

RESEARCH ARTICLE

# Mechanisms for Complex Chromosomal Insertions

Shen Gu<sup>1</sup>, Przemyslaw Szafranski<sup>1</sup>, Zeynep Coban Akdemir<sup>1</sup>, Bo Yuan<sup>1</sup>, Mitchell L. Cooper<sup>1</sup>, Maria A. Magriñá<sup>2</sup>, Carlos A. Bacino<sup>1,3</sup>, Seema R. Lalani<sup>1,3</sup>, Amy M. Breman<sup>1</sup>, Janice L. Smith<sup>1</sup>, Ankita Patel<sup>1</sup>, Rodger H. Song<sup>1</sup>, Weimin Bi<sup>1</sup>, Sau Wai Cheung<sup>1</sup>, Claudia M. B. Carvalho<sup>1</sup>, Paweł Stankiewicz<sup>1\*</sup>, James R. Lupski<sup>1,3,4,5\*</sup>

**1** Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, Texas, United States of America, **2** Medical Specialties Unit From City Hall São José dos Campos, São Paulo, Brazil, **3** Department of Pediatrics, Baylor College of Medicine, Houston, Texas, United States of America, **4** Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, United States of America, **5** Texas Children's Hospital, Houston, Texas, United States of America

\* [jlupski@bcm.edu](mailto:jlupski@bcm.edu) (JRL); [pawels@bcm.edu](mailto:pawels@bcm.edu) (PS)



CrossMark  
click for updates

 OPEN ACCESS

**Citation:** Gu S, Szafranski P, Akdemir ZC, Yuan B, Cooper ML, Magriñá MA, et al. (2016) Mechanisms for Complex Chromosomal Insertions. *PLoS Genet* 12(11): e1006446. doi:10.1371/journal.pgen.1006446

**Editor:** Nancy B. Spinner, University of Pennsylvania, UNITED STATES

**Received:** June 29, 2016

**Accepted:** October 24, 2016

**Published:** November 23, 2016

**Copyright:** © 2016 Gu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data are included in the paper, including sequences and array findings. Relevant CNV array data has been deposited in GEO (Series No. GSE89306)

**Funding:** This work was supported in part by grants from the US National Institute of Neurological Disorders and Stroke (R01NS058529), National Institute of General Medical Sciences (R01GM106373) and National Human Genome Research Institute (NHGRI)/ National Heart Lung and Blood Institute (NHLBI) (U54 HG006542) to the Baylor-Hopkins Center for Mendelian Genomics. This work was also

## Abstract

Chromosomal insertions are genomic rearrangements with a chromosome segment inserted into a non-homologous chromosome or a non-adjacent locus on the same chromosome or the other homologue, constituting ~2% of nonrecurrent copy-number gains. Little is known about the molecular mechanisms of their formation. We identified 16 individuals with complex insertions among 56,000 individuals tested at Baylor Genetics using clinical array comparative genomic hybridization (aCGH) and fluorescence *in situ* hybridization (FISH). Custom high-density aCGH was performed on 10 individuals with available DNA, and breakpoint junctions were fine-mapped at nucleotide resolution by long-range PCR and DNA sequencing in 6 individuals to glean insights into potential mechanisms of formation. We observed microhomologies and templated insertions at the breakpoint junctions, resembling the breakpoint junction signatures found in complex genomic rearrangements generated by replication-based mechanism(s) with iterative template switches. In addition, we analyzed 5 families with apparently balanced insertion in one parent detected by FISH analysis and found that 3 parents had additional small copy-number variants (CNVs) at one or both sides of the inserting fragments as well as at the inserted sites. We propose that replicative repair can result in interchromosomal complex insertions generated through chromothripsis-like chromoanasythesis involving two or three chromosomes, and cause a significant fraction of apparently balanced insertions harboring small flanking CNVs.

## Author Summary

By traditional cytogenetic techniques, the incidence of microscopically visible chromosomal insertions was estimated to be 1 in 80,000 live births. More recently, by aCGH in conjunction with FISH confirmation of the aCGH findings, insertion events were demonstrated to occur much more frequently (1 in ~500 individuals tested). Although frequently

supported by BCM IDDRC Grant Number 5P30HD024064-23 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** JRL has stock ownership in 23andMe, is a paid consultant for Regeneron Pharmaceuticals, has stock options in LaserGen, Inc., is on the Scientific Advisory Board of Baylor Genetics, and is a co-inventor on multiple United States and European patents related to molecular diagnostics for inherited neuropathies, eye diseases and bacterial genomic fingerprinting. The Department of Molecular and Human Genetics at Baylor College of Medicine derives revenue from the chromosomal microarray analysis and clinical exome sequencing offered in the Baylor Genetics (<https://www.bcm.edu/research/medical-genetics-labs/>).

detected, little is known about the molecular mechanisms of their formation. In this study, we identified 16 individuals with complex chromosomal insertions among 56,000 individuals tested at Baylor Genetics using clinical microarray analysis (CMA) and FISH. Custom high-density aCGH was performed on 10 individuals with available DNA, and breakpoint junctions were fine-mapped at nucleotide resolution by long-range PCR and DNA sequencing in 6 individuals to glean insights into potential mechanisms of formation. In addition, we analyzed 5 families with apparently balanced insertion in one parent detected by FISH analysis and found that 3 parents had additional small copy-number variants (CNVs) at one or both sides of the inserting fragments as well as at the inserted sites. We propose that replicative repair can result in interchromosomal complex insertions generated through chromothripsis-like chromoanasythesis involving two or three chromosomes, and cause a significant fraction of apparently balanced insertions harboring small flanking CNVs.

## Introduction

Chromosomal insertion occurs when a segment of one chromosome is translocated and inserted into an interstitial region of another non-homologous chromosome (interchromosomal insertion), or into a different region of the same chromosome (intrachromosomal insertion). Insertions are considered as complex chromosomal rearrangements (CCRs) since they require at least three chromosome breakage events. [1] Chromosomal insertions are also considered as complex genomic rearrangements (CGRs) as they consist of more than one simple rearrangement, and have two or more DNA breakpoint junctions. [2, 3] By cytogenetic techniques, the incidence of microscopically visible insertions was estimated to be 1 in 80,000 live births.[4] More recently, by array-comparative genomic hybridization (aCGH) in conjunction with fluorescence *in situ* hybridization (FISH) confirmation of the aCGH findings, insertion events were demonstrated to occur much more frequently, with estimated incidence of 1 in 500[1] or 1 in 563[5] individuals tested. Another study demonstrated that ~2.1% of apparently *de novo*, interstitial CNVs were actually consequences of imbalances resulted from parents with balanced insertions.[6] These data highlight the importance of identifying such parental genomic information for reproductive counseling and potential recurrence risk estimates.

Phenotypic consequences of insertions vary, depending on the size, gene content and orientation of the inserted fragment, in addition to possible disruption or dysregulation of a gene or topologically associating domain (TAD) at the inserted genomic locus. Complex insertions are defined as insertions generated by more than three DNA breakages and joining events.[1, 7] Usually, additional copy-number gain or loss is observed at the inserted site for these events complicating the interpretation of potential phenotypic consequences observed. Little is known regarding the molecular mechanisms for the formation of insertions; particularly with regards to the mechanism(s) of formation of complex insertions. Thus, we aimed to elucidate the potential underlying mechanisms generating complex insertions. Surprisingly, we observed complex insertions as part of apparent chromothripsis-like, chromoanasythesis events involving two or three chromosomes.

Chromothripsis was first described as a catastrophic phenomenon in cancer genomes and observed as highly complex somatic rearrangements, with a distinct pattern of frequent oscillations between neutral and deleted copy-number states and seemingly focused on one chromosome.[8] A similar apparent chromosome shattering mechanism has been observed as *de novo*

mutations in individuals with neurodevelopmental abnormalities, and this type of germline chromothripsis involves complex balanced rearrangement among several chromosomes.[9] These events may appear as balanced rearrangements by conventional metaphase chromosome analysis. Both somatic and germline chromothripsis were proposed to be caused by a similar chromosome shattering mechanism that undergoes repair through non-homologous end-joining (NHEJ).[10, 11] A third type of chromothripsis-like event, defined as chromoanagenesis, was observed as *de novo* constitutional CGRs involving region-focused copy-number changes including duplications and triplications. These chromoanagenesis events were proposed to be generated through replication-based mechanisms, such as fork stalling and template switching and/or microhomology-mediated break-induced replication (FoSTeS/MMBIR) with iterative template switching resulting in extensive complexity.[12] The molecular analysis and findings in complex insertions we report here mostly resembled the patterns observed in constitutional genomic chromoanagenesis events.

In this study, we identified 16 individuals with distinct complex insertions among 56,000 individuals tested at Baylor Genetics (BG) using clinical aCGH and FISH. We fine-mapped DNA breakpoint junctions in 6 complex insertions at nucleotide resolution, and three of them resembled chromoanagenesis events with multiple chromosomes involved. In addition, we analyzed 5 families with unbalanced insertions detected in probands and inherited from parents with apparently balanced insertion detected by FISH analysis. We found that 3 parents had additional small CNVs at one or both sides of the inserting fragments as well as at the inserted sites likely generated during formation of the structural variant. We propose that these events are due to DNA replicative repair errors generated by replication-based mechanism(s) using iterative template switching.[3]

## Results

### Identify complex chromosomal insertions

Previously, we demonstrated that by performing confirmatory FISH of the copy-number gains identified in clinical chromosome microarray analysis (CMA) testing, some duplications were shown not to be represented by tandem duplication events, but were rather translocated and inserted at another locus in the genome.[1] This approach allowed the discovery of 40 individuals with insertions among the 18,000 individuals tested in the CMA laboratory at BG from July 2005 to January 2009. Among these 40 individuals, 8 were found to carry complex insertions (S1 Table, individuals Cplex1–8).[1] In this study, we expanded the cohort to 56,000 individuals tested from July 2005 to November 2014, and identified an additional 76 individuals with chromosome insertions (out of the subsequent 38,000 individuals tested), therefore resulting in the incidence of insertions being consistently about 1 in 500. This incidence is likely underestimated given that some of the insertions are too small to be verified by FISH. Among these latter 76 individuals, we identified another 8 individuals with complex insertions (S1 Table, individuals Cplex9–16).

Among the 16 individuals with complex insertions, 2 are intrachromosomal insertions (Cplex1 and Cplex2), and the remaining 14 are interchromosomal insertions (Cplex3–16) (S1 Table). Cplex2 was previously demonstrated in detail with all proposed breakpoint junctions mapped (BAB3105 from Ref. 12) in the paper that first defined the chromoanagenesis phenomenon and thus was excluded from the current study. For the remaining 15 individuals, we were able to obtain genomic DNA from 10 individuals (Cplex1, 3, 4, 5, 6, 7, 9, 11, 12, and 16) and repeated the CMA testing using the latest version (Baylor CMA version 10.2 oligo).[13] To map the breakpoint junctions to nucleotide resolution, we further designed high-density custom aCGH specifically targeting the inserted fragment and the potential inserting loci in 8

individuals based on the CMA results (Cplex3, 4, 5, 6, 7, 9, 11 and 12). By long-range PCR with Sanger sequencing, we were able to map (or partially map) the breakpoint junctions to nucleotide resolution in 6 individuals (Cplex4, 5, 6, 9, 11 and 12). For the remaining 4 individuals, we were unable to map breakpoint junctions, probably due to the limitations of the techniques applied in this study and potential further complexity at those breakpoints. Parental samples were not available for these 16 individuals.

