Nested Inversion Polymorphisms Predispose Chromosome 22q11.2 to Meiotic Rearrangements

Wolfram Demaerel,1 Matthew S. Hestand,1 Elfi Vergaelen,1 Ann Swillen,1 Marcos López-Sánchez,2,3,4 Luis A. Pérez-Jurado,2,3,4 Donna M. McDonald-McGinn,5,6 Elaine Zackai,5,6 Beverly S. Emanuel,5,6 Bernice E. Morrow,7 Jeroen Breckpot,1 Koenraad Devriendt,1 Joris R. Vermeesch,1,* and International 22q11.2 Brain and Behavior Consortium

Inversion polymorphisms between low-copy repeats (LCRs) might predispose chromosomes to meiotic non-allelic homologous recombination (NAHR) events and thus lead to genomic disorders. However, for the 22q11.2 deletion syndrome (22q11.2DS), the most common genomic disorder, no such inversions have been uncovered as of yet. Using fiber-FISH, we demonstrate that parents transmitting the de novo 3 Mb LCR22A–D 22q11.2 deletion, the reciprocal duplication, and the smaller 1.5 Mb LCR22A–B 22q11.2 deletion carry inversions of LCR22B–D or LCR22C–D. Hence, the inversions predispose chromosome 22q11.2 to meiotic rearrangements and increase the individual risk for transmitting rearrangements. Interestingly, the inversions are nested or flanking rather than coinciding with the deletion or duplication sizes. This finding raises the possibility that inversions are a prerequisite not only for 22q11.2 rearrangements but also for all NAHR-mediated genomic disorders.

De novo 22q11.2 deletions (MIM: 192430) occur with an estimated frequency of 1/3,000–1/6,000 live births1,2 and 1/1,000 pregnancies.3 Major clinical findings include palatal abnormalities; immunodeficiency; hypocalcemia; congenital heart disease; gastrointestinal, renal, and skeletal problems; mild dysmorphic features; developmental delays; and learning and behavioral difficulties including ADHD, anxiety disorders, autism, and psychotic illnesses, such as schizophrenia.4–6 The 22q11.21 region contains four LCRs associated with the condition; these are termed LCR22A, LCR22B, LCR22C, and LCR22D.7,8 Individual LCR22s show a heterogeneous architecture comprised of many repeat subunits9–11 (Figure S1). The paralogous LCR22s show a heterogeneous architecture comprised of many repeat subunits9–11 (Figure S1). The paralogous sequences show similarity of up to 99.6%, driving meiotic NAHR.12,13 LCR22A and LCR22D have the most complex organization and flank the typical ~3 Mb deletion observed in ~90% of individuals.8,11 In the remaining individuals, smaller nested deletions are observed.13

Reciprocal duplications are the meiotic NAHR complement and are causal for the 22q11.2 duplication syndrome (MIM: 608363).15 Inversion polymorphisms in the genome might make NAHR more likely and have been demonstrated to exist between flanking LCRs for many genomic disorders, including Williams-Beuren (7q11.23, MIM: 194050), Prader-Willi/Angelman (15q11-q13, MIM: 176270), 17q11.31 microdeletion (MIM: 610443), and 8p23.1 microdeletion (MIM: 222400) syndromes.17–20 Not surprisingly, we and others have searched for the existence of an inversion polymorphism in 22q11.2. Because inversion sizes underlying other genomic disorders are largely similar to deletion sizes, interphase FISH experiments were performed within and flanking the LCR22A–D interval in parents of individuals with de novo 3 Mb 22q11.2 deletions. No inversion polymorphism between LCR22A and LCR22D could be identified.21,22

Scrubutizing the LCR22 subunit organization in the reference genome showed the potential presence of paralogous sequences of inverted, duplicated subunits.21 These make the existence of inversions possible.24 In addition, we applied the algorithm InvClust25 to SNP data from the 1000 genomes26 population to predict and determine the global distribution of the inverted allele on the basis of reduced recombination patterns. Variants with a MAF below 0.1 were discarded, and window sizes ranged from 0.2 to 3 Mb by 0.2 Mb steps. We then used InvClust to predict haplotypic signatures induced by putative inversions.27 Multidimensional scaling analysis was targeted to genotypes between LCR22B and LCR22D and between LCR22C and LCR22D. We kept the first two eigen-components and used a mixture-model classification with the invClust algorithm. We assigned inversion genotypes to each cluster whenever the clustering confidence for each sample was greater than 99%. We detected several positive signals (BIC > 0) of LD differences at 0.2–3 Mb window sizes. The most significant of those signals corresponds to the 0.4 Mb segment flanked by LCR22C and LCR22D. This suggests the existence of a nested inversion polymorphism between LCR22C and LCR22D (Figure S2). In addition, several approaches to mapping inversions genome-wide have been developed.28–30 On the basis of mapping information of paired-end sequences, an inversion is

