altres <sup>a</sup>	
Controls						
C3	87	102 (58,62)	98 (96,08)	1 (0,98)	3 (2,94)	N/N
C7	69	101 (73,19)	96 (95,05)	2 (1,98)	3 (2,97)	N/N
C9	77	102 (66,23)	96 (94,12)	3 (2,94)	3 (2,94)	N/N
C11	77	104 (67,53)	101 (97,12)	1 (0,96)	2 (1,92)	N/N
C14	75	103 (68,67)	101 (98,06)	1 (0,97)	1 (0,97)	N/N
% ±SEM		66,85 ±2,37	96,08 ±0,70	1,57 ±0,40	2,35 ±0,40	
Parens transmissors de la SDG/VCF						
DG1P	77	103 (66,88)	100 (97,09)	3 (2,91)	0	N/N
DG2P	112	119 (53,13)	116 (97,48)	0	3 (2,52)	N/N
DG3P	86	102 (59,30)	101 (99,02)	1 (0,98)	0	N/N
DG5P	103	110 (53,40)	104 (94,55)	3 (2,73)	3 (2,73)	N/N
DG6M	78	103 (66,03)	101 (98,06)	2 (1,94)	0	N/N
DG7P	93	102 (54,84)	101 (99,02)	0	1 (0,98)	N/N
DG8P	93	102 (54,84)	97 (95,10)	2 (1,96)	3 (2,94)	N/N
DG9M	80	101 (63,13)	96 (95,05)	2 (1,98)	3 (2,97)	N/N
% ± SEM		57,96 ±2,02	96,76 ±0,64	1,51 ±0,40	1,73 ±0,50	

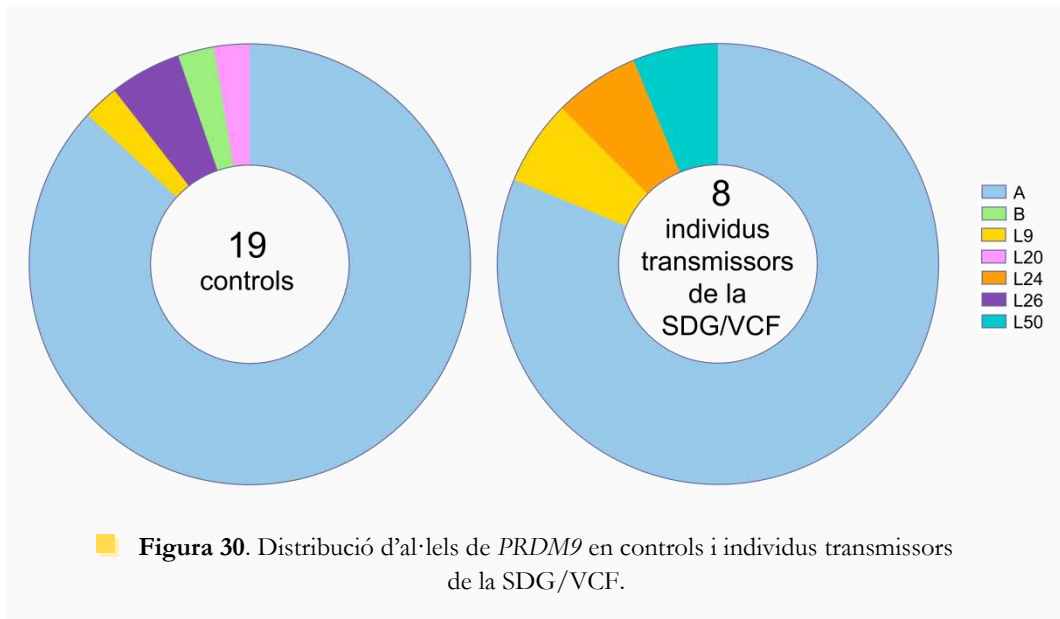
<sup>a</sup> Associacions no esperades.

N = normal; inv = inversió

En tots els casos, la freqüència de la inversió va ser significativament inferior a la freqüència esperada en una situació d'heterozigosi (Test de khi-quadrat;  $p = <0,05$ ) (**Taula 12**). En conseqüència, en tots els individus es va descriure un genotip normal /normal (N/N) i no es van observar diferències entre la freqüència de inversions de la població control i la població d'individus transmissors de la SDG/VCF (Test de Mann Whitney;  $p = 0,941$ ).

#### VARIANTS AL·LÈLIQUES DE PRDM9

La seqüenciació del domini de dits de zinc de *PRDM9* es va realitzar en 19 individus control i vuit individus transmissors de la SDG/VCF (**Taula 13**). L'al·lel A es va observar en el 86,64% de la població control analitzada i en el 81,25% de la població problema. Es va identificar un al·lel no descrit prèviament amb més de 13 dits de zinc (L50), tot i que degut a limitacions tècniques no es va poder descriure tot el domini de dits de zinc (**Figura 30**).



Les freqüències al·lèliques observades en controls i en individus transmissors de la SDG/VCF no van mostrar diferències significatives (Test de Fisher;  $p = 0,088$ ). Els pares DG2 i DG5 que havien mostrat increments en la susceptibilitat a la NAHR de la regió 22q11.2 van presentar els genotips A/L50 i A/A, respectivament (**Taula 13**).



■ **Taula 13.** Genotipatge de PRDM9 en controls i individus transmissors de la SDG/VCF (pare (P), mare (M)).

	Mostres	Domini dits de zinc haplotip 1	Domini dits de zinc haplotip 2	Al·lels	
Controls	C7	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C8	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C9	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C10	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C11	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C12	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C13	ABCDDECFGHFIJ	ABCDDECFGPFQJ	A/L9	
	C14	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C15	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C16	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C17	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C19	ABCDDECFGHFIJ	ABCDDECFGHFQJ	A/L26	
	C20	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C21	ABCDDECFGHFIJ	ABCDDECFGKFQJ	A/L20	
	C22	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C23	ABCDDECFGHFIJ	ABCDDECFGHFQJ	A/L26	
	C24	ABCDDECFGHFIJ	ABCDCCFGHFIJ	A/B	
	C25	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	C26	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	
	Pares transmissors de la SDG/VCF	DG1P	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A
		DG2P	ABCDDECFGHFIJ	ABCCCC?FGHFIJ <sup>a</sup>	A/L50
		DG3P	ABCDDECFGHFIJ	ABCDDECFGPFQJ	A/L9
		DG5P	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A
		DG6M	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A
		DG7P	ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A
		DG8P	ABCDDECFGHFIJ	ABCDDECFTPFQJ	A/L24
DG9M		ABCDDECFGHFIJ	ABCDDECFGHFIJ	A/A	

<sup>a</sup> Limitacions tècniques impedeixen la seqüenciació completa de tots els dits de zinc de l'al·lel.



# DISCUSSIÓ

- 5.1 NAHR EN PARES TRANSMISSORS  
SDG/VCF
- 5.2 FACTORS DE PREDISPOSICIÓ A LA  
SDG/VCF



## 5.1 NAHR EN PARES TRANSMISSORS SDG/VCF

### Consideracions metodològiques de l'anàlisi de la NAHR

La FISH en nuclis descondensats d'espermatozoides s'ha aplicat àmpliament en estudis citogenètics (MARTIN 2005). L'anàlisi microscòpica es realitza a partir de la identificació de la presència o absència de senyals d'hibridació (BOSCH ET AL. 2003; SLOTER ET AL. 2007; BOSCH ET AL. 2001; TEMPLADO ET AL. 2011; MOLINA ET AL. 2011; ANTON ET AL. 2004; GODO ET AL. 2015) o de canvis en la seva disposició (MOLINA ET AL. 2012A; SLOTER ET AL. 2000). En les diferents situacions en què s'ha aplicat, el punt de partida és el disseny d'una combinació de sondes de DNA específiques i l'aplicació d'uns criteris estrictes de valoració. La tècnica permet inferir la constitució cromosòmica de cada espermatozoide amb una alta sensibilitat i especificitat.

L'estudi de reorganitzacions cromosòmiques en regions flanquejades per LCR mitjançant la FISH en espermatozoides ha permès avaluar la susceptibilitat a la NAHR en diferents individus, així com determinar el tipus de recombinació no al·lèlica predominant. Per a l'anàlisi de les freqüències de delecions i duplicacions de les regions 7q11.23, 15q11q13 i 22q11.2, es van utilitzar combinacions de tres sondes: una sonda locus específica, una sonda de control d'hibridació i una sonda de control de ploïdia.

La presència/absència dels senyals de la sonda de control d'hibridació i la sonda locus específica permet diferenciar les delecions de la regió crítica (senyal control d'hibridació) de les nul·lisomies (cap senyal de les dues sondes) i les duplicacions (un senyal de la sonda de control d'hibridació en presència de dos senyals de la sonda locus específica) de les disomies (dos senyals de cadascuna de les dues sondes). En el cas particular de les duplicacions, els dos senyals observats com a producte de la duplicació de cada regió crítica es van poder distingir gràcies a què la distància estimada entre els locus duplicats (3 Mb per 7q11.23, 10 Mb per 15q11q13 i 6 Mb per 22q11.2) en tots els casos era superior al límit de resolució de la FISH en nuclis interfàsics (50 kb) (Heng et al. 2005).

Per altra banda, la presència de la sonda de control de ploïdia permet diferenciar les aneuploïdies del cromosoma que conté la regió crítica estudiada de les diploïdies (dos

senyals de les tres sondes). També s'ha utilitzat per identificar els espermatozoides amb una descondensació insuficient per possibilitar la hibridació de les sondes (cap senyal de les tres sondes).

### Susceptibilitat a la NAHR

La població de pares transmissors de la SDG/VCF no va mostrar una susceptibilitat incrementada a la generació de delecions i duplicacions de la regió 22q11.2 en comparació a les dades de la població control (MOLINA ET AL. 2011). Tot i així, a nivell individual es van observar increments significatius de delecions 22q11.2 en dos pares transmissors de la SDG/VCF.

Els pares transmissors de la SDG/VCF en els quals es va estimar un risc de transmissió incrementat presentaven una probabilitat de generar espermatozoides portadors de delecions 22q11.2 dos cops superior a la de la població control (Oportunitat relativa 2,03). Malgrat que l'increment de reorganitzacions és moderat i que el nombre de mostres analitzades és limitat, els resultats publicats sobre anomalies en gàmetes de pares transmissors d'altres síndromes donen suport a què l'increment de delecions podria ser rellevant en l'avaluació del risc de transmissió. En aquest sentit, s'han identificat pares transmissors de la síndrome de Prader-Willi que presentaven un increment significatiu de delecions 15q11q13 en espermatozoides (MOLINA ET AL. 2010). La presència d'individus amb un increment de gàmetes portadores de les anomalies cromosòmiques que afecten a la descendència no és un fenomen aïllat en els trastorns genòmics. S'han observat increments disomies del cromosoma 21 en pares amb descendència afectada per la síndrome de Down (BLANCO ET AL. 1998) o aneuploidies dels cromosomes sexuals en pares amb descendència afectada per la síndrome de Turner (MARTÍNEZ-PASARELL ET AL. 1999) i la síndrome de Klinefelter (ARNEDO ET AL. 2006).

La susceptibilitat a la NAHR observada en alguns pares transmissors de la SDG/VCF pot ser deguda a un increment de la inestabilitat focalitzat en la regió 22q11.2 o a la presència d'una inestabilitat meiòtica generalitzada. La fallada dels mecanismes que eviten la NAHR s'ha plantejat com una possible explicació per la inestabilitat generalitzada. En concret, s'han proposat alteracions de l'activitat de les helicasses que

dificulten la formació dels complexos de proteïnes que participen en la formació d'unions de Holliday com RAD51 i DMC1. També s'han proposat alteracions de proteïnes que comproven la fidelitat dels heterodúplex de DNA durant la reparació dels DSB (en anglès *mismatch repair proteins*) com PMS1 i MSH2 (SASAKI ET AL. 2010).

La presència de fenòmens de susceptibilitat generalitzada a la NAHR ha estat proposada en parets d'individus portadors de trastorns genòmics. En concret, Molina (2011) va analitzar la freqüència de delecions i duplicacions de les regions 7q11.23 i 22q11.2 en espermatozoides de parets transmissors de la síndrome de Prader-Willi que presentaven un increment de delecions i duplicacions de la regió 15q11q13. Les dades obtingudes mostraren un increment de la susceptibilitat a la NAHR per les regions crítiques 7q11.23 i/o 22q11.2 en un 60% dels individus.