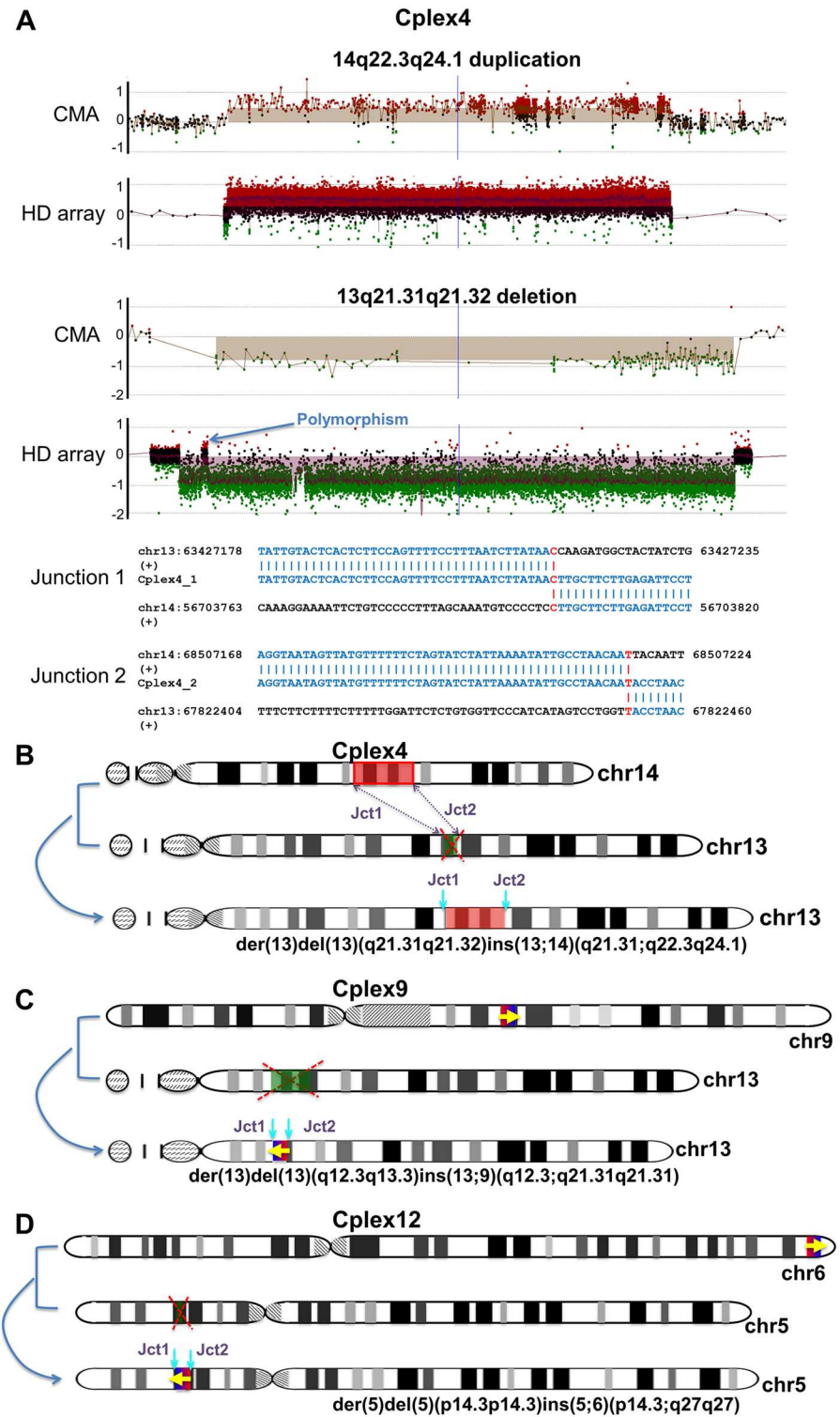
### Basic complex insertions

Among the 6 individuals with breakpoint junctions mapped, three individuals showed basic complex insertions, with a duplicated fragment translocated and inserted into another genomic locus with a deletion at the inserting position (Fig 1, S1 Table). Cplex4 demonstrated an ~11.8 Mb duplication on chromosome 14 (14q22.3q24.1) and an ~4.4 Mb deletion on chromosome 13 (13q21.31q21.32) revealed by array results (Fig 1A). FISH analysis and breakpoint junctions mapping demonstrated that the third copy of the duplicated segment on chr14 (chromosome 14) was inserted into chr13 at the position of the deletion (Fig 1B, FISH images previously published in a case report). [14] Similarly, in individual Cplex9, array results revealed an ~2.2 Mb duplication on chr9 at band 9q21.31 and an ~8.3 Mb deletion on chr13 at bands 13q12.3q13.3, while in Cplex12, array results revealed an ~0.8 Mb duplication on chr6 at band 6q27 and an ~0.5 Mb deletion on chr5 at band 5p14.3. Both FISH and breakpoint junctions mapping demonstrated that the duplicated fragment was inserted into the locus at which the deletion was observed in these latter two individuals Cplex9 and Cplex12 (Fig 1C and 1D and S1, S2A and S3A Figs). Note that in individual Cplex4, the inserted fragment was in the same orientation as the reference genome, however, in both Cplex9 and Cplex12, the inserted fragments were inverted when insertionally translocated (Fig 1B–1D). The CGRs in all three individuals were proposed to be generated through two breakpoint junctions, with 1 bp microhomology observed at both junctions in Cplex4 (Fig 1A, Table 1), 2 bp and 3 bp microhomologies observed at the junctions in Cplex9 (S2B Fig, Table 1), and 2 bp and 3 bp small insertions at junctions in Cplex12 (S3B Fig, Table 1).

### Chromothripsis-like, chromoanasythesis insertions

In contrast to the three individuals described above, Cplex5, Cplex6, and Cplex11 showed multiple CNVs in addition to the insertions and were generated through multiple breakpoint junctions (Table 1). Cplex5 exhibited 4 CNVs from the array results on both chr6 and chrX: an ~1.3 Mb duplication at 6q21, an ~0.4 Mb deletion at 6q24.2, an ~8.6 Mb deletion at 6q25.1q25.3 (resulting in an overall duplication-normal-deletion-normal-deletion CGR pattern on chr6), and an ~1.5 Mb duplication at Xq28 (Fig 2A). FISH analysis revealed that the duplicated fragment of Xq28 was inserted and translocated to chr6, potentially at the deleted locus of 6q24.2 (S1 Table). Breakpoint junction mapping confirmed the findings observed by FISH, and the 4 mapped junctions enabled developing a parsimonious model accounting for all available data potentially explaining the rearrangement in this individual (Fig 2B). In brief, the duplicated fragment of Xq28 was inserted into 6q24.2, replacing the deleted region of 6p24.2 (Junction 1 and 2), while a duplicated fragment of 6p21 was inserted into 6q25.1, again replacing the other deleted region of 6q25.1q25.3 (Junction 3 and 4). Breakpoint junction sequencing revealed a 7 bp templated insertion (copied from nearby sequences) at Junction 1, 2 bp microhomology at Junctions 2 and 4, and 5 bp microhomology at Junction 3 (Table 1; S4 Fig).

In individual Cplex6, CMA showed an ~0.58 Mb duplication at 5p15.33, and an ~0.07 Mb duplication at Xq28. High-density aCGH revealed that the duplication on Xq28 actually



**Fig 1. Proposed mechanisms in individuals with basic complex insertions.** (A) Upper panel: CMA and high-density aCGH results of Cplex4. Lower panel: breakpoint junction sequences in Cplex4. Microhomology between distal and proximal sequences are highlighted in red. HD array, high-density aCGH; (+), sequences in the positive strand in the hg19 reference genome. (B) Chromosome ideograms of individual Cplex4 demonstrating the duplicated fragment (segment highlighted in red) on chr14 was inserted and translocated to chr13 with where the deletion of chr13 (segment highlighted in green) occurred. (C) Chromosome ideograms



of individual Cplex9 demonstrating the duplicated fragment (right facing arrow) on chr9 was inserted and translocated to chr13 with where the deletion of chr13 (segment highlighted in green) occurred. (D) Chromosome idiograms of individual Cplex12 demonstrating the duplicated fragment (right facing arrow) on chr6 was inserted and translocated to chr5 with where the deletion of chr5 (segment highlighted in green) occurred. Note that in Cplex9 and Cplex12, the inserted fragments were both inverted after the insertion in comparison to the reference genome (left facing arrows). Jct1, Junction 1; Jct2, Junction 2. Polymorphism is defined as the observation that similar CNVs have been documented in multiple healthy, clinically unaffected individuals according to the Database of Genomic Variants (DGV).

doi:10.1371/journal.pgen.1006446.g001

contained a small triplication (~6 kb) embedded in the duplication (S5A Fig). FISH analysis and breakpoint junction mapping demonstrated that the duplicated fragment of 5p15.33 was inserted in an inverted orientation to Xq28. In addition, the triplication was also embedded in the duplication in an inverted orientation (Fig 2C, S1 Table), revealing a duplication—inverted triplication—duplication; a CGR pattern analogous to that previously observed and designated DUP-TRP/INV-DUP.[15] The proximal side of the duplication at Xq28 was joined to the distal side of the duplication at 5p15.33 (Fig 2C, Junction 1), while the proximal side of the