*Correspondence: joris.vermeesch@kuleuven.be
© 2017
predicted between LCR22B and LCR22D in three out of nine individuals (invFEST database). To ascertain the validity of these predictions, we designed a fiber-FISH assay and applied it on immortalized EBV-transformed lymphoblastoid cell lines derived from peripheral blood samples. To detect potential inversions proximal to LCR22D, we designed three fluorescently labeled oligonucleotide probes flanking the LCR (Figure 1). The probes were labeled with red (Cy3), green (FITC), and a mix of both, generating a yellow signal (Agilent Technologies, California, USA). Fibers were counterstained with DAPI.

DNA fibers were stretched as previously described. Visualization was performed with an epifluorescence microscope (Zeiss Axioplan) at 600x magnification. Using Cytovision (Leica Biosystems, Nussloch), we recorded signals if (1) all three or four fluorescent signal strings were ordered in a consecutive and linear way, (2) distances between signal strings were concordant with the design, if LCR variability was taken into account, and (3) local fiber density was low, such that the chance of detecting "hybrid" signals from different fibers was reduced. Hybridization of this pattern on a control cell line (GM19238) resulted in ten fibers with the predicted wild-type pattern in the order of

Figure 1. Design and Fluorescence Patterns
(A) First probe design aligned with schematic representation of the LCR22B-D region. LCR22s are depicted with the Segmental Dups track in UCSC. Locations of known segmental duplications between LCR22B, LCR22C, and LCR22D are shown as connecting, colored lines. Red and green oligos (hg19, chr22:21,104,802–21,311,933 and chr22:21,223,781–21,424,739) partially label the unique sequence proximally from LCR22D and have 50 Kb of overlapping sequence depicted by the central yellow bar. A reference probe distal to LCR22D (chr22:21,931,955–22,124,705) is labeled with a green-red mix, shown as a yellow bar. Oligos hybridized to a reference allele show the expected wild-type (WT) sequence of red, overlapping yellow, and green signal strings. Subsequently, an unlabeled gap marks the position of LCR22D, followed by the mixed reference signal. Inversion (INV) signals show a signal sequence of green, overlapping yellow, and red, followed by the LCR22D gap and the mixed reference probe.

(B) Second probe design aligned with the LCR22B-D region. An additional green probe, starting proximally from LCR22C to the middle of the repeat (chr22:20,499,302–20,699,303), is added. From this location, a red (chr22:20,699,530–20,899,531) and a green (chr22:20,890,041–21,103,073) probe are designed to consecutively label 200 Kb each. A red reference probe marks the region distal to LCR22D (chr22:21,572,091–21,773,001). Probe patterns suggest a LCR22C–D inversion polymorphism.

(C) A probe pattern suggesting a LCR22B–D inversion.
red, green, and yellow (Figure 1A). In addition, one fiber had an inversion signal in the order green, red, and yellow (Figure 1A). LCR22D can be seen as a gap in the string between the green and yellow signals (Figure 1A). To determine the frequency of false-positive signals for an inversion, we used the assay to screen two hydatidiform mole hTERT-immortalized cell lines (CHM1 and CHM13). These hydatidiform molar cell lines are haploid. Hence, either all or none of the fibers should be inverted. 1/20 (5%) and 4/23 (17.39%) of the fiber patterns presented inverted alleles in the CH1 and CHM13 cell lines, respectively. Because the moles are presumably wild-type, we consequently defined cells with less than 20% inverted signals to be wild-type and cells with more than 20% inverted signals to have the inversion.

Because our a priori assumption was that an inversion in a parent would predispose a child to having a 22q11.2 deletion, we initially performed the assay on family 1 (Table S1). Both parents are unaffected and have a child with a de novo 22q11.2 deletion. Fiber-FISH showed that one of the parents in family 1 had two different patterns: the wild-type pattern and the pattern indicative for an inversion, we initially performed the assay on family 1 (Table S1). Both parents are unaffected and have a child with a de novo 22q11.2 deletion. Fiber-FISH showed that one of the parents in family 1 had two different patterns: the wild-type pattern and the pattern indicative for an inversion proximal to LCR22D (Figure 1A). On DNA fibers from the other parent, only the wild-type pattern was observed. Short tandem-repeat polymorphism analysis of both parents and a sibling demonstrated that the 22q11.2 deletion was derived from chromosome 22 of the parent carrying the inversion.