De forma similar al treball desenvolupat per Molina (2011), en el nostre projecte es van analitzar les freqüències de delecions i duplicacions en espermatozoides d'altres regions inestables del genoma, amb una arquitectura genòmica similar a la de la SDG/VCF. L'anàlisi de múltiples regions va aportar una visió global sobre la susceptibilitat a la NAHR en els parets transmissors de la SDG/VCF. Les regions analitzades van ser la 7q11.23, corresponent a la síndrome de Williams-Beuren, i la regió 15q11q13, corresponent a la síndrome de Prader-Willi/Angelman. En ambdós casos, es tracta de trastorns genòmics causats majoritàriament per una delecio. Els parets transmissors de la SDG/VCF no van mostrar increments de susceptibilitat a la NAHR per la regió 7q11.23 ni per la regió 15q11q13. Per tant, en els dos individus en els quals s'havia determinat prèviament una susceptibilitat a la NAHR per la regió 22q11.2 es va inferir una susceptibilitat focalitzada en aquesta regió crítica.

### Efecte de l'edat en la NAHR

---

En referència a la freqüència de delecions i duplicacions en espermatozoides i la seva relació amb l'edat dels individus, no hi ha consens en la literatura especialitzada sobre si la inestabilitat cromosòmica s'incrementa amb l'edat. Alguns estudis conclouen que l'edat no es relaciona amb el nombre d'anomalies cromosòmiques (KURAHASHI I EMANUEL 2001; PLASTIRA ET AL. 2007). Per contra, altres autors han descrit una correlació entre els increments d'anomalies cromosòmiques i l'edat (SLOTTER ET AL. 2007;

BOSCH ET AL. 2003; PRESTES SARTORELLI ET AL. 2001). A més a més, s'ha suggerit que els increments de reorganitzacions estructurals observats en individus d'edats avançades no mostren una distribució homogènia en tots els cromosomes sinó que els més grans tendeixen a presentar freqüències de delecions i duplicacions més elevades (TEMPLADO ET AL. 2011).

En relació a l'efecte de l'edat en les freqüències de delecions i duplicacions de la regió 22q11.2, no es va observar una correlació entre les freqüències de les reorganitzacions i l'edat dels individus analitzats. En l'anàlisi de reorganitzacions 22q11.2 en espermatozoides d'homes control, Molina et al. (2011) tampoc havien observat aquesta correlació. En conjunt, aquests resultats apunten a què la susceptibilitat a la NAHR de la regió 22q11.2, durant la meiosi masculina, és independent de l'edat. Aquest resultat concorden amb estudis poblacionals que indiquen que l'edat paterna no constitueix un factor de predisposició per aquesta síndrome (DELIO ET AL. 2013).

### Cromàtides involucrades en la NAHR

La sinapsi no al·lèlica que condueix a la NAHR pot produir-se a partir de diferents formes d'aparellament que es classifiquen de forma genèrica en inter-cromàtide (cromàtides homòlogues, cromàtides no homòlogues o cromàtides germanes) i intra-cromàtide (**Figura 3**). Les freqüències de les reorganitzacions resultants de la NAHR permeten deduir el tipus d'aparellament no al·lèlic predominant en cada cas. A nivell poblacional, es va observar una freqüència equivalent de delecions i duplicacions de les regions 7q11.23 i 15q11q13 en els pares transmissors de la SDG/VCF analitzats. En canvi, les freqüències de delecions i duplicacions de la regió 22q11.2 van mostrar diferències significatives. Tot i així, a nivell individual, tan sols es va observar una distorsió de la proporció 1:1 entre delecions i duplicacions 22q11.2 dels dos individus que també presentaven un increment significatiu de delecions 22q11.2.

Aquestes dades suggereixen una contribució majoritària de la NAHR inter-cromàtide per les tres regions analitzades. No obstant, apunten a un augment dels fenòmens de NAHR intra-cromàtide en individus amb increments significatius de delecions 22q11.2. Altres autors han descrit també una participació de la NAHR intra-cromàtide en l'origen d'algunes reorganitzacions causants de la SDG/VCF. Particularment, s'ha estimat que

un 14% de les reorganitzacions que originen la SDG/VCF tenen origen en un esdeveniment de NAHR intra-cromàtide (THOMAS ET AL. 2006).

De forma similar, en pares transmissors de la síndrome de Prader-Willi es determinà que l'increment de delecions 15q11q13 observat en pares amb descendència afectada s'atribuïa a la participació de la NAHR intra-cromàtide (MOLINA ET AL. 2010). Aquests resultats, doncs, reforcen la idea de què l'increment de reorganitzacions en individus que presenten una susceptibilitat incrementada a la NAHR s'associa a l'augment de l'activitat de la NAHR intra-cromàtide.

## 5.2 FACTORS DE PREDISPOSICIÓ A LA SDG/VCF

L'increment d'esdeveniments de NAHR a la regió 22q11.2 pot ser degut a diferents característiques genètiques. En concret, en aquest projecte s'ha estudiat l'arquitectura genòmica dels LCR involucrats en la reorganització, la presència d'inversions de la regió crítica en heterozigosi i el genotip de *PRDM9*.

### ARQUITECTURA GENÒMICA DELS LCR

#### Consideracions metodològiques de l'estudi dels LCR

La naturalesa repetitiva dels LCR constitueix una important limitació per la seva anàlisi, independentment de la tècnica d'estudi aplicada. En conseqüència, el coneixement de la variabilitat estructural en aquestes regions és complicat d'assolir.

En l'aplicació de mètodes de seqüenciació del DNA, les regions duplicades originen assemblatges i assignacions errònies que donen lloc a una cobertura incompleta de les seqüències dels LCR. Aquesta situació explica la imprecisió de les seqüències d'aquestes regions en les bases de dades del genoma (BAILEY ET AL. 2001). L'aplicació de les tècniques de seqüenciació massiva en l'anàlisi de CNV, juntament amb la millora de les eines computacionals, han aconseguit millorar la informació sobre l'arquitectura dels LCR (ALKAN ET AL. 2009; GUO ET AL. 2011; ITSARA ET AL. 2012; NUTTLE ET AL. 2014). Tot i així, aquestes tècniques segueixen subestimant el número total de seqüències repetitives presents en els LCR (ALKAN ET AL. 2010; RETTERER ET AL. 2015).



Les metodologies d'anàlisi de regions específiques, basades en les tècniques de la FISH i de la PCR, s'han aplicat amb èxit en l'estudi de la variabilitat estructural dels LCR (BABCOCK ET AL. 2007; CARVALHO I LUPSKI 2008; CUSCÓ ET AL. 2008; ANTONACCI ET AL. 2010; BOETTGER ET AL. 2012; MOLINA ET AL. 2012B). Concretament, en aquest projecte de tesi es va abordar l'anàlisi de l'arquitectura genòmica a través de la *fiber*-FISH i la ddPCR.

La *fiber*-FISH proporciona una anàlisi directe de l'arquitectura genòmica amb una resolució de 10 – 20 kb (HENG ET AL. 2005). Gràcies a la visualització directa dels patrons de repeticions, la *fiber*-FISH permet diferenciar haplotips i analitzar els LCR de forma individual. Tot i així, la resolució de la tècnica resulta insuficient per discriminar les còpies de seqüències en tàndem. Així doncs, en l'anàlisi de seqüències altament repetitives, com són les regions de LCR, és necessari considerar la possibilitat de què cada senyal observat en una fibra de cromatina pugui no correspondre únicament a una sola còpia de la sonda hibridada. En concordança amb aquest plantejament, al llarg d'aquest projecte vam considerar els senyals dels fòsmids analitzats per *fiber*-FISH com a blocs de seqüències en tàndem. Un altre inconvenient que planteja la *fiber*-FISH és l'elevat temps que exigeix l'aplicació dels protocols juntament amb l'anàlisi microscòpica. La durada de l'anàlisi dels LCR22-2 i LCR22-4 s'estima en dos mesos per individu. Per això, tot i que els resultats obtinguts per aquesta tècnica proporcionen una descripció exhaustiva de l'arquitectura genòmica dels LCR, els seus requeriments temporals dificulten la seva aplicació generalitzada.

D'altra banda, tot i que els mètodes basats en la tecnologia de la PCR permeten distingir les seqüències repetides en tàndem, impedeixen diferenciar variacions de còpies entre els haplotips que presenta cada individu i analitzar variacions de les seqüències d'un únic LCR. Per tant, cal tenir en compte que els estudis en els quals s'apliquen tècniques basades en la PCR proporcionen resultats genèrics de la seqüència diana en tots els LCR paràlegs. Tot i així, els avantatges més destacables respecte la *fiber*-FISH són la baixa complexitat dels protocols, l'obtenció de resultats en període de temps breu i la possibilitat d'incloure un elevat número de mostres en un mateix assaig.

Malgrat les limitacions metodològiques exposades, la *fiber*-FISH constitueix la tècnica més precisa per determinar l'arquitectura genòmica de regions complexes com els LCR

(ATEN ET AL. 2008). Juntament amb els resultats exposats en aquest projecte de tesi, podem trobar altres exemples en els quals s'ha aplicat aquesta metodologia per l'anàlisi de l'arquitectura genòmica de LCR i per l'anàlisi d'altres tipus de seqüències (**Taula 14**).

■ **Taula 14.** Anàlisi de la variabilitat estructural mitjançant *fiber*-FISH.

Variació estructural analitzada	Referència
CNV gen <i>AMY1</i>	Perry et al. 2007
Arquitectura genòmica dels LCR17	Barbouti et al. 2004
CNV gen <i>PLP1</i>	Shimojima et al. 2010b
CNV seqüències regió 22q11.2 fora dels LCR	Shimojima et al. 2010a
Arquitectura genòmica dels LCR22	Molina et al. 2012b

### Variabilitat de blocs de seqüències en tàndem en controls

La variabilitat interindividual en la mida dels LCR ha estat descrita en les bases de dades de variants genòmiques (REDON ET AL. 2006). Tanmateix, les variacions en la seva arquitectura es troben escassament caracteritzades degut a la complexitat de la seva anàlisi. Els resultats obtinguts mitjançant la tècnica de la *fiber*-FISH avalen l'existència de variacions dels blocs de duplicacions segmentals en els LCR22 i han contribuït a la caracterització de la seva arquitectura genòmica.

En la població control, l'anàlisi dels blocs de seqüències en tàndem va mostrar una variabilitat interindividual major en el LCR22-2 que en el LCR22-4. Els rang de còpies de blocs de seqüències en tàndem dels fòsmid analitzats en ambdós LCR (L9, K3 i B22) són majors en el LCR22-2, especialment en referència als blocs que corresponen a les seqüències dels fòsmids L9 i K3. Al contrari, en el LCR22-4 el nombre de blocs és menys variable per tots els fòsmids analitzats i les diferències interindividuals també són menors. En conjunt, aquestes observacions suggereixen que el LCR22-2 ocupa una regió més extensa que el LCR22-4. Coincidentment, l'anàlisi de l'arquitectura genòmica dels LCR22 realitzat per Shaikh et al. (2000) estimà que l'LCR22-2 presentava una mida aproximada de ~350 kb mentre que l'LCR22-4 tenia una mida menor amb ~250 kb (SHAIKH ET AL. 2000).

Els factors que condicionen la freqüència de la NAHR són la longitud, el grau d'homologia entre els blocs de duplicacions segmentals i la densitat de motius de recombinació (DITTWALD ET AL. 2013B; PENG ET AL. 2014). Respecte les dues primeres variables, els LCR22-2 i LCR22-4 tenen una major similitud pel que fa a estructura i a identitat de seqüència respecte la resta de LCR22. Aquests LCR comparteixen gairebé la mateixa estructura de blocs de duplicacions segmentals i mostren una alta conservació de la seqüència amb una identitat del 97-98% en les 250 kb duplicades en ambdues regions (SHAIKH ET AL. 2000; BABCOCK ET AL. 2003). D'acord amb el model de sinapsis ectòpica, les variacions en la longitud dels LCR podrien afavorir contactes presinàptics entre regions pròximes que presenten una alta homologia, que actuarien com a precursors de la NAHR (LIU ET AL. 2011). Així doncs, la variabilitat dels blocs de duplicacions segmentals entre l'LCR22-2 i LCR22-4 podria explicar la major susceptibilitat d'aquest LCR als fenòmens de NAHR.

En referència a la densitat de motius de recombinació, es va descriure el número de motius de PRDM9 en les seqüències dels fòsmids (**Taula 15**). Les seqüències amb una variabilitat interindividual major, corresponents als fòsmid L9 i K3, van mostrar 2 i 3 motius de recombinació, respectivament. En canvi, les seqüències dels fòsmids L21 i M12, que presentaven una baixa variabilitat interindividual, van mostrar les densitats més elevades de motius de recombinació amb 8 i 14 motius, respectivament. Per tant, no es va observar una relació entre l'increment de la variabilitat interindividual dels fòsmids analitzats i la densitat de motius de recombinació. D'altra banda, es va observar com els fòsmids L21 i M12, amb una major densitat de motius de recombinació, cobreixen els gens *DGCR6* (OMIM 601279) i *PRODH* (OMIM 606810), i *RIMBP3C* (OMIM 612701), respectivament, implicats en els punts de trencament recorrents descrits en la deleció causant de la SDG/VCF (**Taula 15**) (BITTEL ET AL. 2009). Així doncs, l'alta densitat de motius de recombinació podria explicar la localització dels punts de trencament associats a la SDG/VCF.