**Table 1. Features of sequenced breakpoint junction of chromosomal insertions.**

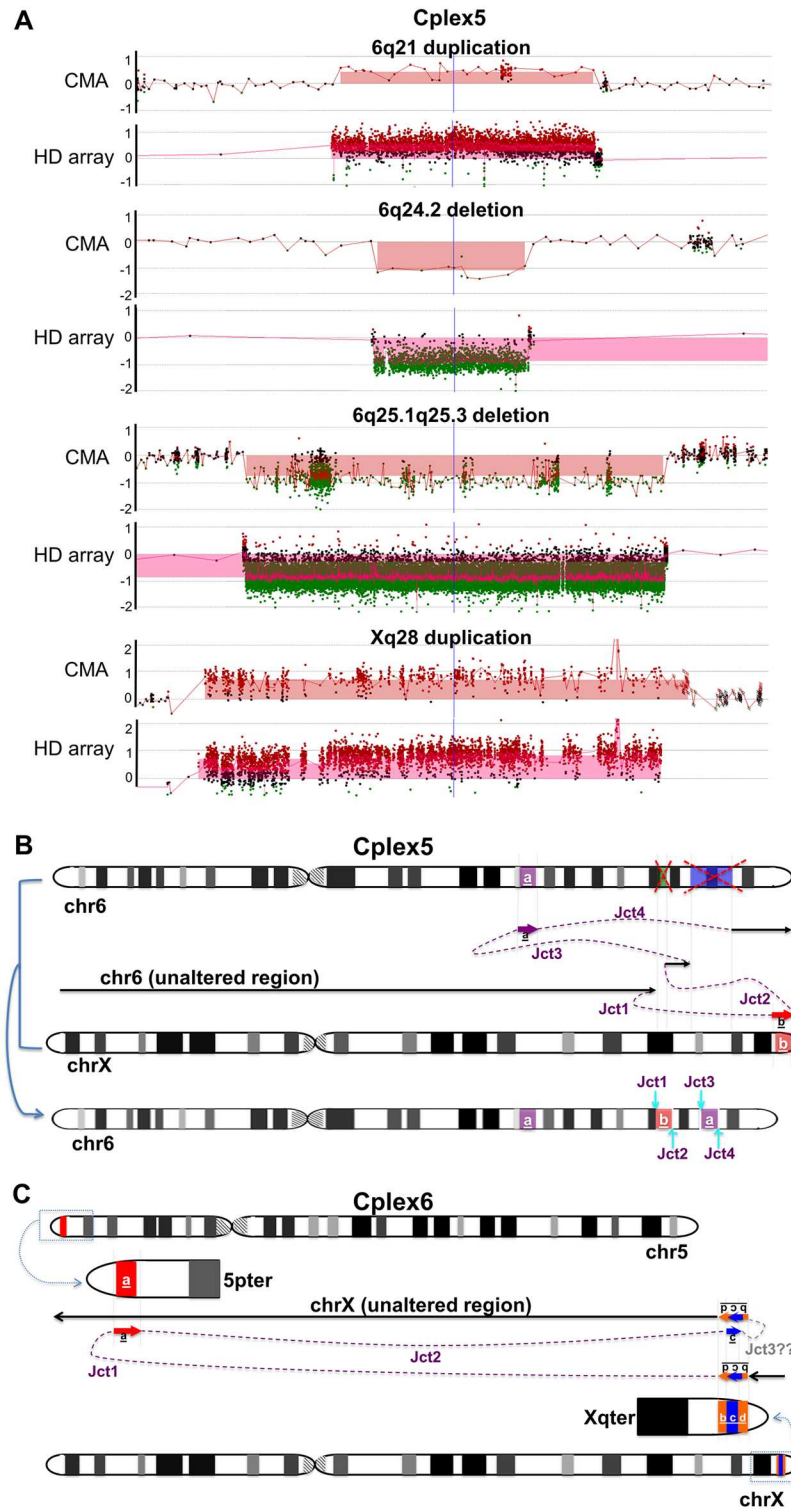
Sequenced junction	CNV description	Breakpoint junction feature	Distal Repetitive Element <sup>a</sup>	Proximal Repetitive Element
Cplex4_Jct1 Cplex4_Jct2	14p22.3q24.1 DUP, 13q21.31q21.32 DEL, ins(13;14)(q21.31;q22.3q24.1)	1bp MH <sup>b</sup> 1bp MH	MER44B MER113B	- L2a
Cplex9_Jct1 Cplex9_Jct2	9q21.31 DUP, 13q12.3q13.3 DEL, ins(13;9)(q12.3;q21.31q21.31)	2bp MH 3bp MH	- <i>AluSq</i>	- -
Cplex12_Jct1 Cplex12_Jct2	6p27 DUP, 5p14.3 DEL, ins(5;6)(p14.3;q27q27)	2bp insertion 3bp insertion	- MER5B	- L2a
Cplex5_Jct1 Cplex5_Jct2 Cplex5_Jct3 Cplex5_Jct4	6q21q25.3 DUP-NML-DEL-NML-DEL, Xq28 DUP, ins(6;X)(q25.3;q28q28)	7bp templated insertion 2bp MH 5bp MH 2bp MH	HERVH48-int L2a (in LCR) MER102a -	in LCR - - -
Cplex6_Jct1 Cplex6_Jct2	Xq28 DUP- TRP/INV-DUP, 5p15.33 DUP, ins(X;5)(q28;p15.3p15.3)	blunt ends 376bp templated insertion	<i>AluSx1</i> -	HERVH-int (in LCR) -
Cplex11_Jct1 Cplex11_Jct2 Cplex11_Jct3 Cplex11_Jct4 Cplex11_Jct6	13q33.2q34 DUP-NML-DUP-NML-DEL, Xq21.1 DUP-NML-DUP-TRP-DUP, ins(13;X)(q33.3;q21.1q21.1)	blunt ends 13,357bp templated insertion blunt ends blunt ends 2bp MH	- - L1MA2 L1MC4a L1MEc	L1PB1 - L1PB3 THE1D-int -
BAB1381_Jct1 BAB1381_Jct2 BAB1381_Jct3	Xq22.2 DEL-NML-DUP, 19q13.4 DUP, ins(19;X)(q22.2;q13.4q13.4)	3bp MH 15bp templated insertion 18bp MH	L2a <i>AluSg4</i> <i>AluSx1</i>	<i>AluSx3</i> L1PA3 <i>AluSz</i>
Mat3_Jct1 Mat3_Jct2 Mat3_Jct3	7p15.1 DEL, ins(9;7)(p24;p15.1p15.1)	815bp templated insertion 6bp templated insertion 6bp MH	- HERVK14-int MER21A	HERVK14-int MIR3 -
Mat12_DEL Jct	19q13.33 DUP, 19q13.31 DUP, ins(19)(p13;q13.33q13.33)	18bp insertion	<i>AluY</i>	<i>AluY</i> (in LCR)

Abbreviations: DEL, deletion; DUP, duplication; TRP, triplication; NML, normal; DUP-TRP/INV-DUP, inverted triplication embedded in duplication; LCR, low copy repeats.

<sup>a</sup>According to the “RepeatMasker” track in UCSC genome browser (GRCh37/hg19);

<sup>b</sup>MH, microhomology.

doi:10.1371/journal.pgen.1006446.t001



**Fig 2. Proposed mechanisms in individuals Cplex5 and Cplex6 with chromothripsis-like, chromoanasythesis insertions.** (A) CMA and high-density aCGH results of Cplex5. (B) Chromosome ideograms of individual Cplex5 demonstrating a duplication (segment highlighted in magenta) and two deletions (segments highlighted in green and blue, respectively) on chr6, plus a duplication (segment “b” highlighted in red) on chrX. Breakpoint junction mapping indicated that the duplicated fragment of Xq28 (red segment “b”) was inserted to 6q24.2, replacing the deleted region of 6p24.2 (green segment) through Junction

1 and 2, while the duplicated fragment of 6p21 (magenta segment “a”) was inserted to 6q25.1, again replacing the other deleted region of 6q25.1q25.3 (blue segment) through Junction 3 and 4. The overall result in Cplex5 was a rearranged chr6 with inserted fragment from chrX. (C) Chromosome ideograms of individual Cplex6 demonstrating a duplication (segment “a” highlighted in red) at 5pter and a triplication (segment “c” highlighted in blue) embedded in a duplication (segment “b+c+d” highlighted in orange) at Xqter. Junction 1 and 2 led to the insertion and joining of the duplicated region on chr5 (red segment “a”) to the triplication (blue segment “c”) on chrX, both in an inverted orientation in comparison to the reference genome. Note that Junction 3 was a hypothetical breakpoint junction to most parsimoniously explain the putative mechanism for this rearrangement. The overall result in individual Cplex6 was a rearranged chrX with inserted fragment from chr5. Jct1, Junction1; Jct2, Junction 2; Jct3, Junction 3. Dashed purple lines represent potential template switching paths during the generation of the CGRs. ‘?’ indicates hypothetical junction.

doi:10.1371/journal.pgen.1006446.g002

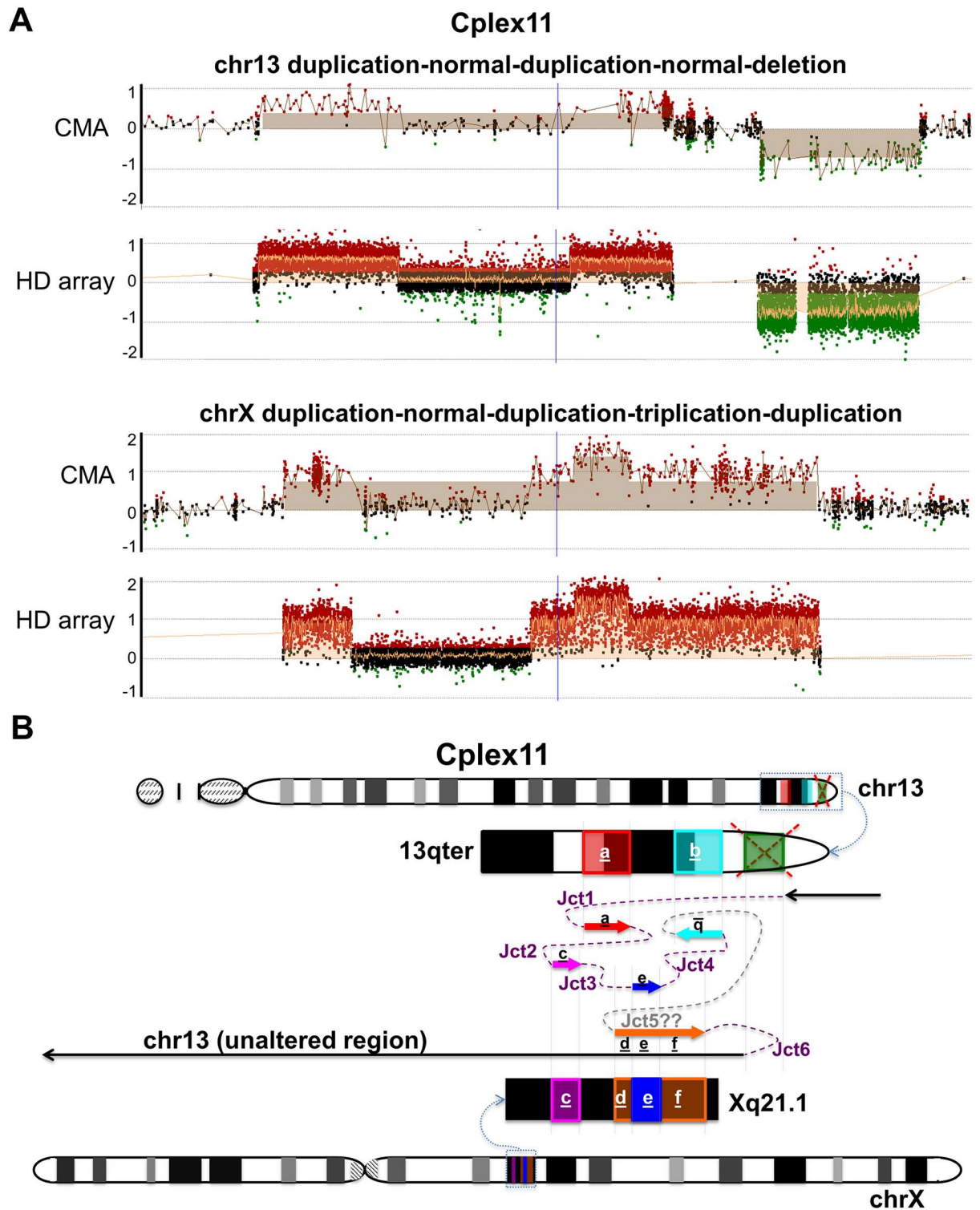
5p15.33 duplication was joined to the proximal side of the triplication embedded in the Xq28 duplication, leading to the triplication being inverted (Junction 2). We hypothesized a third junction connecting both distal sides of the triplication and the duplication at Xq28 should be present to generate the overall CGR in this individual; however, we were unable to uniquely position and map this breakpoint, possibly due to the presence of a low copy repeat (LCR) (S5A Fig). Sequences of Junction 1 in this individual showed blunt ends, while Junction 2 showed an insertion of 376 bp templated from at least three nearby genomic loci on both chr5 and chrX (Table 1, S5B Fig).