To confirm that this inversion polymorphism would predispose a parent to having a child with a de novo 22q11.2 rearrangement, we recruited an additional 16 trios. All individuals in the study were informed of the project’s outlines and gave written consent for their EBV cell lines and DNA to be used for sequencing and genotyping purposes. The study was approved by the medical ethics committee of the University hospital of Katholieke Universiteit Leuven. The research study subjects had been diagnosed on the basis of a clinical examination and molecular testing, by FISH with N25 probe or with TUPLE1 and ARSA probes (Abbot Molecular), by Illumina HumanCytoSNP-12 V2.1 BeadChip analysis, or by analysis with the MLPA SALSA P250 DiGeorge diagnostic probe kit (MRC-Holland).

All probands had a de novo 22q11.2 rearrangement, whereas their parents were unaffected. Fifteen probands carried the most common 3 Mb LCR22A–D deletion. One other proband carried the 1.5 Mb LCR22A–B deletion, and an additional proband had an LCR22A–D duplication (Table S1). To determine the parents-of-origin of the 22q11.2 deletions, we amplified four microsatellites in the deleted region and two outside. All marker sizes were compared visually for trios, and the parent-of-origin was determined. At least three STRs were informative in every family, and paternity was confirmed for all (data not shown).

Lymphoblastoid cell lines of all parents were screened for the presence or absence of an inversion. In one parent-of-origin only, the regular string of FISH signals was observed indicating that this parent matched the homozygous reference structure, whereas the remaining sixteen parents-of-origin were carriers of an inverted allele (94.1%). To confirm the fiber-FISH findings, we also performed interphase FISH on a subset of cells derived from twelve individuals (Figure S3). In the six individuals fiber-FISH suggested to be wild-type, inversion signals were observed in 10%-15% of nuclei. A small percentage of apparent inversions is expected as a result of chromatin folding. In all seven individuals suggested to be heterozygous for the inversion by fiber-FISH, interphase inversion signals were observed in 36%-47% of the nuclei. Because this is significantly higher than the percentage in wild-types, we conclude that the interphase results are concordant with the fiber-FISH data (Table S1).

Because 16 out of 17 parents-of-origin carry an inversion, it is highly likely that the inversion is a risk factor for 22q11.2 rearrangements. Surprisingly, the inversion was also detected in eight out of 17 non-transmitting parents (47.1%) who had passed the intact chromosome 22 to the affected child (Table S1). This observation suggested that the inversion is a common polymorphism in the population.

To determine the population frequency, we screened 11 additional individuals (Table S1). Seven random EBV cell lines established during unrelated routine tests by the hospital were used as normal population controls. Three additional HapMap control cell lines (Coriell Cell Repository) were randomly picked from those available in the lab and cultured according to standard protocols.

Out of the 13 controls, 11 random individuals and the two haploid cell lines, five were detected as heterozygous (38.5%), a proportion similar to the occurrence of the inversion in the non-transmitting parent. When combined with the parental ratios, the population incidence of the inversion allele is 45%. In contrast, the incidence of the inversion in the parent-of-origin for the children affected with the 22q11.2 deletion or duplication is 94% (Figure 2). This is significantly higher than the occurrence in the general population (Fisher exact p = 0.0012). Hence, the 22q11.2 inversion polymorphism makes NAHR events on chromosome 22q11.2 more likely. The population incidence of 22q11DS is about 1/3,000. If we assume that the inversion is a prerequisite for NAHR, the deletion should only occur in inversion carriers, which make up 47.9% of the population. As a consequence, the individual risk for an inversion carrier is about twice the population incidence. Thus, a parent that is a carrier of an inversion has an estimated chance of 1/1,350 of having a child with 22q11DS. A variable frequency of the inversion allele might also underlie differences in the 22q11DS frequency among populations.36

We hypothesized that the inversions observed in the fiber-FISH probe pattern could be the result of both LCR22C–D and LCR22B–D inversions. To specify which of these rearrangements occurs, we designed a second
The inversion is considered a susceptibility factor. However, whereas inversions in these genomic disorders are larger or of similar size to the region of deletion or duplication, we demonstrate here that 22q11.2 inversions are nested in or even flanking the larger rearrangements for which they cause a predisposition. Perhaps also all parents-of-origin of the Williams-Beuren syndrome and the Prader-Willi and Angelman syndromes carry nested or flanking small inversion polymorphisms that remain undetected. In the Williams-Beuren syndrome region, three LCRs are present. One third of Williams-Beuren syndrome parents-of-origin have a 1.79 Mb to 2.56 Mb proximal-distal inversion that predisposes them to having a child with the most-common 1.55 Mb proximal-medial LCR deletions. Interestingly, two additional small inversion polymorphisms have been reported between the medial and distal LCRs. These are hypothesized to increase susceptibility to nested deletions in the distal LCRs. The potential existence of these inversions in parents-of-origin has not yet been investigated.