#### Variabilitat de blocs de seqüències en tàndem en parets SDG/VCF

---

Un cop establerta la variabilitat dels LCR22-2 i LCR22-4 en individus control, es va analitzar l'arquitectura genòmica d'aquests LCR en parets transmissors de la SDG/VCF.

■ **Taula 15.** Anàlisi de les seqüències dels fòsmids (UCSC *Assembly Feb 2009*).

Codi del fòsmid	LCR22	Gens UCSC	Motius de recombinació*
L9	LCR22-2, LCR22-4	<i>USP18</i> , <i>AK129567</i> <sup>a, b</sup>	3
K3	LCR22-2, LCR22-4	<i>AK302545</i> , <i>GGT3P</i>	2
B22	LCR22-2, LCR22-4	<i>GGT3P</i>	5
L21	LCR22-2	<i>DGCR6</i> <sup>a</sup> , <i>PRODH</i> <sup>a</sup>	8
M12	LCR22-4	<i>RIMBP3C</i> <sup>b</sup>	14
I18	LCR22-4	<i>PI4KAP2</i>	2

Els gens en gris es troben coberts parcialment pel fòsmid corresponent.

\* Motius de recombinació reconeguts per PRDM9, CCNCCNTNCCNC.

<sup>a</sup> Gens descrits en el punt de trencament proximal (LCR22-2) (BITTEL ET AL. 2009).

<sup>b</sup> Gens descrits en el punt de trencament distal (LCR22-4) (BITTEL ET AL. 2009).

Degut a la limitació temporal de la tècnica de la *fiber*-FISH, es va prioritzar l'anàlisi dels pares amb increments de delecions 22q11.2 i que, per tant, presentaven una major susceptibilitat a la NAHR en aquesta regió. És important remarcar que els resultats previs de l'anàlisi de microsatèl·lits en ambdós individus van permetre determinar que la delecio que afectava a la seva descendència corresponia al segment de 3 Mb flanquejat pels LCR22-2 i LCR22-4 (**Taula 8**).

En ambdós pares SDG/VCF, es van identificar números de còpies de blocs de seqüències en tàndem fora dels intervals de normalitat definits en controls pels fòsmids L9 i K3. L'increment de delecions 22q11.2 en aquests individus constata que les variacions observades en els blocs de duplicacions segmentals podrien modificar l'homologia entre algunes regions dels LCR22-2 i LCR22-4 i promoure la recombinació no al·lèlica durant la meiosi. Tanmateix, l'anàlisi de les seqüències dels fòsmids (**Taula 15**) demostra que majoritàriament els punts de trencament a partir dels qual succeïrien els esdeveniments de NAHR no es localitzen en les seqüències en les quals hem observat variacions. La susceptibilitat a patir un DSB podria localitzar-se en altres seqüències contigües com les dels fòsmids L21, a LCR22-2, i M12, a LCR22-4. Aquestes seqüències contenen una alta densitat de motius de recombinació de *PRDM9* i cobreixen completament alguns dels gens descrits en els punts de trencaments habituals de la delecio 22q11.2 (**Taula 15**).

Així doncs, els resultats obtinguts suggereixen que les variacions de les seqüències dels fòsmids L9 o K3 podrien promoure l'aparellament no al·lèlic dels LCR. En conseqüència, els DSB produïts en seqüències contigües durant la meiosi tindrien una probabilitat més elevada de generar esdeveniments de NAHR i originar gàmetes portadores de reorganitzacions (**Figura 31A**).

### Variabilitat de seqüències paràlogues

Tenint en compte que els CNV observats en els pares SDG/VCF es restringien a les seqüències dels fòsmids L9 i K3 es van intentar validar els resultats a través de l'anàlisi del número de còpies de gens continguts en les regions reconegudes per aquests fòsmids. A part de la validació dels resultats obtinguts per *fiber*-FISH, un segon objectiu dels assajos de ddPCR va ser la recerca d'una tècnica que pogués detectar els mateixos CNV amb una metodologia més senzilla. La ddPCR, en contraposició a la *fiber*-FISH, permetria analitzar un nombre més elevat d'individus i requeriria un període de temps menor.

L'anàlisi del número total de seqüències dels gens *AK129567* i *AK302545* coberts pels fòsmids L9 i K3 i amb seqüències paràlogues localitzades en múltiples LCR22, va mostrar una alta variabilitat de còpies en la població control. L'assemblatge del Desembre de 2013 (UCSC Genome Browser) assenyala l'existència de dues de còpies d'aquests gens. Tanmateix els resultats obtinguts posen de manifest que el mínim de còpies estimables en la població control és molt superior al descrit en les bases de dades amb 12 còpies per *AK129567* i de 10 còpies per *AK302545*. La manca de concordança entre els nostres resultats i les bases de dades del genoma evidencia la infravaloració de les seqüències contingudes en les regions de LCR.

D'altra banda, en la comparació entre les dades obtingudes en controls i en individus transmissors de la SDG/VCF, no es va poder establir cap relació entre els CNV de les seqüències paràlogues dels fòsmids L9 i K3 i el risc de SDG/VCF. A més a més, els resultats obtinguts en individus en els quals es va aplicar tan la ddPCR com la *fiber*-FISH (6 controls i 2 pares transmissors de la SDG/VCF amb increments d'espermatozoides portadors de la deleció 22q11.2) determinaren que no hi havia una correlació entre les dades obtingudes per ambdues tècniques. L'anàlisi de CNV en múltiples LCR22

mitjançant la ddPCR emmascarava les variacions significatives a LCR22-2 i LCR22-4 detectades mitjançant *fiber*-FISH. Aquesta observació reforça la necessitat d'utilitzar tècniques que permetin analitzar l'arquitectura genòmica de LCR22 de forma específica.

En altres trastorns genòmics com la síndrome de Williams-Beuren, l'anàlisi de les variacions de còpies de gens i pseudogens presents en LCR paràlegs sí que ha permès la identificació de CNV significatius entre la població control i els pares transmissors d'aquesta síndrome (CUSCÓ ET AL. 2008). Tanmateix, la regió 7q11.23, relacionada amb la síndrome de Williams-Beuren, presenta tan sols 3 LCR mentre que en la regió 22q11.2 se'n han comptabilitzat un total de 8. La menor complexitat dels LCR facilita l'observació de CNV relacionats amb els LCR que originen la reorganització causant del trastorn genòmic. A part dels trastorns genòmics, també s'han observat CNV de gens i pseudogens amb còpies paràlogues en LCR, com *CCL3L1*, *DEFB* o *FCGR3B*, que s'han assenyalat com a factors de predisposició a malalties del sistema immunitari (MARCINKOWSKA-SWOJAK ET AL. 2014; FRANKE ET AL. 2015; CANTSILIERIS I WHITE 2013). En aquests casos, els CNV no es relacionen amb la susceptibilitat a reorganitzacions entre LCR sinó amb una susceptibilitat directa a malalties com a conseqüència de modificacions en la dosi gènica i, per tant, són independents de la complexitat dels LCR en els quals es localitzen.

A nivell individual, es va observar un pare transmissor de la SDG/VCF amb un CNV de seqüències del fòsmid K3 fora del rang de normalitat establert en controls. Tot i això, considerant que aquest individu no havia presentat una susceptibilitat a la NAHR en les anàlisis d'espermatozoides i que, degut a les característiques tècniques de la ddPCR, no podem atribuir l'increment de còpies a un LCR22 específic, considerem que l'efecte d'aquesta dada en l'avaluació dels risc de transmissió d'aquest individu és incert. Per tal de concretar la variació de còpies del fòsmid K3 en els LCR22-2 i LCR22-4, caldria realitzar un anàlisi dels blocs de seqüències en tàndem mitjançant la *fiber*-FISH.

### Variabilitat de pseudogens

---

Els assajos de ddPCR van detectar dues intensitats de fluorescència diferents en l'amplificació del gen *AK129567*. En assajos amb sondes TaqMan s'ha descrit una relació inversament proporcional entre la fluorescència emesa i la mida dels amplicons

resultants de la reacció (LAURIE ET AL. 2013). Per tant, la presència de dues intensitats diferents en una mateixa reacció es pot explicar per l'amplificació de dos tipus de seqüències del fòsmid L9 amb mides diferents. En l'anàlisi dels resultats es va interpretar que els amplicons amb una intensitat de fluorescència inferior, i representats minoritàriament en el conjunt de les reaccions, corresponien a l'amplificació d'un pseudogen de *AK129567* (Servei tècnic de BioRad; comunicació personal).

Els CNV de pseudogens localitzats en LCR s'han proposat com un factor modificador de la susceptibilitat a la NAHR. Fusco et al. (2009) relacionaren la variació de la seqüència d'un pseudogen amb la promoció de reorganitzacions entre LCR de la regió Xq28. En el seu treball, aquests autors analitzaren la variabilitat estructural dels LCR L1 i L2 de la regió Xq28 en famílies amb descendència afectada per la incontinència pigmentària. Aquesta malaltia s'origina per una reorganització entre L1 i L2 que dona lloc a la deleció del gen *IKBKG* (OMIM 300248). Ambdós LCR comparteixen una elevada homologia de seqüència, tot i que presenten diferències pel què fa al contingut de les seqüències repetitives *MER67B* i a la presència del gen *IKBKG*, a L1, i el seu corresponent pseudogen, a L2. Els seus resultats demostraren com els polimorfismes d'una deleció del pseudogen *IKBKG* i d'una duplicació de *MER67B* incrementaven la susceptibilitat a l'aparellament no al·lèlic dels LCR L1 i L2 i promovien la NAHR durant la meiosi (FUSCO ET AL. 2009). Així doncs, les variacions de seqüències de pseudogens condicionen l'homologia dels LCR i poden constituir un factor de risc per la generació de delecions.

De manera similar, en el context del nostre estudi, l'absència del pseudogen del fòsmid L9 podria alterar el grau d'homologia entre els LCR22-2 i LCR22-4 incentivant el seu aparellament no al·lèlic i, per tant, incrementant els esdeveniments de NAHR (**Figura 31B**). No obstant, en el nostre cas la complexitat de la regió 22q11.2 amb un total de 8 LCR dificulta la comprovació d'aquesta hipòtesi. Cal tenir present que no podem assegurar que els pseudogens detectats es trobin en els LCR involucrats en la reorganització que causa la SDG/VCF, atès que la tècnica de la ddPCR impedeix l'assignació dels CNV observats a cap LCR22.

Els resultats obtinguts suggereixen una tendència a la disminució de còpies del pseudogen de *AK129567p* en la població d'individus transmissors de la SDG/VCF

respecte als controls. Tanmateix, la mida de la cohort d'individus transmissors analitzada és massa petita per obtenir una significança estadística que permeti establir una relació entre CNV de *AK129567p* i la SDG/VCF. Per tal de precisar la rellevància dels CNV d'aquest pseudogen, així com d'altres seqüències que conformen els LCR, caldria recopilar més dades en una cohort major de parens transmissors de la SDG/VCF.

## INVERSIONS EN HETEROZIGOSI DE LA REGIÓ CRÍTICA

### Consideracions metodològiques

La FISH interfàsica ha estat prèviament aplicada per altres autors amb l'objectiu d'identificar haplotips d'inversions de diferents regions crítiques (GIMELLI ET AL. 2003; OSBORNE ET AL. 2001; BAYÉS ET AL. 2003; GEBHARDT ET AL. 2003). En aquests estudis, l'assignació dels haplotips normals i invertits es va realitzar tenint en compte l'ordre dels senyals. En canvi, l'anàlisi d'inversions 22q11.2 desenvolupat en el present projecte, tal i com s'havia realitzat en l'estudi d'inversions 15q11q13 de Molina et al. (2012), es va basar en la distribució dels senyals.

Mitjançant la nostra metodologia es van observar tres tipus de distribucions dels senyals de FISH per cada haplotip: 2 senyals associats i un senyal independent (2+1), 3 senyals associats i 3 senyals separats. Les distribucions dels 3 senyals associats o separats poden ser conseqüència de la formació de llaços de cromatina propis de l'empaquetament del DNA en els nuclis interfàsics (VORSANOVA ET AL. 2010). En conseqüència, únicament es van considerar informatius els haplotips amb una distribució 2+1.

Per tal d'identificar la disposició dels senyals en forma 2+1, és necessari que la distància entre les sondes agrupades i la sonda independent sigui superior al límit de resolució establert per la FISH interfàsica (>50 kb). En el nostre cas, els senyals agrupats es trobaven a una distància de 600 kb, mentre que el senyal independent es trobava a 3,6 – 4,2 Mb. Cal destacar, que no totes les regions flanquejades per LCR permeten el disseny de sondes per la detecció d'inversions que donin lloc a una distribució 2+1 dels senyals.