Individual Cplex11 exhibited the most complicated rearrangement in this study. Array results demonstrated a duplication-normal-duplication-normal-deletion pattern at 13q33.2 to 13q34 and a duplication-normal-duplication-triplication-duplication pattern at Xq21.1 (Fig 3A); FISH analysis showed that both of the two duplicated regions on chr13 were inserted into chrX (S6 Fig, S1 Table). Breakpoint mapping further demonstrated that the rearrangement between chr13 and chrX could be potentially generated through 6 junctions (Fig 3B). With the exception of the hypothetical Junction 5, we were able to map the remaining 5 junctions to nucleotide resolution. Based on the information of the five junctions mapped and the CNVs observed, we proposed the existence of Junction 5 to most parsimoniously explain the observed overall rearrangement in this individual (Fig 3B). Upon careful examination of the junctions, we observed that sequences of Junction 2 contained an 8,192 bp insertion from Xq13.2, followed by a 5,167 bp insertion from 4q13.1, leading to the discovery of the involvement of a third chromosome, chromosome 4, in this individual’s CGR (Table 1, S7 Fig). The remaining mapped junctions showed 2 bp microhomology (Junction 6) or blunt ends (Junction 1, 3 and 4).

## CNVs inherited from parents with apparently balanced insertions

Previously, we reported a child (BAB1379) with *PLP1* deletion that resulted from a maternal balanced insertion (BAB1381) of a segment on chrX containing the entire *PLP1* gene translocated and inserted into the telomeric region of the q arm of chr19 (Fig 4A).[16] The *PLP1* deletion breakpoint junction was mapped in the previous publication, showing an *Alu-Alu* mediated rearrangement (Junction 3 in S8 Fig, re-drawn in hg19). This junction was present in the mother (BAB1381) and her affected son with Pelizaeus-Merzbacher disease (BAB1379), but not in the unaffected son (BAB1380). To fine map other breakpoint junctions involving the insertion, we designed high-density aCGH targeting both the regions on chrX containing *PLP1*, and the potential insertion site at 19qter. Surprisingly, in the mother, we did not see complete copy-number neutral genomic intervals around the *PLP1* region as expected for her balanced insertion, but instead observed small CNVs that map at the exact loci of both ends of the deletion position in her affected son (Fig 4B). More specifically, an ~10 kb deletion at the proximal boundary, and an ~22 kb duplication at the distal boundary of the deletion position





**Fig 3. Proposed mechanisms in individual Cplex11 with chromothripsis-like, chromoanasythesis insertions.** (A) CMA and high-density aCGH results of Cplex11. (B) Chromosome ideograms of individual Cplex11 demonstrating two duplications (segments “a” and “b” highlighted in red and cyan) and a deletion (segment highlighted in green) on chr13, plus a duplication (segment “c” highlighted in magenta) and a triplication embedded in the other duplication (segment “e” highlighted in blue embedded in segment “d+e+f” in orange) on chrX. Junction 1 joined the distal side of the chr13 deletion (green) to the proximal side of the first duplication on chr13 (red “a”). Junction 2 joined the distal side of the first duplication on chr13 (red “a”) to the proximal side of the first duplication

on chrX (magenta “c”). Junction 3 joined the distal side of the first duplication on chrX (magenta “c”) to the proximal side of the triplication on chrX (blue “e”). Junction 4 joined the distal side of the triplication on chrX (blue “e”) to the distal side of the second duplication on chr13 (cyan “d”). Junction 5 (note this junction is hypothetical) joined the proximal side of the second duplication on chr13 (cyan “d”) to the proximal side of the second duplication on chrX (orange “d+e+f”). Lastly, Junction 6 joined the distal side of the second duplication on chrX (orange “d+e+f”) to the proximal side of the deletion on chr13 (green). The overall result in individual Cplex11 was a rearranged chr13 with multiple inserted fragments from chrX. Jct1 to Jct6, Junction 1 to Junction 6. Dashed purple lines represent potential template switching paths during the generation of the CGRs. ‘??’ indicates hypothetical junction.

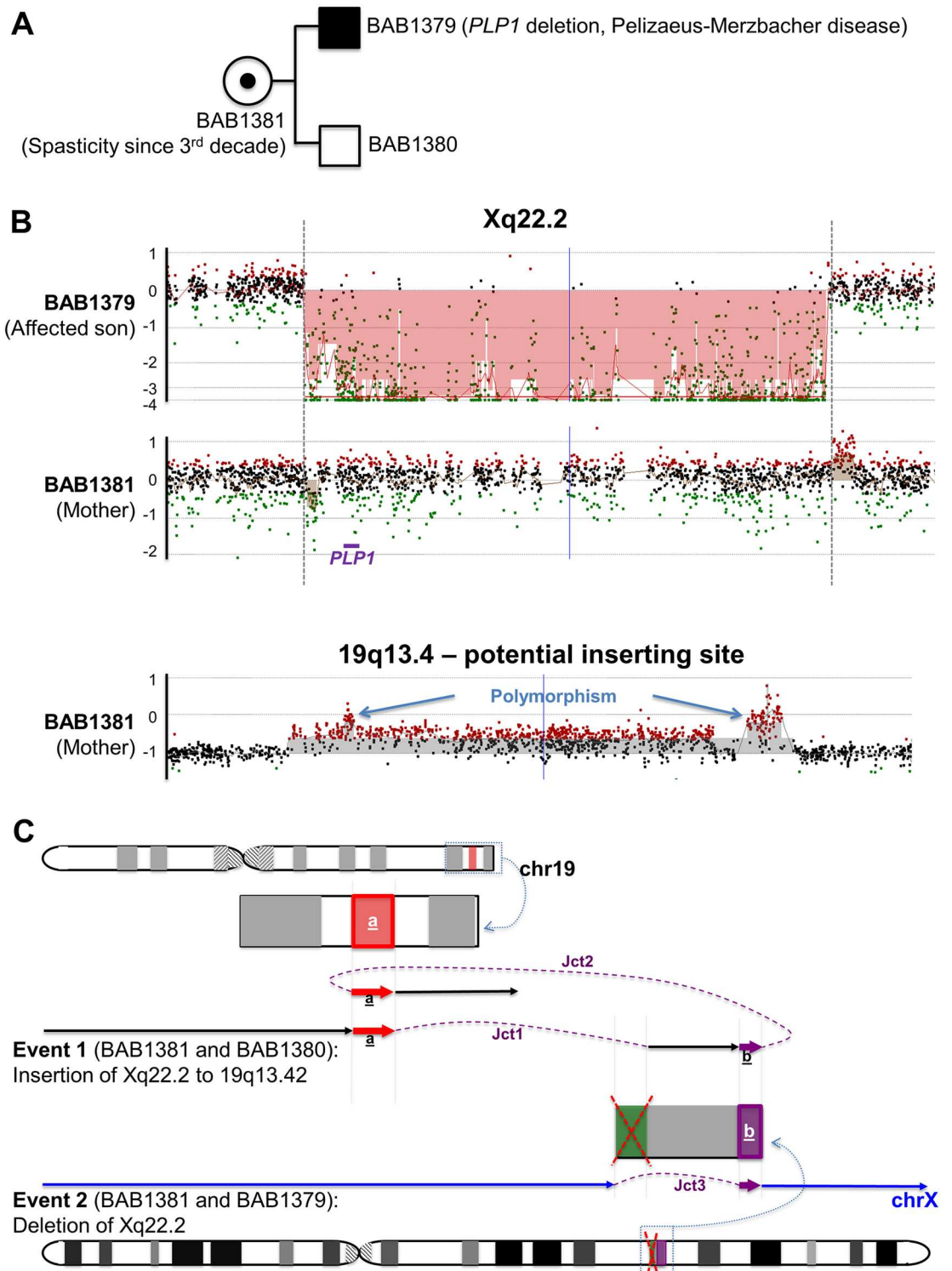
doi:10.1371/journal.pgen.1006446.g003

in her son (Fig 4B). In addition, an ~182 kb duplication was detected at 19q13.4, the potential inserting site, in the mother (Fig 4B). Further breakpoint junction mapping in the mother revealed that the distal side of the duplication on chr19 joined the distal side of the small deletion on chrX (Junction 1 in Fig 4C), while the proximal side of the chr19 duplication joined the distal side of the small duplication on chrX (Junction 2 in Fig 4C). The two small CNVs detected on chrX in the mother were actually due to unbalanced insertion from chrX to chr19, together with a duplication at the inserting site at 19q13.4. Sequences of the junctions showed 3 bp microhomology (Junction 1) and 15 bp templated insertion from nearby sequences at Junction 2 (Table 1, S8 Fig).

Observations in this family intrigued us to consider that the phenomenon may not be unique—CNVs inherited from parents with apparently balanced insertions may not be completely balanced at the molecular level. Given the small size of the potential CNVs, some may evade detection by clinical CMA. Therefore, we searched for similar families in the CMA database at BG, and found 12 families with a proband having a CNV inherited from a parent with apparently balanced insertion (named Family 1 to Family 12). We consented 4 families (Family 3, 4, 7 and 12) for further research studies, and discovered that in 2 families (Family 3 and Family 12), the apparently balanced insertions in the parents were not completely balanced, but actually had additional complexities revealed by molecular analyses.

In Family 3, Proband 3 (P3) showed a ~4.588 Mb deletion at 7p15.2p14.3 from array results; this deletion was further found by FISH analysis to be inherited from Mother 3 (Mat3) with apparently balanced insertion from chr7 into 9p24 (S9A Fig, S1 Table). Upon careful interpretation of high-density aCGH results, a small deletion (~4 kb) was observed in the mother at the exact boundary of the deletion in her child (Fig 5). We were able to fine map the identical deletion breakpoint junction present in both P3 and Mat3. Interestingly, an 815 bp insertion from 9p24 (chr9:5874574–5875388) was found at the chr7 junction sequences (Jct1)—the potential insertion locus observed from FISH in Mat3 (Fig 5). We further performed high-density aCGH in both Mat3 and P3 targeting the entire short arm of chr9. No promising CNVs were identified in either Mat3 or P3, however, three probes covering chr9:5874574–5875388 showed elevated ratio only in P3 but not Mat3 (S9B Fig). Based on this observation, we suspected an exchange of genetic material between chr7 and chr9 in the mother Mat3—the ~4.588 Mb fragment from 7p15.2p14.3 was inserted to chr9, replaced by a small fragment from chr9p24.1 (815 bp from chr9:5874574–5875388). Note that the large fragment of 7p15.2p14.3 broke and re-joined during the inserting process based on the observation of mapped breakpoint junction 2 (Jct2), and additional junction(s) should be present that connect the inserted fragments from chr7 to chr9 except for the mapped junction 3 (Jct3, S9C Fig).