Two mechanisms have been invoked to explain how inversions trigger NAHR. An inversion can lead to unstable asynaptic bubbles during meiosis; these bubbles are prone to breaks and secondary rearrangements. Alternatively, the inversion could re-orient the paralogous subunits from being in opposing to be in identical direction. NAHR of subunits in opposing orientation would result in an inversion, whereas NAHR of subunits in an identical orientation would generate deletions or duplications. In addition, the inversion could cause paralogous sequences to be more equally spread among LCRs, resulting in longer stretches of similarity and hence a higher chance for nonallelic homologous pairing and recombination. Thus, the risk of NAHR might increase between different LCR22s depending on the exact position of the inversion breakpoints within the LCRs. Because all but one of the parents-of-origin carry an inversion, we speculate that the inversion polymorphisms might be a necessary prerequisite to any of the known 22q11.2 rearrangements. In the one parent-of-origin without an inversion of LCR22B–D, we speculate that there might be another structural variation, proximal or distal to the LCR22B–D interval.

Surprisingly, no individuals homozygous or hemizygous for the inversion were observed. If the 29 diploid population samples and non-transmitting parents were in Hardy-Weinberg equilibrium, at least one individual should be homozygous for the inversion (chi-square test: $p = 0.1198$). Additionally, in eight individuals for whom both parents are inversion carriers, none carried the inversion ($p = 0.0039$). One possible explanation could be that the inversion breakpoints alter or disrupt the function or regulation of one or multiple essential genes. LCR22s contain four known genes, USP18 (Ubiquitin Specific Peptidase 18, MIM: 607057), GGT (Gamma Glutamyltranspeptidase, MIM: 612346), GGTS (Gamma Glutamyltransferase 5, MIM: 137168) and BCR (Breakpoint Cluster Region, MIM: 151410). Each has been (partially) duplicated during primate evolution and transposed on 22q11.2 to create a mosaic of largely pseudogenes. Additionally, inversion polymorphisms are known to affect topological associations between functional genes and their regulating elements, leading to more-distant effects. For several other genomic disorders, it has been demonstrated that genes within LCRs can modulate

---

Figure 2. Genotype Distributions in Parents-of-Origin, the Non-transmitting parents, a Population Sample, and Eight Individuals Whose Parents Are Both Carriers of an Inversion Polymorphism

No individuals homozygous for the inversion were detected, nor did any of the individuals carry a remaining allele with an inversion.

probe set containing an additional green-labeled oligonucleotide probe in the region between LCR22B and LCR22C. Red and green probes were extended in comparison to the first pattern, and the reference probe distal to LCR22D was only labeled red. The proximal green and red probes were designed so that they would both extend to the middle of LCR22C. Thirteen normal individuals were tested with the new probes. Three carried inversions between LCR22C and LCR22D, and 10 carried them between LCR22B and LCR22D (Figures 1B and 1C and Table S1).

Now, 22q11.2DS can be added to the growing list of genomic disorders triggered by the presence of inversions. For the 5q35 deletion in Sotos syndrome (MIM: 117550), the 8p23 rearrangements, and the 17q21.31 deletions and duplications, all parents-of-origin are carriers of an inversion polymorphism between flanking LCRs. For these syndromes, the parental inversion is thought to be an obligate prerequisite that triggers these rearrangements. On the other hand, in Williams-Beuren syndrome and the Prader-Willi and Angelman syndromes, respectively, 30% and 66% of parents-of-origin carry an inversion polymorphism. The inversion is considered a susceptibility factor. However, whereas inversions in these genomic disorders are larger or of similar size to the region of deletion or duplication, we demonstrate here that 22q11.2 inversions are nested in or even flanking the larger rearrangements for which they cause a predisposition. Perhaps also all parents-of-origin of the Williams-Beuren syndrome and the Prader-Willi and Angelman syndromes carry nested or flanking small inversion polymorphisms that remain undetected. In the Williams-Beuren syndrome region, three LCRs are present. One third of Williams-Beuren syndrome parents-of-origin have a 1.79 Mb to 2.56 Mb proximal-distal inversion that predisposes them to having a child with the most-common 1.55 Mb proximal-medial LCR deletions. Interestingly, two additional small inversion polymorphisms have been reported between the medial and distal LCRs. These are hypothesized to increase susceptibility to nested deletions in the distal LCRs. The potential existence of these inversions in parents-of-origin has not yet been investigated.