Es van observar tres tipus de associacions de senyals amb distribució 2+1 corresponents a l'haplotip normal, l'haplotip invertit i una associació no esperada atribuïble a l'efecte de la disposició dels llaços de cromatina. Tan el recompte



d'haplotips invertits com el recompte d'associacions no esperades van mostrar una freqüència inferior al 3%. Aquesta observació va suggerir que, al igual que les associacions no esperades, els haplotips designats com a invertits eren el resultat de la disposició de la cromatina i que no eren veritables inversions. Tenint en compte aquestes dades, la presència d'un haplotip invertit només pot ser considerada quan la freqüència observada és superior al llindar d'error establert en un 6%.

### Inversions de la regió 22q11.2

---

La inversió de 3 Mb de la regió crítica de la SDG/VCF no va ser observada en cap haplotip dels individus transmissors de la SDG/VCF ni en els individus control analitzats. En conjunt, els individus estudiats en aquest projecte i els analitzats prèviament per Gebhardt et al. (2003) i Saitta et al. (2004) demostren l'absència d'inversions 22q11.2 en un total de 43 controls i 48 pares transmissors de la síndrome de SDG/VCF. Aquest resultat descarta la presència d'aquest polimorfisme i la seva rellevància com a factor de predisposició de la SDG/VCF.

En altres regions crítiques corresponents a altres desordres genòmics sí que s'ha demostrat l'existència d'inversions polimòrfiques. Aquestes inversions en heterozigosi constitueixen un factor de susceptibilitat reorganitzacions secundàries causants de desordres genòmics. Per exemple, la inversió 7q11.23 afecta a un 5% de la població general i s'ha detectat en un 25-33% dels pares transmissors de la síndrome de Williams-Beuren (BAYÉS ET AL. 2003; SCHERER ET AL. 2005; OSBORNE ET AL. 2001; HOBART ET AL. 2010). El fet que no totes les regions flanquejades per LCR mostrin inversions polimòrfiques apunta a una susceptibilitat diferent a la formació d'aquesta reorganització en funció de l'arquitectura genòmica dels LCR involucrats (**Taula 16**).

---

### GENOTIP DE *PRDM9*

#### Consideracions metodològiques del genotipatge de *PRDM9*

---

Tal i com hem descrit a l'apartat 3.3, l'anàlisi del genotip de *PRDM9* requereix la realització d'una PCR i una posterior reacció de seqüenciació Sanger sobre el producte amplificat. En al·lels amb 13 dits de zinc, com l'al·lel majoritari A de *PRDM9*, el

■ **Taula 16.** Inversions polimòrfiques causant de reorganitzacions secundàries relacionades amb trastorns genòmics (Adaptat de (FEUK 2010)).

Inversió	Trastorn genòmic	Referència
inv(3)(q29)	Síndrome de la microdeleció 3q29	Antonacci et al. 2009
inv(5)(q35)*	Síndrome de Sotos	Visser et al. 2005
inv(7)(q11.23)*	Síndrome de Williams-Beuren	Hobart et al. 2010
inv(8)(p23)	Síndrome de la microdeleció 8p23	Sugawara et al. 2003
inv(15)(q11q13)*	Síndrome de Prader-Willi/Angelman	Gimelli et al. 2003
inv(15)(q13.3)*	Síndrome de la deleció 15q13.3	Sharp et al. 2008
inv(15)(q24)	Síndrome de la microdeleció 15q24	Kidd et al. 2008
inv(17)(p11.2p13)	Síndrome de Smith-Magenis	Yang et al. 1997
inv(17)(q12)	Síndrome de quists renals i diabetis	Antonacci et al. 2009
inv(17)(q21.31)*	Síndrome de la deleció 17q21.31	Rao et al. 2010

\* Freqüències de la inversió més elevades en pares transmissors que en la població general.

Producte de la PCR presenta una longitud de 1900 pb dels quals se'n seqüencien 1240 pb. Les metodologies de seqüenciació Sanger ofereixen lectures amb una longitud aproximada de 800 pb. Per aquest motiu, per seqüenciar els 1240 pb que conformen el domini de dits de zinc es realitzen dues lectures diferents del producte de PCR, amb encebadors *forward* i *reverse* respectivament. Les dues lectures realitzades s'alineen per tal d'identificar tots els dits de zinc.

La seqüenciació dels al·lels de *PRDM9* amb més de 13 dits de zinc, és més complexa ja que cada dit de zinc incrementa 84 pb en la seqüència analitzada i fa que el solapament de la lectura en *forward* i en *reverse* s'hagi de realitzar en segments en els quals la fiabilitat de la seqüència és menor. En conseqüència, l'augment de dits de zinc pot impedir l'obtenció d'un alineament fidedigne entre les dues lectures realitzades. Aquesta situació genera incertesa en l'assignació dels dits de zinc centrals en el domini i en la caracterització de canvis de nucleòtids que poden comportar la descripció de nous tipus de dits de zinc i de nous al·lels de *PRDM9*. En el nostre estudi, aquestes limitacions han impedit l'assignació dels dits de zinc centrals en el domini de l'al·lel L50, format per més de 13 dits de zinc. En qualsevol cas, la combinació de dits de zinc identificada no corresponia a cap dels al·lels descrits anteriorment. Per aquest motiu, vam poder assegurar que es tractava d'un al·lel identificat per primera vegada.

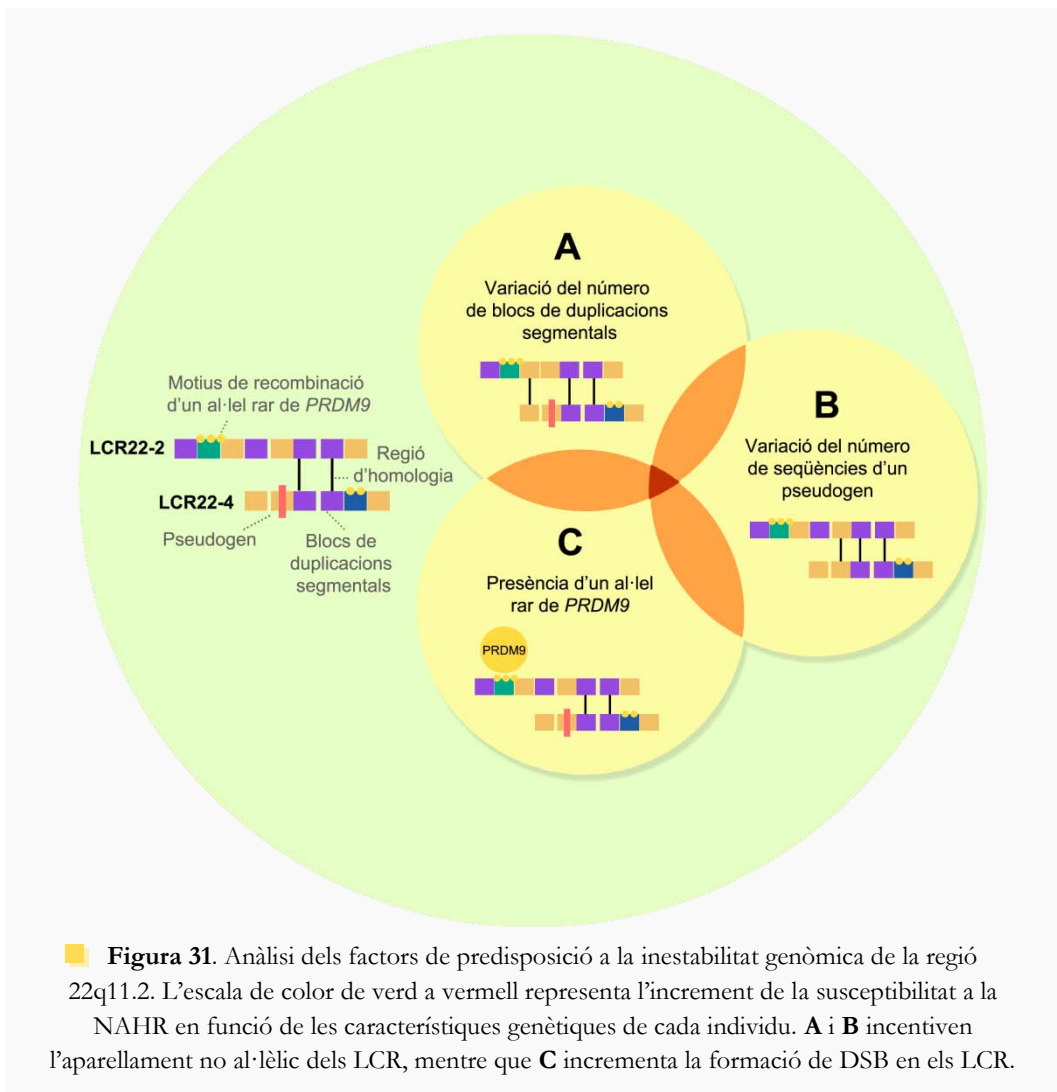
### Al·lels de *PRDM9*

S'ha demostrat que hi ha una correlació positiva entre la densitat de motius de recombinació reconeguts per *PRDM9* i la freqüència de NAHR (DITTWALD ET AL. 2013B). A més a més, s'han descrit enriquiments de seqüències motiu reconegudes per *PRDM9* en els punts de trencament de les reorganitzacions estructurals causades per NAHR (Myers et al. 2008; Pratto et al. 2014). Aquestes observacions han contribuït a la proposició del genotip de *PRDM9* com a factor de predisposició de desordres genòmics. En aquest sentit, l'al·lel A de *PRDM9* s'ha proposat com a factor de predisposició de la síndrome de Charcot-Marie-Tooth tipus 1A i la neuropatia hereditària amb paràlisi sensible a la pressió, entre altres trastorns genòmics (BERG ET AL. 2010; PRATTO ET AL. 2014). També s'ha suggerit que els portadors d'al·lels rars de *PRDM9* poden tenir una major predisposició a la generació de delecions 7q11.23 associada a la síndrome de Williams-Beuren (BOREL ET AL. 2012). D'altra banda, s'han proposat alguns al·lels de *PRDM9*, com l'al·lel C, que poden ser protectors ja que les seqüències motiu que reconeixen no es troben en els punts de trencament coincidents amb trastorns genòmics (PRATTO ET AL. 2014).

En l'anàlisi del genotip de *PRDM9* com a factor de predisposició de la SDG/VCF realitzat en aquest estudi, no es van observar diferències significatives entre les freqüències al·lèliques de 19 individus control i 8 individus transmissors de la SDG/VCF. Coincidentment, altres autors tampoc han pogut identificar cap associació entre les freqüències al·lèliques i la SDG/VCF (ALEMANY-SCHMIDT ET AL. 2014; BOREL ET AL. 2012). Si analitzem les dades publicades per Alemany-Schmidt et al. (2014) conjuntament amb els nostres resultats, observem com en la població espanyola no hi ha al·lels específics de *PRDM9* que actuïn com a factor de predisposició per la SDG/VCF.

En l'anàlisi dels al·lels de *PRDM9* en parets transmissors de la SDG/VCF, es va identificar un al·lel nou de *PRDM9* denominat L50. Els al·lels rars de *PRDM9*, com L50, poden generar patrons de recombinació diferents a través de la creació de DSB inusuals als extrems de la regió crítica de la SDG/VCF que incentivin els fenòmens de la NAHR (**Figura 31C**).

L'al·lel L50 es va descriure en un individu que havia mostrat CNV dels blocs de seqüències en tàndem dels LCR22-2 i LCR22-4 i que a la vegada presentava un increment de delecions 22q11.2 en espermatozoides. La confluència d'un al·lel rar de *PRDM9* i de variacions en l'arquitectura genòmica dels LCR22 en un mateix individu podria reflectir la contribució de múltiples factors de predisposició en la definició de la seva susceptibilitat a la NAHR de la regió 22q11.2 durant la meiosi. Els CNV dels blocs de seqüències en tàndem modificarien la probabilitat d'un aparellament no al·lèlic mentre que l'al·lel rar de *PRDM9* podria induir l'inici de la NAHR a través del reconeixement de motius de recombinació inusuals i la conseqüent formació de punts de trencaments (**Figura 31**).







# CONCLUSIONS



## PRIMERA

La freqüència de delecions de la regió 22q11.2 en espermatozoides es troba incrementada en alguns pares transmissors de la SDG/VCF a causa d'una major susceptibilitat a la NAHR intra-cromàtide. Aquesta susceptibilitat a la NAHR es restringeix a la regió 22q11.2.

## SEGONA

L'arquitectura genòmica dels LCR22-2 i LCR22-4 condiciona la freqüència d'esdeveniments de NAHR:

- Els LCR22-2 i LCR22-4 presenten variacions estructurals diferencials en pares transmissors de la SDG/VCF amb una major susceptibilitat a la NAHR en la regió 22q11.2. Les variacions en l'arquitectura genòmica dels LCR22 poden modificar el seu grau d'homologia i afavorir els aparellaments no al·lèlics.
- L'anàlisi de seqüències paràlogues en els LCR22 suggereix una possible implicació dels CNV del pseudogen de *AK129567* en l'alteració de la homologia.