Her child P3 inherited the deleted chr7 with the 815 bp insertion from chr9, together with an unaltered paternal chr9. We also suspected that the insertion site on chr9 was around chr9:5874574–5875388, and therefore performed walk-in PCRs and successfully pinpointed the insertion site on chr7 (S9C Fig). Another interesting observation is the presence of human endogenous retroviral elements (HERVs) at the boundaries of both Jct1 and Jct2, which are known to promote genome instability and induce CNV formation.[17]



**Fig 4. Proposed mechanisms in the *PLP1* deletion/insertion family with apparently balanced insertion in the mother.** (A) Pedigree of the family. (B) High-density aCGH results of BAB1379 and BAB1381. (C) Chromosome ideograms of the mother (BAB1381) demonstrating a duplication (segment “a” highlighted in red) on chr19 and a small deletion (segments highlighted in green) plus a small duplication (segment “b” highlighted in magenta) on chrX. The insertion event from chrX to chr19 were generated through Junction 1 and 2: Junction 1 joined the distal side of the duplication on chr19 (red “a”) to the distal side of the small deletion on chrX (green), while Junction 2 joined the distal side

of the small duplication on chrX (magenta “b”) to the proximal side of the duplication on chr19 (red “a”). The deletion event on chrX was generated through Junction 3 joining the proximal side of the small deletion on chrX (green) to the proximal side of the small duplication on chrX (magenta “b”). Her affected son (BAB1379) inherited the chrX with the deletion and an intact chr19, while her unaffected son (BAB1380) inherited an intact chrX and a chr19 with the insertion. Note that colored fragments are not in proportion to the actual CNVs’ sizes; i.e. not to scale. Jct1, Junction1; Jct2, Junction 2; Jct3, Junction 3. Dashed purple lines represent potential template switching paths during the generation of the CGRs. Polymorphism is defined as the observation that similar CNVs have been documented in multiple healthy, clinically unaffected individuals according DGV.

doi:10.1371/journal.pgen.1006446.g004

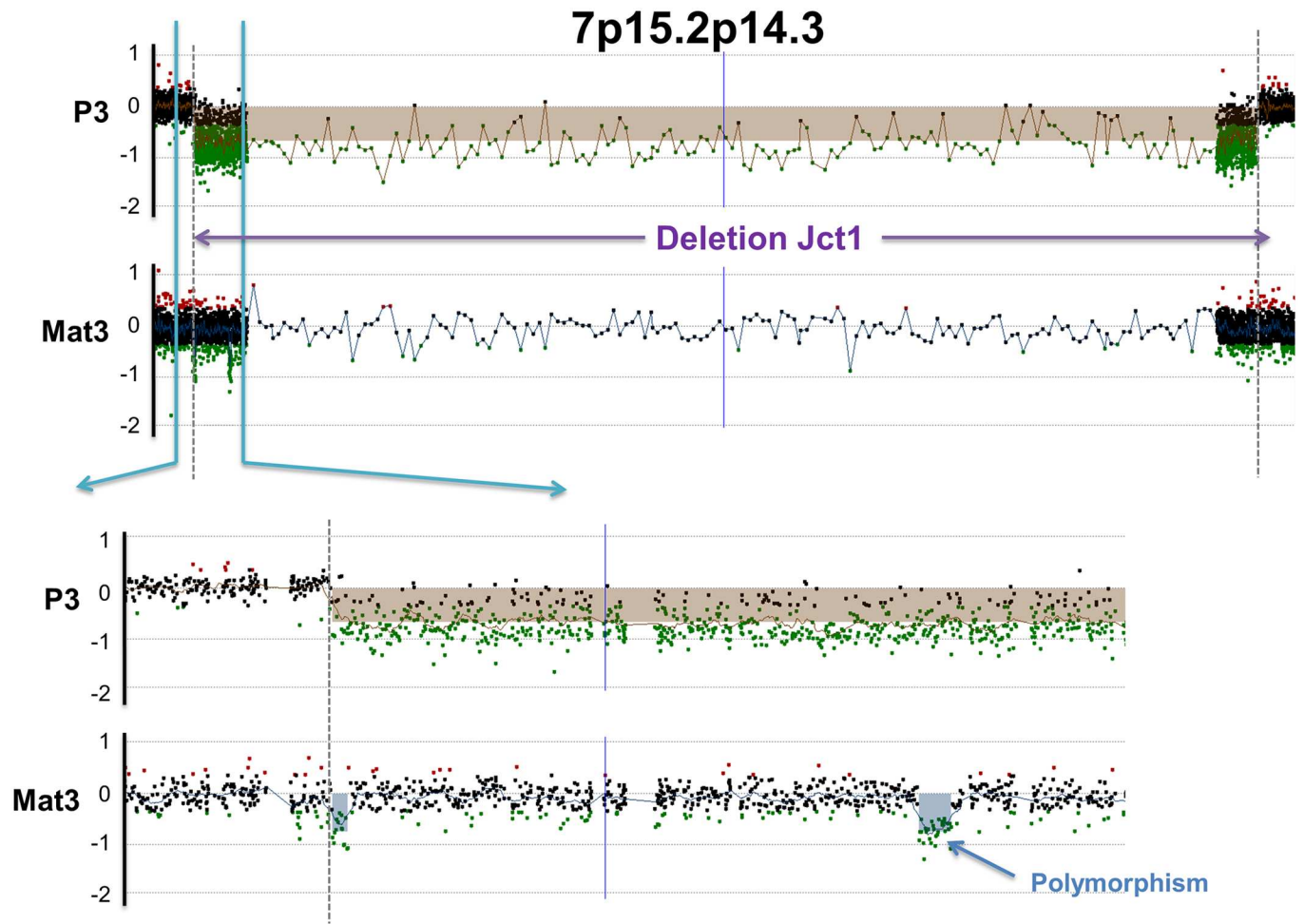
In Family 12, Mother 12 (Mat12) had two children with CNVs in the long arm of chr19: an ~3.5 Mb duplication at 19q13.33q13.41 in her daughter (P12\_dup), and a slightly smaller deletion (~3.352 Mb) at the same locus in her son (P12\_del) (S1 Table). FISH analysis demonstrated an apparently balanced insertion of a segment at 19q13.33 into the short arm of chr19 at 19p13 in Mat12; FISH analysis also demonstrated the same insertion in P12\_dup, indicating the duplication present in P12\_dup was likely a recombination product of intrachromosomal maternal insertion (S10 Fig). The reciprocal deletion present in P12\_del was also likely a recombination product (S11B Fig). High-density aCGH revealed that the apparently balanced insertion in Mat12 was not balanced—at both proximal and distal boundaries of the duplication/deletion in her two children, there were two small duplications of ~111 kb and ~77 kb in size, respectively. In addition, a small triplication (~33 kb) was found embedded in the duplication near the proximal side in P12\_dup (S11A Fig). These additional complexities were likely accompanying events with the insertion in Mat12 that was subsequently transmitted and inherited by her two children, similar to the rearrangement events involving the *PLP1* observed in BAB1381 mentioned above (S11C Fig, refer to S11 Fig for details of proposed rearrangements in Family 12).

## Discussion

Previously, we demonstrated that confirmatory and parental studies of CNVs by FISH analysis, especially the copy-number gains identified through CMA testing, led to the discovery of chromosomal insertions at a rate as high as 1 in ~500 individuals tested.[1] A similar high rate of 1 in ~563 individuals was independently reported.[5] Although it is now widely recognized that chromosomal insertions are not rare events,[1, 5, 6] the underlying mechanisms for their formation remain largely unknown. Most of the previous studies on insertions were based on relatively low resolution genome analysis by clinical arrays in combination with molecular cytogenetics, FISH, and chromosome analysis; only a few breakpoint junctions have been mapped to nucleotide resolution.[5, 18, 19]

In this study we focused on a subset of chromosomal insertions—complex insertions with additional copy-number gain or loss at the inserted site. High-density aCGH revealed additional complexities that evaded detection by CMA testing, including small triplications embedded in duplications (in individuals Cplex6 and P12\_dup) and small CNVs in individuals with apparently balanced insertions (in individuals BAB1381, Mat3, and Mat12). In addition, breakpoint junction mapping and careful examination of the junction sequences provided insights into the potential mechanisms for formation of these complex insertions, leading to the observation of distinct molecular characteristics of apparently basic complex insertions versus chromothripsis-like, chromoanasythesis insertions. Of note, only individuals with CNVs large enough to be detected by clinical microarray, and subsequently with copy-number gains large enough to be verified as insertions by FISH, were initially identified and molecularly studied. Therefore, copy-number neutral insertions, and insertions with smaller CNVs





**Jct1 (in both Mat3 and P3)**

```

chr7:25988845  AACCCAAACCTCCATTCCACTAGGGATGAGAGAGCAGTCTCTCCACCTAAAGGATCT  25988901
(+)
Mat3/P3_1     AACCCAAACCTCCATTCCACTAGGGAATGTGTGCTTAGCATGCACCTGTAATCCCAG
               |||
chr9:5874547  CGTCTCCATAAAAGGCCCAAGAGAAGCTGTGTGCTTAGCATGCACCTGTAATCCCAG  5874603
(+)
.....
[815bp insertion at the breakpoint junction from chr9:5874574-5875388]
.....
chr9:5875354  CTAAAGGAAGGGGCTCAGTGACTTGGTCGAGCCTGGACTGGCGGGGTTACTTCTGC  5875410
(+)
Mat3/P3      CTAAAGGAAGGGGCTCAGTGACTTGGTCGAGCCTGTGAGGCAGATTCAGATGAGGAA
               |||
chr7:30576589  ATAAAAGCAACTCTTGAGCCATTTCAAACAGATGTGAGGCAGATTCAGATGAGGAA  30576645
(+)
    
```

**Fig 5. High-density aCGH results and breakpoint junction in Family 3 with apparently balanced insertion in the mother.** Upper panel: High-density aCGH showed the deletion at 7p15.2p14.3 in P3, and a small deletion in Mat3 at the distal boundary of the large deletion in her child P3. Lower panel: deletion breakpoint junction sequences shared in P3 and Mat3. An 815 bp insertion from chr9:5874574–5875388 was observed at this junction. (+), sequences in the positive strand in the hg19 reference genome. Jct1, Junction1.

doi:10.1371/journal.pgen.1006446.g005



that escaped detection by clinical array or FISH validation, were selected against inclusion in this study.