Two mechanisms have been invoked to explain how inversions trigger NAHR. An inversion can lead to unstable asynaptic bubbles during meiosis; these bubbles are prone to breaks and secondary rearrangements. Alternatively, the inversion could re-orient the paralogous subunits from being in opposing to be in identical direction. NAHR of subunits in opposing orientation would result in an inversion, whereas NAHR of subunits in an identical orientation would generate deletions or duplications. In addition, the inversion could cause paralogous sequences to be more equally spread among LCRs, resulting in longer stretches of similarity and hence a higher chance for nonallelic homologous pairing and recombination. Thus, the risk of NAHR might increase between different LCR22s depending on the exact position of the inversion breakpoints within the LCRs. Because all but one of the parents-of-origin carry an inversion, we speculate that the inversion polymorphisms might be a necessary prerequisite to any of the known 22q11.2 rearrangements. In the one parent-of-origin without an inversion of LCR22B–D, we speculate that there might be another structural variation, proximal or distal to the LCR22B–D interval.

Surprisingly, no individuals homozygous or hemizygous for the inversion were observed. If the 29 diploid population samples and non-transmitting parents were in Hardy-Weinberg equilibrium, at least one individual should be homozygous for the inversion (chi-square test: $p = 0.1198$). Additionally, in eight individuals for whom both parents are inversion carriers, none carried the inversion ($p = 0.0039$). One possible explanation could be that the inversion breakpoints alter or disrupt the function or regulation of one or multiple essential genes. LCR22s contain four known genes, USP18 (Ubiquitin Specific Peptidase 18, MIM: 607057), GGT (Gamma Glutamyltranspeptidase, MIM: 612346), GGTS (Gamma Glutamyltransferase 5, MIM: 137168) and BCR (Breakpoint Cluster Region, MIM: 151410). Each has been (partially) duplicated during primate evolution and transposed on 22q11.2 to create a mosaic of largely pseudogenes. Additionally, inversion polymorphisms are known to affect topological associations between functional genes and their regulating elements, leading to more-distant effects. For several other genomic disorders, it has been demonstrated that genes within LCRs can modulate

---

The American Journal of Human Genetics 101, 616–622, October 5, 2017 619
neurodevelopmental phenotypes.49,50 In analogy, a similar phenomenon might take place at 22q11. Negative selection for inversion homozygotes during embryogenesis might explain the observed difference between the incidence of the 22q11DS in fetuses in low-risk women without evidence of congenital heart disease or palatal anomalies on ultrasound (~1/10000)3,51 and the population incidence of 1/3000–1/6000 live births. On the other hand, negative selection of inversion homozygotes is in contradiction with the high frequency of heterozygotes. Another explanation would be that the population sampled is too small. Larger population studies and better confirmation of the inversion polymorphisms are required to confirm this observation.

Supplemental Data

Supplemental Data contain three figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2017.09.002.

Consortia


Acknowledgments

W.D. is a fellow at Ku Leuven and is supported by IWT (131625). This work was made possible by grants from the KUL PFV/10/016 SymbioSys to J.R.V. and GOA/12/015 to J.R.V. and K.D. to the FWO to J.V. (G.0E1117N), the National Institute of Mental Health (5U01MH101723-02), and the Belgian Science Policy Office Inter-university Attraction Poles (BELSPO-IAP) program through the project IAP P7/43-BeMGI. I.A.P.J.'s lab was funded by the Marató de TV3 (20153230), the Spanish Ministry of Economy and Competitiveness (PI1302481), and the Catalan Department of Economy and Knowledge (2014SGR1468 and the ICREA Acadèmia award). J.V.K. was funded by the Jerome Lejeune Foundation (project #1665). The molar cell lines CHM1 and CHM13 were kindly provided by E. Eichler. D.M.M.-M has given lectures for Natera on 22q11.2 deletion syndrome. L.A.P.J. is a scientific advisor of qGenomics Laboratory, S.L.

Received: February 22, 2017
Accepted: August 16, 2017
Published: September 28, 2017

Web Resources

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org

References


620 The American Journal of Human Genetics 101, 616–622, October 5, 2017