## TERCERA

L'haplotip de la inversió 22q11.2 no s'observa constitutivament en individus control ni en individus transmissors de la SDG/VCF. La inexistència d'aquest polimorfisme exclou la implicació de les inversions de la regió crítica com a factor de predisposició de la SDG/VCF.



#### QUARTA

Les freqüències al·lèliques de *PRDM9* no presenten diferències significatives entre individus control i individus transmissors de la SDG/VCF. Tanmateix, l'al·lel L50 descrit en un pare transmissor de la SDG/VCF amb un increment de delecions 22q11.2 en espermatozoides, indica la seva possible contribució en la susceptibilitat a la NAHR d'aquest individu.

#### CINQUENA

El risc a la NAHR en la regió 22q11.2 és complex i atribuïble a la confluència de diferents característiques genètiques.

# BIBLIOGRAFIA

## A

- Aitman TJ, Dong R, Vyse TJ, Norsworthy PJ, Johnson MD, Smith J, Mangion J, Robertson-Lowe C, Marshall AJ, Petretto E, et al. 2006. Copy number polymorphism in *Fcgr3* predisposes to glomerulonephritis in rats and humans. *Nature* **439**: 851-855.
- Alberti A, Romano C, Falco M, Cali F, Schinocca P, Galesi O, Spalletta A, Di Benedetto D, Fichera M. 2007. 1.5 Mb de novo 22q11.21 microduplication in a patient with cognitive deficits and dysmorphic facial features. *Clin Genet* **71**: 177-182.
- Aleman-Schmidt A, Navarro-Palou M, Voltes-Cobo A, Rosell J, Heine-Suñer D, Picornell A, Oliver-Bonet M. 2014. Geographical genetic variability: a factor to consider when assessing clinical implications of PRDM9. *Mol Genet genomic Med* **2**: 201-203.
- Alkan C, Kidd JM, Marques-Bonet T, Aksay G, Antonacci F, Hormozdiari F, Kitzman JO, Baker C, Malig M, Mutlu O, et al. 2009. Personalized copy number and segmental duplication maps using next-generation sequencing. *Nat Genet* **41**: 1061-1067.
- Alkan C, Sajjadian S, Eichler EE. 2010. Limitations of next-generation genome sequence assembly. *Nat Methods* **8**: 61-65.
- Angelo CSD, Jehue FS, Koiffmann P. 2007. An Inherited Atypical 1 Mb 22q11.2 Deletion Within the DGS / VCFS 3 Mb Region in a Child With Obesity and Aggressive Behavior. **1932**: 1928-1932.
- Anton E, Blanco J, Egozcue J, Vidal F. 2004. Sperm FISH studies in seven male carriers of Robertsonian translocation t(13;14)(q10;q10). *Hum Reprod* **19**: 1345-1351.
- Antonacci F, Kidd JM, Marques-Bonet T, Teague B, Ventura M, Girirajan S, Alkan C, Campbell CD, Vives L, Malig M, et al. 2010. A large and complex structural polymorphism at 16p12.1 underlies microdeletion disease risk. *Nat Genet* **42**: 745-750.
- Antonacci F, Kidd JM, Marques-Bonet T, Ventura M, Siswara P, Jiang Z, Eichler EE. 2009. Characterization of six human disease-associated inversion polymorphisms. *Hum Mol Genet* **18**: 2555-2566.
- Armengol L, Pujana MA, Cheung J, Scherer SW, Estivill X. 2003. Enrichment of segmental duplications in regions of breaks of synteny between the human and mouse genomes suggest their involvement in evolutionary rearrangements. *Hum Mol Genet* **12**: 2201-2208.

- Armengol L, Villatoro S, González JR, Pantano L, García-Aragónés M, Rabionet R, Cáceres M, Estivill X. 2009. Identification of copy number variants defining genomic differences among major human groups. *PLoS One* **4**: e7230.
- Arnedo N, Templado C, Sánchez-Blancque Y, Rajmil O, Nogués C. 2006. Sperm aneuploidy in fathers of Klinefelter's syndrome offspring assessed by multicolour fluorescent in situ hybridization using probes for chromosomes 6, 13, 18, 21, 22, X and Y. *Hum Reprod* **21**: 524-528.
- Aten E, White SJ, Kalf ME, Vossen RH a M, Thygesen HH, Ruivenkamp C a., Kriek M, Breuning MHB, Den Dunnen JT. 2008. Methods to detect CNVs in the human genome. *Cytogenet Genome Res* **123**: 313-321.
- Azofeifa J, Fauth C, Kraus J, Maierhofer C, Langer S, Bolzer A, Reichman J, Schuffenhauer S, Speicher MR. 2000. An optimized probe set for the detection of small interchromosomal aberrations by use of 24-color FISH. *Am J Hum Genet* **66**: 1684-1688.


  
B

- Babcock M, Pavlicek A, Spiteri E, Kashork CD, Ioshikhes I, Shaffer LG, Jurka J, Morrow BE. 2003. Shuffling of Genes Within Low-Copy Repeats on 22q11 (LCR22) by Alu-Mediated Recombination Events During Evolution. *Genome Res* **13**: 2519-2532.
- Babcock M, Yatsenko S, Hopkins J, Brenton M, Cao Q, de Jong P, Stankiewicz P, Lupski JR, Sikela JM, Morrow BE. 2007. Hominoid lineage specific amplification of low-copy repeats on 22q11.2 (LCR22s) associated with velo-cardio-facial/digeorge syndrome. *Hum Mol Genet* **16**: 2560-2571.
- Bailey JA, Eichler EE. 2006. Primate segmental duplications: crucibles of evolution, diversity and disease. *Nat Rev Genet* **7**: 552-564.
- Bailey JA., Yavor AM, Massa HF, Trask BJ, Eichler EE. 2001. Segmental duplications: Organization and impact within the current human genome project assembly. *Genome Res* **11**: 1005-1017.
- Bailey JA, Gu Z, Clark RA, Reinert K, Samonte R V, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE. 2002. Recent Segmental Duplications in the Human Genome. *Science* **297**: 1003-1007.
- Barbouti A, Stankiewicz P, Nusbaum C, Cuomo C, Cook A, Höglund M, Johansson B, Hagemeyer A, Park S-S, Mitelman F, et al. 2004. The Breakpoint Region of the Most Common Isochromosome, i(17q), in Human Neoplasia Is Characterized by a Complex Genomic Architecture with Large, Palindromic, Low-Copy Repeats. *Am J Hum Genet* **74**: 1-10.
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**: 836-840.

- Baumer A. 1998. High level of unequal meiotic crossovers at the origin of the 22q11.2 and 7q11.23 deletions. *Hum Mol Genet* **7**: 887-894.
- Bayés M, Magano LF, Rivera N, Flores R, A. Pérez Jurado L. 2003. Mutational Mechanisms of Williams-Beuren Syndrome Deletions. *Am J Hum Genet* **73**: 131-151.
- Ben-shachar S, Ou Z, Shaw CA, Belmont JW, Patel MS, Hummel M, Amato S, Tartaglia N, Berg J, Sutton VR, et al. 2008. 22q11.2 Distal Deletion: A Recurrent Genomic Disorder Distinct from DiGeorge Syndrome and Velocardiofacial Syndrome. *Am J Hum Genet* **82**: 214-221.
- Berg IL, Neumann R, Lam K-WG, Sarbajna S, Odenthal-Hesse L, May CA, Jeffreys AJ. 2010. PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat Genet* **42**: 859-863.
- Bittel DC, Yu S, Newkirk H, Kibiryeva N, Holt A, Butler MG, Cooley LD. 2009. Refining the 22q11.2 deletion breakpoints in DiGeorge syndrome by aCGH. *Cytogenet Genome Res* **124**: 113-120.
- Blanco J, Gabau E, Gómez D, Baena N, Guitart M, Egozcue J, Vidal F. 1998. Chromosome 21 disomy in the spermatozoa of the fathers of children with trisomy 21, in a population with a high prevalence of Down syndrome: increased incidence in cases of paternal origin. *Am J Hum Genet* **63**: 1067-1072.
- Boettger LM, Handsaker RE, Zody MC, McCarroll SA. 2012. Structural haplotypes and recent evolution of the human 17q21.31 region. *Nat Genet* **44**: 881-885.
- Borel C, Cheung F, Stewart H, Koolen D a, Phillips C, Thomas NS, Jacobs PA, Eliez S, Sharp AJ. 2012. Evaluation of PRDM9 variation as a risk factor for recurrent genomic disorders and chromosomal non-disjunction. *Hum Genet* **131**: 1519-1524.
- Bosch M, Rajmil O, Egozcue J, Templado C. 2003. Linear increase of structural and numerical chromosome 9 abnormalities in human sperm regarding age. *Eur J Hum Genet* **11**: 754-759.
- Bosch M, Rajmil O, Martínez-Pasarell O, Egozcue J, Templado C. 2001. Linear increase of diploidy in human sperm with age: a four-colour FISH study. *Eur J Hum Genet* **9**: 533-538.
- Botto LD, May K, Fernhoff PM, Correa A, Coleman K, Rasmussen SA, Merritt RK, O'Leary LA, Wong L-Y, Elixson EM, et al. 2003. A Population-Based Study of the 22q11.2 Deletion: Phenotype, Incidence, and Contribution to Major Birth Defects in the Population. *Pediatrics* **112**: 101-107.



- Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, Nagamani SCS, Erez A, Bartnik M, Wiśniowiecka-Kowalnik B, et al. 2014. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am J Hum Genet* **95**: 173-182.

- Cantsilieris S, White SJ. 2013. Correlating Multiallelic Copy Number Polymorphisms with Disease Susceptibility. *Hum Mutat* **34**: 1-13.
- Carlson C, Sirotkin H, Pandita R, Goldberg R, McKie J, Wadey R, Patanjali SR, Weissman SM, Anyane-Yeboah K, Warburton D, et al. 1997. Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am J Hum Genet* **61**: 620-629.
- Carvalho CMB, Lupski JR. 2008. Copy number variation at the breakpoint region of isochromosome 17q. *Genome Res* **18**: 1724-1732.
- Chowdhury R, Bois PRJ, Feingold E, Sherman SL, Cheung VG. 2009. Genetic Analysis of Variation in Human Meiotic Recombination ed. G.P. Copenhagen. *PLoS Genet* **5**: e1000648.
- Cooper GM, Nickerson DA, Eichler EE. 2007. Mutational and selective effects on copy-number variants in the human genome. *Nat Genet* **39**: S22-9.
- Coppinger J, McDonald-McGinn D, Zackai E, Shane K, Atkin JF, Asamoah A, Leland R, Weaver DD, Lansky-Shafer S, Schmidt K, et al. 2009. Identification of familial and de novo microduplications of 22q11.21-q11.23 distal to the 22q11.21 microdeletion syndrome region. *Hum Mol Genet* **18**: 1377-1383.
- Coughlin CR, Scharer GH, Shaikh TH. 2012. Clinical impact of copy number variation analysis using high-resolution microarray technologies: advantages, limitations and concerns. *Genome Med* **4**: 80.
- Cuscó I, Corominas R, Bayés M, Flores R, Rivera-Brugués N, Campuzano V, Pérez-Jurado LA. 2008. Copy number variation at the 7q11.23 segmental duplications is a susceptibility factor for the Williams-Beuren syndrome deletion. *Genome Res* **18**: 683-694.


 D

- Darai-ramqvist E, Sandlund A, Müller S, Klein G, Imreh S, Kost-alimova M. 2008. Segmental duplications and evolutionary plasticity at tumor chromosome break-prone regions. *Genome Res* **18**: 370-379.
- Davis A, Chen D. 2013. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res* **2**: 130-143.
- Delio M, Guo T, McDonald-McGinn DM, Zackai E, Herman S, Kaminetzky M, Higgins AM, Coleman K, Chow C, Jalbrzikowski M, et al. 2013. Enhanced maternal origin of the 22q11.2 deletion in velocardiofacial and DiGeorge syndromes. *Am J Hum Genet* **92**: 439-447.
- Demaerel W, Hosseinzadeh M, Nouri N, Sedghi M, Dimitriadou E, Salehi M, Abdali H, Memarzadeh M, Zamani M, Vermeesch JR. 2016. Reciprocal 22q11.2 Deletion and Duplication in Siblings with Karyotypically Normal Parents. *Cytogenet Genome Res* **148**: 1-5.

- Dittwald P, Gambin T, Gonzaga-Jauregui C, Carvalho CMB, Lupski JR, Stankiewicz P, Gambin A. 2013a. Inverted Low-Copy Repeats and Genome Instability-A Genome-Wide Analysis. *Hum Mutat* **34**: 210-220.
- Dittwald P, Gambin T, Szafranski P, Li J, Amato S, Divon MY, Rodríguez Rojas LX, Elton LE, Scott D a, Schaaf CP, et al. 2013b. NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits. *Genome Res* **23**: 1395-1409.