We categorized individuals Cplex4, Cplex9, and Cplex12 as basic complex insertions (S12 Fig) based on the following observations: first, only one duplication was observed in these individuals, in contrast to the multiple copy-number gains observed in other individuals in this study; second, a deletion was always present at the inserting site; third, none of them were *de novo* events (S1 Table). Breakpoint junctional sequences in these individuals showed 1–3 bp microhomology or 2–3 bp small insertions; these features represent mutational signatures of breakpoint junctions observed in structural variants potentially generated by either non-homologous end-joining (NHEJ), or alternatively, microhomology-mediated end-joining (MMEJ) or FoSTeS/MMBIR with a single template switch.[3, 20–23]

In contrast to the individuals with basic complex insertions that were potentially generated by a number of different mechanisms, individuals Cplex5, Cplex6 and Cplex11 showed multiple CNVs including triplications. In addition, Cplex5 and Cplex11 are *de novo* events (S1 Table, inheritance mode in Cplex6 is unknown). Breakpoint junctions' sequences in these individuals showed longer homology (>4 bp) and hundreds to thousands of base pairs of templated insertions. CNVs in these individuals resembled chromoanasythesis events[12], and their breakpoint junctions features are signature findings observed in structural variants generated through replicative repair based mechanism, e.g. FoSTeS/MMBIR with multiple iterative template switch events.[24, 25] Interestingly, one of the 16 individuals identified with complex insertions initially included in this study, Cplex2, was included and analyzed in detail in the paper that first defined the chromoanasythesis phenomenon (BAB3105 from Ref. 12). This further strengthens our proposal that complex insertions could be part of a chromoanasythesis event.

Currently, three similar yet distinct types of chromothripsis, or chromothripsis-like events have been described, together they were referred to as 'chromoanagenesis'. [26] In somatic changes in the cancer genomes, chromothripsis was shown to be a catastrophic, one-step event leading to a signature pattern of frequent oscillations between unaltered and deleted copy-number states.[8] In cancer chromothripsis, most CNVs observed from genomic sequence analyses are deletions, with much less duplications resolved, and usually involves one chromosome. In contrast to the frequent copy-number loss in cancer chromothripsis, germline chromothripsis observed in individuals with neurodevelopmental abnormalities was shown to be balanced rearrangements—although several chromosomes were shattered and rejoined, the overall complex rearrangement involved almost no copy-number changes (except for deleting or inserting short sequences at breakpoint junctions).[9, 10] A recent study on unbalanced interchromosomal translocations revealed two individuals with *de novo* chromothripsis translocations generated through at least 18 or 33 breakpoint junctions, respectively, and both individuals only carried two large deletions (from 800 kb to 6.6 Mb) but no copy-number gains. [27] Both somatic and constitutional chromothripsis were proposed to be generated by NHEJ, given that the vast majority of the breakpoint junctions in these events showed blunt ends, 1 or 2 bp microhomology, or small insertions.[10, 11, 28]

In contrast to the balanced germline chromothripsis involving shattering and rejoining of several chromosomes, another type of chromothripsis-like events, observed by high-density aCGH and mechanistically defined as chromoanasythesis, was shown to involve multiple copy-number changes, particularly multiple gains of copy-number including duplications and triplications.[12] Notably, chromoanasythesis was considered to be region-focused events. [12, 29, 30] In the original paper that defined the chromoanasythesis phenomenon, all 17 individuals studied showed CNVs on the same chromosome, more specifically, 15 out of 17 individuals showed CNVs confined within the distal half of the involved chromosome arms.[12] It

was proposed that co-occurrence of CNVs with substantial interchromosomal exchanges would result in a non-viable offspring.[10] Here, we demonstrated that chromoanasythesis could involve two or even three chromosomes, as we observed a templated insertion as long as 5,167 bp from a third chromosome in addition to the two chromosomes involved in the rearrangements in Cplex11 (S7 Fig).

We categorized individuals Cplex5, Cplex6, and Cplex11 as chromoanasythesis events (S12 Fig) based on the observations that included: i) multiple copy-number gains, including triplications, were detected, ii) longer microhomology (>4 bp) observed at breakpoint junctions and iii) long templated insertions from multiple genomic loci also present at breakpoint junctions. Similar to previously reported region-focused chromoanasythesis events, these features are likely found in CGRs generated by iterative template switching during replicative repair based mechanisms, e.g. FoSTeS/MMBIR.[3, 24, 25] Note that in individuals Cplex6 and Cplex11, some of their breakpoint junction sequences showed blunt ends (S5B and S7 Figs); it is not uncommon to observe that a portion of the junctions in CGRs potentially generated through replication based mechanisms can show blunt ends, small insertions or short microhomology (1 or 2 bp).[15, 31, 32] In studies conducted in the yeast model organism, FoSTeS/MMBIR has been demonstrated to occur in the absence of microhomology (with 0–6 bp homology at breakpoint junctions). [33] Although breakpoint junctions with short microhomology (1–3 bp) have been observed in rearrangements proposed to be potentially generated through NHEJ, MMEJ and MMBIR in the human genome, iterative template switches are unique to the mechanism of FoSTeS/MMBIR. Therefore, it is important to consider not only microhomology length, but also the occurrence of templated insertions at junctions, and other evidence for potential iterative template switch events, in addition to whether copy-number gains (especially triplications) are present, when postulating potential biological mechanisms responsible for the generation of CGRs.

Recent studies on DNA damage in micronuclei provided a potential further explanation for the chromothripsis and chromoanasythesis events. Micronuclei are common outcomes of cell division defects; they are structurally similar to intact nuclei, but contain only one or a few chromosomes or chromosomal segments.[34] They could undergo defective and asynchronous DNA replication, resulting in DNA damage and extensive chromosomal fragmentation including catastrophic processes like chromothripsis; most importantly, their damaged and rearranged DNA fragments could be integrated back into the genome.[32, 35] Rearrangements proposed to be generated through NHEJ or MMBIR have been observed in micronuclei DNA, and segments from a single chromosome were observed in the majority of the micronuclei—potentially explaining why most observed chromoanasythesis events are chromosome or chromosome region-focused. The rare chromoanasythesis events involving two or three chromosomes we observed in this study are potentially in accordance with the rare observation of micronuclei DNA from two chromosomes undergoing chromothripsis.[35]

In this study, we also discovered that some apparently balanced insertions are actually unbalanced insertions; small deletions and duplications could be generated accompanying the inserting process. From the mechanistic aspect, it is crucial to reveal these small CNVs—a completely balanced insertion could be attributed to mechanisms like NHEJ, however, the additional CNVs, especially the copy-number gains, are more parsimoniously explained by replicative repair based mechanisms. For example, in the family with *PLP1* insertion, the most parsimonious explanation for the small CNVs at both the inserting site 19q13.42 and missing proximal/additional distal segments accompanying the inserted fragment from Xq22.2 is FoSTeS/MMBIR. During the replication process, a stalled replication fork at chr19 invaded and annealed to chrX, and after replication of a genomic interval containing the entire *PLP1* gene on Xq22.2, the replication fork switched back to chr19q13.42, however, to a more

proximal locus, therefore leading to the small duplication on chr19 (Fig 4C). We consider the situation in Family 12 to be similar to the *PLP1* family, due to the two duplications on both boundaries of the inserting fragment potentially generated accompanying the chromosomal insertion in the mother Mat12 (S11 Fig). Whereas in Family 3, the situation may be different—unlike in the chromoanasythesis subjects and in BAB1381, whose CNVs are most parsimoniously explained by template switching during the replication process (copying material from the inserting chromosomes to the inserted loci, always one direction), there was an exchange of genomic segments between chr9 and chr7 in Mat3. In addition, no copy-number gain was observed in this family, and the insertions in Mat3 are mostly balanced except for the 4 kb deletion at chr7. We propose the bi-directional, mostly balanced insertions in Mat3 may result from multiple breakages and re-joining of both chr7 and chr9, therefore may be generated through NHEJ or MMEJ.[9, 27]

LCRs and repetitive elements are known to facilitate genomic rearrangements.[12, 29, 36, 37] Enrichment of breakpoint in these repetitive sequences has been observed in nonrecurrent and complex structural changes at multiple genomic loci. [29, 38, 39] In the current study, we observed involvement of LCRs at breakpoint junctions in individuals Cplex5 and Cplex6 (Table 1), and HERV elements at breakpoint junctions in Family 3 and Cplex6 (S9C Fig). In addition, we observed involvement of other repetitive sequence, e.g. SINEs (short interspersed nuclear elements) at junctions in Cplex9, Cplex6, BAB1381 and in Family 12, also LINEs (long interspersed nuclear elements) at junctions in Cplex4, Cplex12, Cplex5, Cplex11 and BAB1381 (Table 1). These repeat and repetitive sequences may stimulate genomic instability and potentially assist replicative repair catalyzed genomic rearrangements facilitating template switching and the generation of the nonrecurrent and complex insertion events.[3, 29, 40]

In summary, from studies of complex chromosomal insertions, we observed that chromoanasythesis could occur beyond a confined chromosomal region and involve two or three chromosomes. We observed microhomologies and templated insertions at the breakpoint junctions, resembling the breakpoint junction signatures found in CGRs generated through replication-based mechanism(s) and iterative template switches: FoSTeS/MMBIR.[3] We propose that DNA replicative repair mechanisms can potentially result in interchromosomal complex insertions, and cause a significant fraction of apparently balanced insertions; especially those harboring small flanking CNVs.

## Materials and Methods

### Subjects

Sixteen individuals with complex chromosome insertions were identified in the CMA laboratory at Baylor Genetics among the ~56,000 individuals tested from 2007 to 2014. This study was approved by the Institutional Review Board for Human Subject Research at Baylor College of Medicine (IRB H-25466). Informed consent was obtained prior to collecting identifiable DNA samples (BAB1379, BAB1380, BAB1381, P3, Mat3, P12\_del, P12\_dup and Mat12). The remaining DNA samples were de-identified for breakpoint and mechanistic studies (named Cplex1, Cplex2, Cplex3, etc).

### Clinical CMA and FISH

Custom designed BCM OLIGO V6.5, V7, V8, V9 or V10 oligonucleotide arrays were performed as previously described.[41, 42] Arrays were designed to specifically interrogate clinically significant regions with an average resolution of 30 kb between probes. Interphase and metaphase FISH were performed to confirm the CMA findings and tested using available parental samples.[1]

## High-density aCGH

To further characterize the CNVs identified by CMA and FISH involving complex insertions, we designed several 4X 180K oligonucleotide arrays with ~200 bp per probe spacing from Agilent Technologies (AMADID 073188, 073189, 076797, 079204, 071585, 024241 and 015482). Hybridization controls were gender matched (Individual NA10851 as male control and Individual NA15510 as female control). Scanned array images were processed using Agilent Feature Extraction software (version 10) and extracted files were analyzed using Agilent Genomic Workbench (version 7.0.4.0). Array designs and sequence alignment for breakpoint analysis were based on the February 2009 genome build (GRCh37/hg19 assembly).