## E

- Edelmann L, Pandita RK, Spiteri E, Funke B, Goldberg R, Palanisamy N, Chaganti RSK, Magenis E, Shprintzen RJ, Morrow BE. 1999. A common molecular basis for rearrangement disorders on chromosome 22q11. **8**: 1157-1167.
- Edelmann L, Spiteri E, Koren K, Pulijaal V, Bialer MG, Shanske A, Goldberg R, Morrow BE. 2001. AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am J Hum Genet* **68**: 1-13.
- Eichler EE, Nickerson DA, Altshuler D, Bowcock AM, Brooks LD, Carter NP, Church DM, Felsenfeld A, Guyer M, Lee C, et al. 2007. Completing the map of human genetic variation. *Nature* **447**: 161-165.
- Ensenauer RE, Adeyinka A, Flynn HC, Michels V V, Lindor NM, Dawson DB, Thorland EC, Lorentz CP, Goldstein JL, Mcdonald MT, et al. 2003. Microduplication 22q11.2 , an Emerging Syndrome : Clinical , Cytogenetic , and Molecular Analysis of Thirteen Patients. *Am J Hum Genet* **73**: 1027-1040.

## F

- Fernández L, Nevado J, Santos F, Heine-Suñer D, Martínez-Glez V, García-Miñaur S, Palomo R, Delicado A, Pajares IL, Palomares M, et al. 2009. A deletion and a duplication in distal 22q11.2 deletion syndrome region. Clinical implications and review. *BMC Med Genet* **10**: 48.
- Fernandez-Novoa MC, Vargas MT, Vizmanos JL, Garnacho C, Martínez JJ, Sanz P, Lluch D. 2001. Prader-Willi syndrome large deletion on two brothers. Is this the exception that confirm the rule? *Rev Neurol* **32**: 935-938.
- Feuk L. 2010. Inversion variants in the human genome: role in disease and genome architecture. *Genome Med* **2**: 11.
- Feuk L, Carson AR, Scherer SW. 2006. Structural variation in the human genome. *Nat Rev Genet* **7**: 85-97.

Franke L, El Bannoudi H, Jansen DTSL, Kok K, Trynka G, Diogo D, Swertz M, Franssen K, Knevel R, Gutierrez-Achury J, et al. 2015. Association analysis of copy numbers of FC-gamma receptor genes for rheumatoid arthritis and other immune-mediated phenotypes. *Eur J Hum Genet* **24**: 263-270.

Fusco F, Paciolla M, Pescatore A, Lioi Maria B, Ayuso C, Faravelli F, Gentile M, Zollino M, D'Urso M, Miano MG, et al. 2009. Microdeletion/duplication at the Xq28 IP locus causes a de novo IKKBKG/NEMO/IKKgamma exon4\_10 deletion in families with incontinentia pigmenti. *Hum Mutat* **30**: 1284-1291.



Garcia-Miñaur S, Fantes J, Murray RS, Porteous MEM, Strain L, Burns JE, Stephen J, Warner JP. 2002. A novel atypical 22q11.2 distal deletion in father and son. *J Med Genet* **39**: E62.

Gebhardt GS, Devriendt K, Thoelen R, Swillen A, Pijkels E, Fryns J-P, Vermeesch JR, Gewillig M. 2003. No evidence for a parental inversion polymorphism predisposing to rearrangements at 22q11.2 in the DiGeorge/Velocardiofacial syndrome. *Eur J Hum Genet* **11**: 109-111.

Gentile M, De Sanctis S, Cariola F, Spezzi T, Di Carlo A, Tontoli F, Lista F, Buonadonna AL. 2005. FISH approach to determine cat eye syndrome chromosome breakpoints of a patient with cat eye syndrome type II. *Eur J Med Genet* **48**: 33-39.

Gimelli G, Pujana MA, Patricelli MG, Russo S, Giardino D, Larizza L, Cheung J, Armengol L, Schinzel A, Estivill X, et al. 2003. Genomic inversions of human chromosome 15q11-q13 in mothers of Angelman syndrome patients with class II (BP2/3) deletions. *Hum Mol Genet* **12**: 849-858.

Gimelli S, Beri S, Drabkin H a, Gambini C, Gregorio A, Fiorio P, Zuffardi O, Gemmill RM, Giorda R, Gimelli G. 2009. The tumor suppressor gene TRC8/RNF139 is disrupted by a constitutional balanced translocation t(8;22)(q24.13;q11.21) in a young girl with dysgerminoma. *Mol Cancer* **8**: 52.

Godo A, Blanco J, Vidal F, Sandalinas M, Garcia-Guixé E, Anton E. 2015. Altered segregation pattern and numerical chromosome abnormalities interrelate in spermatozoa from Robertsonian translocation carriers. *Reprod Biomed Online* **31**: 79-88.

Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, et al. 2005. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **307**: 1434-1440.

Gotter AL. 2003. A palindrome-mediated mechanism distinguishes translocations involving LCR-B of chromosome 22q11.2. *Hum Mol Genet* **13**: 103-115.

Gotter AL, Nimmakayalu MA, Jalali GR, Hacker AM, Vorstman J, Conforto Duffy D, Medne L, Emanuel BS. 2007. A palindrome-driven complex rearrangement of 22q11.2 and 8q24.1 elucidated using novel technologies. *Genome Res* **17**: 470-481.

- Gregory TR. 2005. Synergy between sequence and size in large-scale genomics. *Nat Rev Genet* **6**: 699-708.
- Gu W, Zhang F, Lupski JR. 2008. Mechanisms for human genomic rearrangements. *Pathogenetics* **1**: 4.
- Guo X, Freyer L, Morrow B, Zheng D. 2011. Characterization of the past and current duplication activities in the human 22q11.2 region. *BMC Genomics* **12**: 71.



## H

- Hajianpour A, Murer-Orlando M, Docherty Z. 1991. Germ line mosaicism for chromosome 5 “Cri-du-chat” deletion? *Am J Hum Genet* **49**: 217.
- Hatchwell E, Long F, Wilde J, Crolla J, Temple K. 1998. Molecular Confirmation of Germ Line Mosaicism for a Submicroscopic Deletion of Chromosome 22q11. *Am J Med Genet* **78**: 103-106.
- Heng HHQ, Windle B, Tsui L-C. 2005. High-Resolution FISH Analysis. En *Current Protocols in Human Genetics*, Vol. Chapter 4 of, p. Unit 4.5, John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Hiroi N, Takahashi T, Hishimoto A, Izumi T, Boku S, Hiramoto T. 2013. Copy number variation at 22q11.2: from rare variants to common mechanisms of developmental neuropsychiatric disorders. *Mol Psychiatry* **18**: 1153-1165.
- Hobart HH, Morris CA, Mervis CB, Pani AM, Doris J, Rios CM, Kimberley KW, Gregg RG, Bray P. 2010. Inversion of the Williams syndrome region is a common polymorphism found more frequently in parents of children with Williams syndrome. *Am J Med Genet C Semin Med Genet* **154C**: 220-228.
- Human Segmental Duplication Database. <http://humanparalogy.gs.washington.edu/>
- Hussin J, Sinnott D, Casals F, Idaghdour Y, Bruat V, Saillour V, Healy J, Grenier JC, De Malliard T, Busche S, et al. 2013. Rare allelic forms of PRDM9 associated with childhood leukemogenesis. *Genome Res* **23**: 419-430.



## I

- Itsara A, Vissers LELM, Steinberg KM, Meyer KJ, Zody MC, Koolen DA., De Ligt J, Cuppen E, Baker C, Lee C, et al. 2012. Resolving the breakpoints of the 17q21.31 microdeletion syndrome with next-generation sequencing. *Am J Hum Genet* **90**: 599-613.



## J

- Jackson EM, Shaikh TH, Gururangan S, Jones MC, Malkin D, Nikkel SM, Zuppan CW, Wainwright LM, Zhang F, Biegel JA. 2007. High-density single nucleotide polymorphism array analysis in patients with germline deletions of 22q11.2 and malignant rhabdoid tumor. *Hum Genet* **122**: 117-27.
- Jefferson A, Volpi E V. 2010. Fluorescence in situ hybridization (FISH) for genomic investigations in rat. *Methods Mol Biol* **659**: 409-426.

## K

- Kara-Mostefa A, Raoul O, Lyonnet S, Amiel J, Munnich A, Vekemans M, Magnier S, Ossareh B, Bonnefont J. 1999. Recurrent Williams-Beuren Syndrome in a Sibship Suggestive of Maternal Germ-Line Mosaicism. *Am J Hum Genet* **64**: 1475-1478.
- Kehrer-Sawatzki H, Cooper DN. 2007. Structural divergence between the human and chimpanzee genomes. *Hum Genet* **120**: 759-778.
- Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F, et al. 2008. Mapping and sequencing of structural variation from eight human genomes. *Nature* **453**: 56-64.
- Kidd JM, Graves T, Newman TL, Fulton R, Hayden HS, Malig M, Kallicki J, Kaul R, Wilson RK, Eichler EE. 2010. A Human Genome Structural Variation Sequencing Resource Reveals Insights into Mutational Mechanisms. *Cell* **143**: 837-847.
- Kokkonen H, Leisti J. 2000. An unexpected recurrence of Angelman syndrome suggestive of maternal germ-line mosaicism of del(15)(q11q13) in a Finnish family. *Hum Genet* **107**: 83-85.
- Koolen DA, Dupont J, de Leeuw N, Vissers LELM, van den Heuvel SPA, Bradbury A, Steer J, de Brouwer APM, Ten Kate LP, Nillesen WM, et al. 2012. Two families with sibling recurrence of the 17q21.31 microdeletion syndrome due to low-grade mosaicism. *Eur J Hum Genet* **20**: 729-733.
- Kurahashi H, Emanuel BS. 2001. Unexpectedly high rate of de novo constitutional t(11;22) translocations. *Nat Genet* **29**: 11-12.
- Kurahashi H, Shaikh T, Takata M, Toda T, Emanuel BS. 2003. The constitutional t(17;22): another translocation mediated by palindromic AT-rich repeats. *Am J Hum Genet* **72**: 733-738.

## L

- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- Laurie MT, Bertout J a, Taylor SD, Burton JN, Jay A, Bielas JH. 2013. Simultaneous digital quantification and fluorescence-based size characterization of massively parallel sequencing libraries. *Biotechniques* **55**: 61-67.
- Lee JA, Carvalho CMB, Lupski JR. 2007. A DNA Replication Mechanism for Generating Nonrecurrent Rearrangements Associated with Genomic Disorders. *Cell* **131**: 1235-1247.
- Liu P, Carvalho CMB, Hastings PJ, Lupski JR. 2012. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* **22**: 211-220.
- Liu P, Lacaria M, Zhang F, Withers M, Hastings PJ, Lupski JR. 2011. Frequency of nonallelic homologous recombination is correlated with length of homology: evidence that ectopic synapsis precedes ectopic crossing-over. *Am J Hum Genet* **89**: 580-588.
- Lupski JR. 2007. Genomic rearrangements and sporadic disease. *Nat Genet* **39**: S43-47.

## M

- Marcinkowska-Swojak M, Klonowska K, Figlerowicz M, Kozłowski P. 2014. An MLPA-based approach for high-resolution genotyping of disease-related multi-allelic CNVs. *Gene* **546**: 257-262.
- Marques-Bonet T, Cheng Z, She X, Eichler EE, Navarro A. 2008. The genomic distribution of intraspecific and interspecific sequence divergence of human segmental duplications relative to human / chimpanzee chromosomal rearrangements. **20**.
- Marques-Bonet T, Girirajan S, Eichler EE. 2009a. The origins and impact of primate segmental duplications. *Trends Genet* **25**: 443-454.
- Marques-Bonet T, Ryder OA, Eichler EE. 2009b. Sequencing Primate Genomes : What Have We Learned ? *Annu Rev Genomics Hum Genet* **10**: 355-386.
- Martin RH. 2005. *Cytogenetics of Human Germ Cells*.
- Martínez-Pasarell O, Nogués C, Bosch M, Egozcue J, Templado C. 1999. Analysis of sex chromosome aneuploidy in sperm from fathers of Turner syndrome patients. *Hum Genet* **104**: 345-349.
- McDonald-McGinn DM, Sullivan KE. 2011. Chromosome 22q11.2 Deletion Syndrome (DiGeorge Syndrome/Velocardiofacial Syndrome). *Medicine (Baltimore)* **90**: 1-18.