## Breakpoint junction mapping

To further confirm the CNVs identified by high-density arrays and map the breakpoint junctions, primers flanking the predicted breakpoints were designed and long-range PCRs were conducted using TaKaRa LA Taq according to the manufacturer's protocol (TaKaRa Bio Company, Cat. No. RR002) as previously described.[29] PCR products were prepared for sequencing using ExoSAP-IT (Affymetrix, Cat. No. 78201) according to the manufacturer's protocol or gel extracted and purified with the Zymoclean Gel DNA Recovery kit (Zymo Research, Cat. No. D4001). Purified PCR products were then sequenced by Sanger di-deoxynucleotide sequencing (BCM Sequencing Core, Houston, TX, USA). To elucidate the insertion site in individual Mat3, the APAGene GOLD genomic walking kit was used according to the company's protocol (BIO S&T, Cat. No. BT901-RT). Generally, this kit enables isolation of unknown sequences which flank known sequences. Three rounds of nested PCR with degenerate random tagging primers provided by the kit were performed, and the end PCR products were cloned into a TA vector (pGEM-T Easy Vector Systems, Promega, Cat. No. A1360) and were further subjected to Sanger sequencing.

## Supporting Information

### **S1 Table. aCGH and FISH results of the individuals studied.**

(XLSX)

### **S1 Fig. FISH images demonstrating CNVs and interchromosomal insertions in individuals Cplex9 (A, B) and Cplex12 (C,D).**

(TIF)

**S2 Fig. Array results and breakpoint junction sequences of individual Cplex9.** (A) CMA and high-density aCGH results of Cplex9. (B) Breakpoint junction sequences in Cplex9. Microhomologies between distal and proximal sequences are highlighted in red. (+), sequences in the positive strand in the hg19 reference genome; (-), sequences in negative strand in the hg19 reference genome.

(TIF)

### **S3 Fig. Array results and breakpoint junction sequences of individual Cplex12.**

(TIF)

### **S4 Fig. Breakpoint junction sequences of individual Cplex5.**

(TIF)

### **S5 Fig. Array results and breakpoint junction sequences of individual Cplex6.**

(TIF)

**S6 Fig. FISH images demonstrating multiple CNVs and interchromosomal insertions in individual Cplex11.**

(TIF)

**S7 Fig. Breakpoint junction sequences of individual Cplex11.**

(TIF)

**S8 Fig. Breakpoint junction sequences of individuals BAB1379, BAB1380 and BAB1381.**

(TIF)

**S9 Fig. Array results and breakpoint junction sequences of individuals P3 and Mat3. (A)**

FISH images demonstrating the deletion of chr7 in individual P3 inherited from the mother Mat3 with apparently balanced insertion of chr7 to chr9. (B) High density aCGH targeting chr9 short arm demonstrated an elevation of three probes covering the 815bp insertion only observed in P3 but not in Mat3. (C) Upper panel: graph demonstrating the chromosomal insertion in Mat3. Note that the large fragment of 7p15.2p14.3 disconnected and re-joined during the inserting process (representing by blue (or "a") and red (or "b") blocks) based on the observation of mapped breakpoint junction 2 (Jct2). Note that the sizes of the blocks are not in exact proportion to the actual sizes of these genomic segments. We postulate additional junction(s) except for the mapped Jct3 that connect the inserted fragment from chr7 to chr9.

Lower panel: breakpoint junction sequences of Jct2 and Jct3 in individual Mat3.

(TIF)

**S10 Fig. FISH images demonstrating CNVs and intrachromosomal insertions in individuals P12\_dup and Mat12.**

(TIF)

**S11 Fig. High-density aCGH results and proposed rearrangements in Family 12. (A)**

Upper panel shows the duplication in the child P12\_dup, copy-number neutral flanking by two small duplications in the mother Mat12, and the deletion in the child P12\_del at 19q13.33q13.41. Middle panel demonstrated the enlarged images of the boundaries on both sides. Child P12\_dup carried a small triplication embedded in the large duplication (three copies of segments a, c and d, plus a fourth copy of segment b). Child P12\_del carried deletion encompassing segment b and c (copy-number neutral for segment a and d). For the mother Mat12, aCGH shows three copies of segment a, b and d, and copy-number neutral (two copies) of segment c, however, the duplicated segment b is likely due to adding up of a triplication and a deletion. Lower panel showed the sequences for the deletion breakpoint junction present in both P12\_del and Mat12. (B) In the mother Mat12, intrachromosomal insertion from chr19 q arm to chr19 p arm led to the reciprocal duplication and deletion of chr19 in her two children. In Mat12, accompanying to the insertion of segment b and c to the p arm, additional amplification of segment a, b and d simultaneously happened, leading to additional materials of one copy of segment a, c and d, plus two copies of segment b in the p arm, together with a deleted fragment containing segment b and c in the q arm. Therefore, the overall CNVs shown in aCGH for Mat12 is duplication for segments a, b and d. Homologous recombination between this rearranged chr19 and the other intact chr19 in Mat12 led to P12\_dup inheriting a chr19 with the abnormal p arm plus a normal q arm, and P12\_del inheriting a chr19 with the normal p arm plus an abnormal q arm. Note that segment a, b, c and d in this figure are only for demonstration of copy numbers and not drawn to scale thus not reflecting the CNV sizes observed from aCGH result. DUP, duplication, TRP, triplication, NML, normal, DEL, deletion.

(TIF)



**S12 Fig. Circos plots of complex insertions.** Red colored blocks represent copy-number gains, while green colored blocks represent copy-number losses. Red lines demonstrate the mapped breakpoint junctions.

(TIF)

## Author Contributions

**Conceived and designed the experiments:** SG CMBC WB SWC PSt JRL.

**Performed the experiments:** SG PSz CMBC.

**Analyzed the data:** SG PSz ZCA BY PSt JRL CMBC.

**Contributed reagents/materials/analysis tools:** MLC MAM CAB SRL AMB JLS AP RHS.

**Wrote the paper:** SG CMBC PSt JRL.

## References

1. Kang SH, Shaw C, Ou Z, Eng PA, Cooper ML, Pursley AN, et al. Insertional translocation detected using FISH confirmation of array-comparative genomic hybridization (aCGH) results. *Am J Med Genet A*. 2010; 152A(5):1111–26. doi: [10.1002/ajmg.a.33278](https://doi.org/10.1002/ajmg.a.33278) PMID: [20340098](https://pubmed.ncbi.nlm.nih.gov/20340098/)
2. Liu P, Carvalho CM, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev*. 2012; 22(3):211–20. doi: [10.1016/j.gde.2012.02.012](https://doi.org/10.1016/j.gde.2012.02.012) PMID: [22440479](https://pubmed.ncbi.nlm.nih.gov/22440479/)
3. Carvalho CM, Lupski JR. Mechanisms underlying structural variant formation in genomic disorders. *Nat Rev Genet*. 2016; 17(4):224–38. doi: [10.1038/nrg.2015.25](https://doi.org/10.1038/nrg.2015.25) PMID: [26924765](https://pubmed.ncbi.nlm.nih.gov/26924765/)
4. Van Hemel JO, Eussen HJ. Interchromosomal insertions. Identification of five cases and a review. *Hum Genet*. 2000; 107(5):415–32. PMID: [11140939](https://pubmed.ncbi.nlm.nih.gov/11140939/)
5. Neill NJ, Ballif BC, Lamb AN, Parikh S, Ravnán JB, Schultz RA, et al. Recurrence, submicroscopic complexity, and potential clinical relevance of copy gains detected by array CGH that are shown to be unbalanced insertions by FISH. *Genome Res*. 2011; 21(4):535–44. doi: [10.1101/gr.114579.110](https://doi.org/10.1101/gr.114579.110) PMID: [21383316](https://pubmed.ncbi.nlm.nih.gov/21383316/)
6. Nowakowska BA, de Leeuw N, Ruivenkamp CA, Sikkema-Raddatz B, Crolla JA, Thoelen R, et al. Parental insertional balanced translocations are an important cause of apparently de novo CNVs in patients with developmental anomalies. *Eur J Hum Genet*. 2012; 20(2):166–70. doi: [10.1038/ejhg.2011.157](https://doi.org/10.1038/ejhg.2011.157) PMID: [21915152](https://pubmed.ncbi.nlm.nih.gov/21915152/)
7. Zhang F, Carvalho CM, Lupski JR. Complex human chromosomal and genomic rearrangements. *Trends Genet*. 2009; 25(7):298–307. doi: [10.1016/j.tig.2009.05.005](https://doi.org/10.1016/j.tig.2009.05.005) PMID: [19560228](https://pubmed.ncbi.nlm.nih.gov/19560228/)
8. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*. 2011; 144(1):27–40. doi: [10.1016/j.cell.2010.11.055](https://doi.org/10.1016/j.cell.2010.11.055) PMID: [21215367](https://pubmed.ncbi.nlm.nih.gov/21215367/)
9. Kloosterman WP, Guryev V, van Roosmalen M, Duran KJ, de Bruijn E, Bakker SC, et al. Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum Mol Genet*. 2011; 20(10):1916–24. doi: [10.1093/hmg/ddr073](https://doi.org/10.1093/hmg/ddr073) PMID: [21349919](https://pubmed.ncbi.nlm.nih.gov/21349919/)
10. Chiang C, Jacobsen JC, Ernst C, Hanscom C, Heilbut A, Blumenthal I, et al. Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat Genet*. 2012; 44(4):390–7, S1. doi: [10.1038/ng.2202](https://doi.org/10.1038/ng.2202) PMID: [22388000](https://pubmed.ncbi.nlm.nih.gov/22388000/)
11. Kloosterman WP, Tavakoli-Yaraki M, van Roosmalen MJ, van Binsbergen E, Renkens I, Duran K, et al. Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and non-homologous repair mechanisms. *Cell Rep*. 2012; 1(6):648–55. doi: [10.1016/j.celrep.2012.05.009](https://doi.org/10.1016/j.celrep.2012.05.009) PMID: [22813740](https://pubmed.ncbi.nlm.nih.gov/22813740/)
12. Liu P, Erez A, Nagamani SC, Dhar SU, Kolodziejska KE, Dharmadhikari AV, et al. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell*. 2011; 146(6):889–903. doi: [10.1016/j.cell.2011.07.042](https://doi.org/10.1016/j.cell.2011.07.042) PMID: [21925314](https://pubmed.ncbi.nlm.nih.gov/21925314/)
13. Wiszniewska J, Bi W, Shaw C, Stankiewicz P, Kang SH, Pursley AN, et al. Combined array CGH plus SNP genome analyses in a single assay for optimized clinical testing. *Eur J Hum Genet*. 2014; 22(1):79–87. doi: [10.1038/ejhg.2013.77](https://doi.org/10.1038/ejhg.2013.77) PMID: [23695279](https://pubmed.ncbi.nlm.nih.gov/23695279/)