- McDonald-McGinn DM, Sullivan KE, Marino B, Philip N, Swillen A, Vorstman JAS, Zackai EH, Emanuel BS, Vermeesch JR, Morrow BE, et al. 2015. 22q11.2 deletion syndrome. *Nat Rev Dis Prim* **1**: 15071.
- McDonald-McGinn DM, Zackai EH. 2008. Genetic counseling for the 22q11.2 deletion. *Dev Disabil Res Rev* **14**: 69-74.
- McVey M, Lee SE. 2008. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* **24**: 529-538.
- Merikangas AK, Segurado R, Heron EA, Anney RJL, Paterson AD, Cook EH, Pinto D, Scherer SW, Szatmari P, Gill M, et al. 2014. The phenotypic manifestations of rare genic CNVs in autism spectrum disorder. *Mol Psychiatry* **20**: 1366-1372.
- Mikhail FM, Burnside RD, Rush B, Ibrahim J, Godshalk R, Rutledge SL, Robin NH, Descartes MD, Carroll AJ. 2014. The recurrent distal 22q11.2 microdeletions are often de novo and do not represent a single clinical entity: a proposed categorization system. *Genet Med* **16**: 92-100.
- Mlynarski EE, Xie M, Taylor D, Sheridan MB, Guo T, Racedo SE, McDonald-McGinn DM, Chow EWC, Vorstman J, Swillen A, et al. 2016. Rare copy number variants and congenital heart defects in the 22q11.2 deletion syndrome. *Hum Genet* **135**: 273-285.
- Molina O. 2011. *Origen i recurrència de trastorns genòmics causats per delecions cromosòmiques*.
- Molina O, Anton E, Vidal F, Blanco J. 2011. Sperm rates of 7q11.23, 15q11q13 and 22q11.2 deletions and duplications: a FISH approach. *Hum Genet* **129**: 35-44.
- Molina O, Anton E, Vidal F, Blanco J. 2012a. High rates of de novo 15q11q13 inversions in human spermatozoa. *Mol Cytogenet* **5**: 11.
- Molina O, Blanco J, Anton E, Vidal F, Volpi E. 2012b. High-resolution fish on DNA fibers for low-copy repeats genome architecture studies. *Genomics* **100**: 380-386.
- Molina O, Blanco J, Vidal F. 2010. Deletions and duplications of the 15q11-q13 region in spermatozoa from Prader-Willi syndrome fathers. *Mol Hum Reprod* **16**: 320-328.
- Myers S, Freeman C, Auton A, Donnelly P, McVean G. 2008. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat Genet* **40**: 1124-1129.



- Nimmakayalu M a, Gotter AL, Shaikh TH, Emanuel BS. 2003. A novel sequence-based approach to localize translocation breakpoints identifies the molecular basis of a t(4;22). *Hum Mol Genet* **12**: 2817-2825.
- Nuttle X, Itsara A, Shendure J, Eichler EE. 2014. Resolving genomic disorder-associated breakpoints within segmental DNA duplications using massively parallel sequencing. *Nat Protoc* **9**: 1496-1513.

## O

- Osborne LR, Li M, Pober B, Chitayat D, Bodurtha J, Mandel A, Costa T, Grebe T, Cox S, Tsui L, et al. 2001. A 1.5 million–base pair inversion polymorphism in families with Williams-Beuren syndrome. *Nat Genet* **29**: 321-325.
- Ottaviani D, LeCain M, Sheer D. 2014. The role of microhomology in genomic structural variation. *Trends Genet* **30**: 85-94.
- Ou Z, Berg JS, Yonath H, Enciso VB, Miller DT, Picker J, Lenzi T, Keegan CE, Sutton VR, Belmont J, et al. 2008. Microduplications of 22q11.2 are frequently inherited and are associated with variable phenotypes. *Genet Med* **10**: 267-77.
- Ou Z, Stankiewicz P, Xia Z, Breman AM, Dawson B, Wiszniewska J, Szafranski P, Cooper ML, Rao M, Shao L, et al. 2011. Observation and prediction of recurrent human translocations mediated by NAHR between nonhomologous chromosomes. *Genome Res* **2**: 33-46.

## P

- Parks MM, Lawrence CE, Raphael BJ. 2015. Detecting non-allelic homologous recombination from high-throughput sequencing data. *Genome Biol* **16**: 72.
- Peng Z, Zhou W, Fu W, Du R, Jin L, Zhang F. 2014. Correlation between frequency of non-allelic homologous recombination and homology properties: evidence from homology-mediated CNV mutations in the human genome. *Hum Mol Genet* **24**: 1225-1233.
- Perry G, Yang F, Marques-Bonet T, Murphy C. 2008. Copy number variation and evolution in humans and chimpanzees. *Genome* 1698-1710.
- Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, et al. 2007. Diet and the evolution of human amylase gene copy number variation. *Nat Genet* **39**: 1256-1260.
- Plastira K, Angelopoulou R, Mantas D, Msaouel P, Lyrakou S, Plastiras A, Bolaris S, Baka S, Papanastasiadis N. 2007. The effects of age on the incidence of aneuploidy rates in spermatozoa of oligoasthenozoospermic patients and its relationship with ICSI outcome. *Int J Androl* **30**: 65-72.
- Portnoi M-F. 2009. Microduplication 22q11.2: a new chromosomal syndrome. *Eur J Med Genet* **52**: 88-93.
- Pratto F, Brick K, Khil P, Smagulova F, Petukhova GV, Camerini-Otero RD. 2014. Recombination initiation maps of individual human genomes. *Science (80- )* **346**: 1256442-1256442.
- Prestes Sartorelli EM, Mazzucatto LF, Monteiro de Pina-Neta J. 2001. Effect of paternal age on human sperm. *Fertil Steril* **76**: 1119-1123.

## R

- Rao P, Li W, Vissers LELM, Veltman J, Ophoff R. 2010. Recurrent inversion events at 17q21.31 microdeletion locus are linked to the MAPT H2 haplotype. *Cytogenet Genome Res* **129**: 275-279.
- Rauch A, Zink S, Zweier C, Thiel CT, Koch A, Rauch R, Lascorz J, Hüffmeier U, Weyand M, Singer H, et al. 2005. Systematic assessment of atypical deletions reveals genotype-phenotype correlation in 22q11.2. *J Med Genet* **42**: 871-876.
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al. 2006. Global variation in copy number in the human genome. *Nature* **444**: 444-454.
- Reiter LT, Hastings PJ, Nelis E, De Jonghe P, Van Broeckhoven C, Lupski JR. 1998. Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. *Am J Hum Genet* **62**: 1023-1033.
- Retterer K, Scuffins J, Schmidt D, Lewis R, Pineda-Alvarez D, Stafford A, Schmidt L, Warren S, Gibellini F, Kondakova A, et al. 2015. Assessing copy number from exome sequencing and exome array CGH based on CNV spectrum in a large clinical cohort. *Genet Med* **17**: 623-629.
- Ribeiro-Bicudo L a, de Campos Legnaro C, Gamba BF, Candido Sandri RM, Richieri-Costa A. 2013. Cognitive deficit, learning difficulties, severe behavioral abnormalities and healed cleft lip in a patient with a 1.2-mb distal microduplication at 22q11.2. *Mol Syndromol* **4**: 292-296.
- Röthlisberger B, Kotzot D. 2007. Recurrence risk in de novo structural chromosomal rearrangements. *Am J Med Genet Part A* **143A**: 1708-1714.

## S

- Saitta SC, Harris SE, Gaeth AP, Driscoll DA, McDonald-McGinn DM, Maisenbacher MK, Yersak JM, Chakraborty PK, Hacker AM, Zackai EH, et al. 2004. Aberrant interchromosomal exchanges are the predominant cause of the 22q11.2 deletion. *Hum Mol Genet* **13**: 417-428.
- Sánchez J, Fernández R, Madruga M, Bernabeu-Wittel J, Antiñolo G, Borrego S. 2013. Somatic and germ-line mosaicism of deletion 15q11.2-q13 in a mother of dizygotic twins with Angelman syndrome. *Am J Med Genet A* **164A**: 370-376.
- Sandrin-Garcia P, Macedo C, Lr M, Es R. 2002. Recurrent 22q11.2 deletion in a sibship suggestive of parental germline mosaicism in velocardiofacial syndrome. *Clin Genet* **61**: 380-383.

- Sarrate Z, Blanco J, Egozcue S, Vidal F, Egozcue J. 2004. Identification of meiotic anomalies with multiplex fluorescence in situ hybridization: Preliminary results. *Fertil Steril* **82**: 712-717.
- Sasaki M, Lange J, Keeney S. 2010. Genome destabilization by homologous recombination in the germ line. *Nat Rev Mol Cell Biol* **11**: 182-195.
- Scherer SW, Gripp KW, Lucena J, Nicholson L, Bonnefont J, Pérez-Jurado L, Osborne LR. 2005. Observation of a parental inversion variant in a rare Williams-Beuren syndrome family with two affected children. *Hum Genet* **117**: 383-388.
- Sellier C, Hwang VJ, Dandekar R, Durbin-Johnson B, Charlet-Berguerand N, Ander BP, Sharp FR, Angkustsiri K, Simon TJ, Tassone F. 2014. Decreased DGCR8 expression and miRNA dysregulation in individuals with 22q11.2 deletion syndrome. *PLoS One* **9**: 13-15.
- Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, et al. 2000. Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* **9**: 489-501.
- Sharp AJ, Mefford HC, Li K, Baker C, Skinner C, Stevenson RE, Schroer RJ, Novara F, De Gregori M, Ciccone R, et al. 2008. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* **40**: 322-328.
- Sheridan MB, Kato T, Haldeman-Englert C, Jalali GR, Milunsky JM, Zou Y, Klaes R, Gimelli G, Gimelli S, Gemmill RM, et al. 2010. A palindrome-mediated recurrent translocation with 3:1 meiotic nondisjunction: the t(8;22)(q24.13;q11.21). *Am J Hum Genet* **87**: 209-218.
- Shimojima K, Imai K, Yamamoto T. 2010a. A de novo 22q11.22q11.23 interchromosomal tandem duplication in a boy with developmental delay, hyperactivity, and epilepsy. *Am J Med Genet A* **152A**: 2820-2826.
- Shimojima K, Inoue T, Hoshino A, Kakiuchi S, Watanabe Y, Sasaki M, Nishimura A, Takeshita-Yanagisawa A, Tajima G, Ozawa H, et al. 2010b. Comprehensive genetic analyses of PLP1 in patients with Pelizaeus-Merzbacher disease applied by array-CGH and fiber-FISH analyses identified new mutations and variable sizes of duplications. *Brain Dev* **32**: 171-179.
- Sloter ED, Lowe X, Moore II DH, Nath J, Wyrobek AJ. 2000. Multicolor FISH analysis of chromosomal breaks, duplications, deletions, and numerical abnormalities in the sperm of healthy men. *Am J Hum Genet* **67**: 862-872.
- Sloter ED, Marchetti F, Eskenazi B, Weldon RH, Nath J, Cabrerros D, Wyrobek AJ. 2007. Frequency of human sperm carrying structural aberrations of chromosome 1 increases with advancing age. *Fertil Steril* **87**: 1077-1086.
- Stalmans I, Lambrechts D, De Smet F, Jansen S, Wang J, Maity S, Kneer P, von der Ohe M, Swillen A, Maes C, et al. 2003. VEGF: a modifier of the del22q11 (DiGeorge) syndrome? *Nat Med* **9**: 173-182.
- Stankiewicz P, Lupski JR. 2002. Genome architecture, rearrangements and genomic disorders. *Trends Genet* **18**: 74-82.
- Stankiewicz P, Lupski JR. 2010. Structural variation in the human genome and its role in disease. *Annu Rev Med* **61**: 437-455.

Sugawara H, Harada N, Ida T, Ishida T, Ledbetter DH, Yoshiura KI, Ohta T, Kishino T, Niikawa N, Matsumoto N. 2003. Complex low-copy repeats associated with a common polymorphic inversion at human chromosome 8p23. *Genomics* **82**: 238-244.

Sung P, Klein H. 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**: 739-750.



T

Tan TY, Collins A, James PA, McGillivray G, Stark Z, Gordon CT, Leventer RJ, Pope K, Forbes R, Crolla J a, et al. 2011. Phenotypic variability of distal 22q11.2 copy number abnormalities. *Am J Med Genet A* **155A**: 1623-1633.

Templado C, Donate A, Giraldo J, Bosch M, Estop A. 2011. Advanced age increases chromosome structural abnormalities in human spermatozoa. *Eur J Hum Genet* **19**: 145-151.

The International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* **437**: 1299-320.

Thomas NS, Durkie M, Potts G, Sandford R, Van Zyl B, Youings S, Dennis NR, Jacobs PA. 2006. Parental and chromosomal origins of microdeletion and duplication syndromes involving 7q11.23, 15q11-q13 and 22q11. *Eur J Hum Genet* **14**: 831-837.

Torgasheva AA, Borodin PM. 2010. Synapsis and recombination in inversion heterozygotes. *Biochem Soc Trans* **38**: 1676-1680.

Torres-Juan L, Rosell J, Sánchez-de-la-Torre M, Fibla J, Heine-Suñer D. 2007. Analysis of meiotic recombination in 22q11.2, a region that frequently undergoes deletions and duplications. *BMC Med Genet* **8**: 14.



V

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al. 2001. The Sequence of the Human Genome. **291**.

Verhagen JMA, Diderich KEM, Oudesluijs G, Mancini GMS, Eggink AJ, Verkleij-Hagoort AC, Groenenberg I a L, Willems PJ, du Plessis F a, de Man S a, et al. 2012. Phenotypic variability of atypical 22q11.2 deletions not including TBX1. *Am J Med Genet A* **158A**: 2412-2420.

Visser R, Shimokawa O, Harada N, Kinoshita A, Ohta T, Niikawa N, Matsumoto N. 2005. Identification of a 3.0-kb major recombination hotspot in patients with Sotos syndrome who carry a common 1.9-Mb microdeletion. *Am J Hum Genet* **76**: 52-67.

Vorsanova SG, Yurov YB, Iourov IY. 2010. Human interphase chromosomes: a review of available molecular cytogenetic technologies. *Mol Cytogenet* **3**: 1.

## Y

- Yagi H, Furutani Y, Hamada H, Sasaki T, Asakawa S, Minoshima S, Ichida F, Joo K, Kimura M, Imamura S ichiro, et al. 2003. Role of TBX1 in human del22q11.2 syndrome. *Lancet* **362**: 1366-1373.
- Yang SP, Bidichandani SI, Figuera LE, Juyal RC, Saxon PJ, Baldini A, Patel PI. 1997. Molecular analysis of deletion (17)(p11.2p11.2) in a family segregating a 17p paracentric inversion: implications for carriers of paracentric inversions. *Am J Hum Genet* **60**: 1184-1193.

## Z

- Zarrei M, MacDonald JR, Merico D, Scherer SW. 2015. A copy number variation map of the human genome. *Nat Rev Genet* **16**: 172-183.
- Zhang F, Gu W, Hurles ME, Lupski JR. 2009. Copy Number Variation in Human Health, Disease, and Evolution. *Annu Rev Genomics Hum Genet* **10**: 451-481.
- Zhang L, Lu HHS, Chung W, Yang J, Li W-H. 2005. Patterns of segmental duplication in the human genome. *Mol Biol Evol* **22**: 135-141.
- Zhang Y-B, Li X, Zhang F, Wang D-M, Yu J. 2012. A Preliminary Study of Copy Number Variation in Tibetans. *PLoS One* **7**: e41768.
- Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. 2013. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics* **14 Suppl 1**: S1.





# ABREVIACIONS

<b>array-CGH</b>	CGH en xips de DNA
<b>BAC</b>	<i>Bacterial Artificial Chromosome</i>
<b>CGH</b>	<i>Comparative genome hybridization</i> Hibridació genòmica comparada
<b>CNV</b>	<i>Copy Number Variations</i> Variacions en el número de còpies
<b>Cy3</b>	<i>Cyanine 3</i>
<b>DAPI</b>	4',6-diamino-2-fenilindol
<b>DMSO</b>	Dimetilsulfòxid
<b>ddPCR</b>	<i>droplet digital PCR</i>
<b>DTT</b>	1,4-ditiotreitol
<b>DECHIPHER</b>	<i>DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources</i>
<b>del</b>	Delecions
<b>del+dup</b>	Delecions i duplicacions
<b>DNA</b>	<i>Deoxyribonucleic acid</i> Àcid desoxiribonucleic
<b>dNTP</b>	Desoxinucleòtids trifosfat
<b>DMC1</b>	<i>Disrupted Meiotic cDNA protein 1</i>
<b>DSB</b>	<i>Double-Stranded Breaks</i> Trencaments de doble cadena del DNA
<b>DTT</b>	1,4-ditiotreitol
<b>dup</b>	Duplicacions
<b>dUTP</b>	Desoxiuridina trifosfat

<b>EDTA</b>	Àcid etilendiamintetraacètic
<b>FAM</b>	<i>Fluorescein amidite</i>
<b><i>fiber</i>-FISH</b>	FISH d'alta resolució sobre fibres de cromatina
<b>FISH</b>	<i>Fluorescence In Situ Hybridization</i> Hibridació <i>in situ</i> fluorescent
<b>FITC</b>	<i>Fluorescein isothiocyanate</i>
<b>FoSTeS</b>	<i>Fork Stalling and Template Switching</i> Força de replicació i canvi de motlle
<b>HaeIII</b>	Tercera endonucleasa del bacteri <i>Haemophilus aegyptius</i>
<b>HEX</b>	<i>Hexachloro-6-carboxyfluorescein</i>
<b>HR</b>	<i>Homologous Recombination</i> Recombinació homòloga
<b>inv</b>	Haplotip inversió 22q11.2
<b>kb</b>	Kilobases
<b>KRAB</b>	<i>KRüppel-Associated Box</i>
<b>LB</b>	Luria-Bertani
<b>LCR</b>	<i>Low-copy repeat</i>
<b>LINEs</b>	<i>Long Interspersed Nuclear Elements</i>
<b>LSI</b>	<i>Locus Specific Identifier</i>
<b>LTR</b>	<i>Long Terminal Repeat</i>
<b>MAPH</b>	<i>Multiplex Amplifiable Probe Hybridization</i>
<b>Mb</b>	Megabases
<b>MER67B</b>	<i>MEdium Reiterated 67B</i>
<b>MEPS</b>	<i>Minimal Efficient Processing Segments</i> Segments mínims complementaris
<b>M-FISH</b>	<i>Multiplex FISH</i>
<b>MLPA</b>	<i>Multiplex Ligation-dependent Probe Amplification</i>

<b>MMEJ</b>	<i>Microhomology-Mediated End Joining</i>
<b>MSH2</b>	<i>MutS protein Homolog 2</i>
<b>N</b>	Haplotip normal
<b>NA</b>	No avaluables
<b>NAHR</b>	<i>Non-Allelic Homologous Recombination</i> Recombinació homòloga no al·lèlica
<b>NHEJ</b>	<i>Non-Homologous End-Joining</i> Fusió d'extremes no homòlegs
<b>OMIM</b>	<i>Online Mendelian Inheritance in Man</i>
<b>pb</b>	Parells de bases
<b>PBS</b>	<i>Phosphate Buffered Saline</i>
<b>PCR</b>	<i>Polymerase Chain Reaction</i> Reacció en cadena de la polimerasa
<b>PMS1</b>	<i>PostMeiotic Segregation protein 1</i>
<b>QMPSF</b>	<i>Quantitative Multiplex PCR of Short Fluorescent fragments</i>
<b>Q-PCR</b>	PCR quantitativa en temps real
<b>RAD51</b>	<i>DNA repair protein 51</i>
<b>RNA</b>	<i>Ribonucleic acid</i> Àcid ribonucleic
<b>RPMI</b>	<i>Roswell Park Memorial Institute</i>
<b>SD</b>	<i>Segmental duplication</i> Duplicació segmental
<b>SDG/VCF</b>	Síndrome de DiGeorge/velocardiofacial
<b>SDSA</b>	<i>Synthesi Dependent Strand Annealing</i> Aparellament de cadena associat a síntesi
<b>SIDA</b>	Síndrome d'ImmunoDeficiència Adquirida
<b>SINEs</b>	<i>Short Interspersed Nuclear Elements</i>
<b>SNP</b>	<i>Single Nucleotide Polymorphisms</i>
<b>SSC</b>	<i>Saline-Sodium Citrate</i>

**VIH** Virus de la Immunodeficiència Humana

## RECOPILATORI DE GENS I PSEUDOGENS

<b><i>AMY1</i></b>	<i>Salivary amylase</i>	OMIM 104700
<b><i>ARSA</i></b>	<i>Arylsulfatase A</i>	OMIM 607574
<b><i>C4B</i></b>	<i>Complement Component 4b</i>	OMIM 120820
<b><i>CCL3L1</i></b>	<i>Chemokine ligand 3 like1</i>	OMIM 601395
<b><i>CECR7</i></b>	<i>Cat Eye syndrome Chromosome Region, candidate 7</i>	
<b><i>DEFB</i></b>	<i>Beta-defensin</i>	OMIM 615997
<b><i>DGCR6</i></b>	<i>DiGeorge syndrome Critical Region gene 6</i>	OMIM 601279
<b><i>DGCR8</i></b>	<i>DiGeorge syndrome Critical Region gene 8</i>	OMIM 609030
<b><i>EIF2C1</i></b>	<i>Eukaryotic translation Initiation Factor 2C, 1</i>	OMIM 606228
<b><i>ELN</i></b>	<i>Elastina</i>	OMIM 130160
<b><i>Factor VIII</i></b>	<i>Factor de coagulació VIII</i>	OMIM 300841
<b><i>FCGR3B</i></b>	<i>Fc gamma receptor 3B</i>	OMIM 610665
<b><i>FOXF1</i></b>	<i>Forkhead Box F1</i>	OMIM 601089
<b><i>GAB4</i></b>	<i>GRB2-Associated Binding protein family, member 4</i>	
<b><i>GGT3P</i></b>	<i>Gamma-Glutamyltransferase 3 Pseudogene</i>	
<b><i>HBD-2</i></b>	<i>Hemoglobin-Delta Locus</i>	OMIM 142000
<b><i>IRGM</i></b>	<i>Immunity-Related GTPase family, M</i>	OMIM 608212
<b><i>IKBKG</i></b>	<i>Inhibitor of Kappa light polypeptide gene enhancer in B cells, Kinase of, Gamma</i>	OMIM 300248
<b><i>LDLR</i></b>	<i>Low Density Lipoprotein Receptor</i>	OMIM 606945

<b>LIMK</b>	<i>Lim Domain Kinase 1</i>	OMIM 601329
<b>LPA</b>	Lipoproteïna A	OMIM 152200
<b>NCF1</b>	<i>Neutrophil Cytosolic Factor 1</i>	OMIM 608512
<b>PI4KA</b>	<i>Phosphatidylinositol 4-Kinase, Catalytic, Alpha</i>	OMIM 600286
<b>PI4KAP2</b>	<i>Phosphatidylinositol 4-Kinase, Catalytic, Alpha Pseudogene 2</i>	
<b>PPIL2</b>	<i>Peptidyl-Prolyl Isomerase-Like 2</i>	OMIM 607588
<b>PRDM9</b>	<i>PR Domain-Containing Protein 9</i>	OMIM 609760
<b>PRODH</b>	<i>Proline dehydrogenase (oxidase)</i>	OMIM 606810
<b>PRSS1</b>	<i>Protease Serine 1</i>	OMIM 276000
<b>RIMBP3C</b>	<i>RIMS binding protein 3C</i>	OMIM 612701
<b>TBX1</b>	<i>T-Box 1</i>	OMIM 602054
<b>TUPLE1</b>	<i>TUP-Like Enhancer of aplit 1</i>	OMIM 600237
<b>UBE2L3</b>	<i>Ubiquitin-Conjugating Enzyme E2l 3</i>	OMIM 603721
<b>USP18</b>	<i>Ubiquitin-Specific Protease 18</i>	OMIM 607057
<b>VEGF</b>	<i>Vascular Endothelial Growth Factor</i>	OMIM 192240
<b><math>\alpha</math>-globina</b>	<i>Hemoglobin--alpha Locus</i>	OMIM 141800

EL GENOMA HUMÀ ESTÀ FORMAT EN UN 5% PER *LOW-COPY REPEATS* (LCR), SEGMENTS DE DNA D'ENTRE 1 I 500 KB QUE ES REPETEIXEN DUES O MÉS VEGADES AL LLARG DEL GENOMA I QUE COMPARTeixEN UNA HOMOLOGIA SUPERIOR AL 90%. L'APARELLAMENT NO AL·LÈLIC D'AQUESTES SEQÜENCIES DURANT LA MEIOSI PROMOU LA GENERACIÓ D'ESDEVENIMENTS DE RECOMBINACIÓ HOMÒLOGA NO AL·LÈLICA (NAHR) QUE DONEN LLOC A REORGANITZACIONS CROMOSÒMIQUES HERETABLES QUE ORIGINEN UN CONJUNT DE MALALTIES ANOMENADES TRASTORNS GENÒMICS. LA DELECIÓ DE LA REGIÓ 22q11.2, ORIGINADA PER LA NAHR ENTRE ELS LCR22, CAUSA LA SÍNDROME DE DiGEORGE/VELOCARDIOFACIAL (SDG/VCF), UN DELS TRASTORNS GENÒMICS MÉS FREQÜENTS EN ELS HUMANS. L'OBJECTIU PRINCIPAL D'AQUESTA TESI DOCTORAL ÉS L'AVAUACIÓ DE LES CARACTERÍSTIQUES GENÈTIQUES QUE MODULEN LA SUSCEPTIBILITAT A LA NAHR EN INDIVIDUS TRANSMISSORS DE LA SDG/VCF.



UAB

Genetics  
of Male  
Fertility