14. Ou Z, Martin DM, Bedoyan JK, Cooper ML, Chinault AC, Stankiewicz P, et al. Branchiootorenal syndrome and oculoauriculovertebral spectrum features associated with duplication of SIX1, SIX6, and OTX2 resulting from a complex chromosomal rearrangement. *Am J Med Genet A*. 2008; 146A(19):2480–9. doi: [10.1002/ajmg.a.32398](https://doi.org/10.1002/ajmg.a.32398) PMID: [18666230](https://pubmed.ncbi.nlm.nih.gov/18666230/)
15. Carvalho CM, Ramocki MB, Pehlivan D, Franco LM, Gonzaga-Jauregui C, Fang P, et al. Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nat Genet*. 2011; 43(11):1074–81. doi: [10.1038/ng.944](https://doi.org/10.1038/ng.944) PMID: [21964572](https://pubmed.ncbi.nlm.nih.gov/21964572/)
16. Inoue K, Osaka H, Thurston VC, Clarke JT, Yoneyama A, Rosenbarker L, et al. Genomic rearrangements resulting in PLP1 deletion occur by nonhomologous end joining and cause different dysmyelinating phenotypes in males and females. *Am J Hum Genet*. 2002; 71(4):838–53. doi: [10.1086/342728](https://doi.org/10.1086/342728) PMID: [12297985](https://pubmed.ncbi.nlm.nih.gov/12297985/)
17. Campbell IM, Gambin T, Dittwald P, Beck CR, Shuvarikov A, Hixson P, et al. Human endogenous retroviral elements promote genome instability via non-allelic homologous recombination. *BMC Biol*. 2014; 12:74. doi: [10.1186/s12915-014-0074-4](https://doi.org/10.1186/s12915-014-0074-4) PMID: [25246103](https://pubmed.ncbi.nlm.nih.gov/25246103/)
18. Baskin B, Stavropoulos DJ, Rebeiro PA, Orr J, Li M, Steele L, et al. Complex genomic rearrangements in the dystrophin gene due to replication-based mechanisms. *Mol Genet Genomic Med*. 2014; 2(6):539–47. doi: [10.1002/mgg3.108](https://doi.org/10.1002/mgg3.108) PMID: [25614876](https://pubmed.ncbi.nlm.nih.gov/25614876/)
19. Masliah-Planchon J, Dupont C, Vartzelis G, Trimouille A, Eymard-Pierre E, Gay-Bellile M, et al. Insertion of an extra copy of Xq22.2 into 1p36 results in functional duplication of the PLP1 gene in a girl with classical Pelizaeus-Merzbacher disease. *BMC Med Genet*. 2015; 16:77. doi: [10.1186/s12881-015-0226-6](https://doi.org/10.1186/s12881-015-0226-6) PMID: [26329556](https://pubmed.ncbi.nlm.nih.gov/26329556/)
20. Kent T, Chandramouly G, McDevitt SM, Ozdemir AY, Pomerantz RT. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat Struct Mol Biol*. 2015; 22(3):230–7. doi: [10.1038/nsmb.2961](https://doi.org/10.1038/nsmb.2961) PMID: [25643323](https://pubmed.ncbi.nlm.nih.gov/25643323/)
21. Lieber MR, Ma Y, Pannicke U, Schwarz K. Mechanism and regulation of human non-homologous DNA end-joining. *Nat Rev Mol Cell Biol*. 2003; 4(9):712–20. doi: [10.1038/nrm1202](https://doi.org/10.1038/nrm1202) PMID: [14506474](https://pubmed.ncbi.nlm.nih.gov/14506474/)
22. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet*. 2008; 24(11):529–38. doi: [10.1016/j.tig.2008.08.007](https://doi.org/10.1016/j.tig.2008.08.007) PMID: [18809224](https://pubmed.ncbi.nlm.nih.gov/18809224/)
23. Pannunzio NR, Li S, Watanabe G, Lieber MR. Non-homologous end joining often uses microhomology: implications for alternative end joining. *DNA Repair (Amst)*. 2014; 17:74–80.
24. Hastings PJ, Ira G, Lupski JR. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet*. 2009; 5(1):e1000327. doi: [10.1371/journal.pgen.1000327](https://doi.org/10.1371/journal.pgen.1000327) PMID: [19180184](https://pubmed.ncbi.nlm.nih.gov/19180184/)
25. Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*. 2007; 131(7):1235–47. doi: [10.1016/j.cell.2007.11.037](https://doi.org/10.1016/j.cell.2007.11.037) PMID: [18160035](https://pubmed.ncbi.nlm.nih.gov/18160035/)
26. Holland AJ, Cleveland DW. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat Med*. 2012; 18(11):1630–8. doi: [10.1038/nm.2988](https://doi.org/10.1038/nm.2988) PMID: [23135524](https://pubmed.ncbi.nlm.nih.gov/23135524/)
27. Weckselblatt B, Hermetz KE, Rudd MK. Unbalanced translocations arise from diverse mutational mechanisms including chromothripsis. *Genome Res*. 2015; 25(7):937–47. doi: [10.1101/gr.191247.115](https://doi.org/10.1101/gr.191247.115) PMID: [26070663](https://pubmed.ncbi.nlm.nih.gov/26070663/)
28. Kloosterman WP, Cuppen E. Chromothripsis in congenital disorders and cancer: similarities and differences. *Curr Opin Cell Biol*. 2013; 25(3):341–8. doi: [10.1016/j.ceb.2013.02.008](https://doi.org/10.1016/j.ceb.2013.02.008) PMID: [23478216](https://pubmed.ncbi.nlm.nih.gov/23478216/)
29. Gu S, Yuan B, Campbell IM, Beck CR, Carvalho CM, Nagamani SC, et al. Alu-mediated diverse and complex pathogenic copy-number variants within human chromosome 17 at p13.3. *Hum Mol Genet*. 2015; 24(14):4061–77. doi: [10.1093/hmg/ddv146](https://doi.org/10.1093/hmg/ddv146) PMID: [25908615](https://pubmed.ncbi.nlm.nih.gov/25908615/)
30. Plaisancie J, Kleinfinger P, Cances C, Bazin A, Julia S, Trost D, et al. Constitutional chromoanagenesis: description of a rare chromosomal event in a patient. *Eur J Med Genet*. 2014; 57(10):567–70. doi: [10.1016/j.ejmg.2014.07.004](https://doi.org/10.1016/j.ejmg.2014.07.004) PMID: [25128687](https://pubmed.ncbi.nlm.nih.gov/25128687/)
31. Beck CR, Carvalho CM, Banser L, Gambin T, Stubbolo D, Yuan B, et al. Complex genomic rearrangements at the PLP1 locus include triplication and quadruplication. *PLoS Genet*. 2015; 11(3):e1005050. doi: [10.1371/journal.pgen.1005050](https://doi.org/10.1371/journal.pgen.1005050) PMID: [25749076](https://pubmed.ncbi.nlm.nih.gov/25749076/)
32. Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, Pan Y, et al. DNA breaks and chromosome pulverization from errors in mitosis. *Nature*. 2012; 482(7383):53–8. doi: [10.1038/nature10802](https://doi.org/10.1038/nature10802) PMID: [22258507](https://pubmed.ncbi.nlm.nih.gov/22258507/)
33. Sakofsky CJ, Ayyar S, Deem AK, Chung WH, Ira G, Malkova A. Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication Leading to Complex Chromosomal Rearrangements. *Mol Cell*. 2015; 60(6):860–72. doi: [10.1016/j.molcel.2015.10.041](https://doi.org/10.1016/j.molcel.2015.10.041) PMID: [26669261](https://pubmed.ncbi.nlm.nih.gov/26669261/)

34. Zhang CZ, Leibowitz ML, Pellman D. Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. *Genes Dev.* 2013; 27(23):2513–30. doi: [10.1101/gad.229559.113](https://doi.org/10.1101/gad.229559.113) PMID: [24298051](https://pubmed.ncbi.nlm.nih.gov/24298051/)
35. Zhang CZ, Spektor A, Cornils H, Francis JM, Jackson EK, Liu S, et al. Chromothripsis from DNA damage in micronuclei. *Nature.* 2015; 522(7555):179–84. doi: [10.1038/nature14493](https://doi.org/10.1038/nature14493) PMID: [26017310](https://pubmed.ncbi.nlm.nih.gov/26017310/)
36. Stankiewicz P, Shaw CJ, Dapper JD, Wakui K, Shaffer LG, Withers M, et al. Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am J Hum Genet.* 2003; 72(5):1101–16. doi: [10.1086/374385](https://doi.org/10.1086/374385) PMID: [12649807](https://pubmed.ncbi.nlm.nih.gov/12649807/)
37. Vissers LE, Bhatt SS, Janssen IM, Xia Z, Lalani SR, Pfundt R, et al. Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. *Hum Mol Genet.* 2009; 18(19):3579–93. doi: [10.1093/hmg/ddp306](https://doi.org/10.1093/hmg/ddp306) PMID: [19578123](https://pubmed.ncbi.nlm.nih.gov/19578123/)
38. Carvalho CM, Zhang F, Liu P, Patel A, Sahoo T, Bacino CA, et al. Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching. *Hum Mol Genet.* 2009; 18(12):2188–203. doi: [10.1093/hmg/ddp151](https://doi.org/10.1093/hmg/ddp151) PMID: [19324899](https://pubmed.ncbi.nlm.nih.gov/19324899/)
39. Yuan B, Harel T, Gu S, Liu P, Burglen L, Chantot-Bastaraud S, et al. Nonrecurrent 17p11.2p12 Rearrangement Events that Result in Two Concomitant Genomic Disorders: The PMP22-RAI1 Contiguous Gene Duplication Syndrome. *Am J Hum Genet.* 2015; 97(5):691–707. doi: [10.1016/j.ajhg.2015.10.003](https://doi.org/10.1016/j.ajhg.2015.10.003) PMID: [26544804](https://pubmed.ncbi.nlm.nih.gov/26544804/)
40. Mayle R, Campbell IM, Beck CR, Yu Y, Wilson M, Shaw CA, et al. DNA REPAIR. Mus81 and converging forks limit the mutagenicity of replication fork breakage. *Science.* 2015; 349(6249):742–7. doi: [10.1126/science.aaa8391](https://doi.org/10.1126/science.aaa8391) PMID: [26273056](https://pubmed.ncbi.nlm.nih.gov/26273056/)
41. Boone PM, Bacino CA, Shaw CA, Eng PA, Hixson PM, Pursley AN, et al. Detection of clinically relevant exonic copy-number changes by array CGH. *Hum Mutat.* 2010; 31(12):1326–42. doi: [10.1002/humu.21360](https://doi.org/10.1002/humu.21360) PMID: [20848651](https://pubmed.ncbi.nlm.nih.gov/20848651/)
42. Ou Z, Kang SH, Shaw CA, Carmack CE, White LD, Patel A, et al. Bacterial artificial chromosome-emulation oligonucleotide arrays for targeted clinical array-comparative genomic hybridization analyses. *Genet Med.* 2008; 10(4):278–89. doi: [10.1097/GIM.0b013e31816b4420](https://doi.org/10.1097/GIM.0b013e31816b4420) PMID: [18414211](https://pubmed.ncbi.nlm.nih.gov/18414211/)