The American Journal of Human Genetics

Supplemental Data

Absence of Heterozygosity due to Template Switching during Replicative Rearrangements

Claudia M.B. Carvalho, Rolph Pfundt, Daniel A. King, Sarah J. Lindsay, Luciana W. Zuccherato, Merryn V.E. Macville, Pengfei Liu, Diana Johnson, Pawel Stankiewicz, Chester W. Brown, DDD Study, Chad A. Shaw, Matthew E. Hurles, Grzegorz Ira, P.J. Hastings, Han G. Brunner, and James R. Lupski

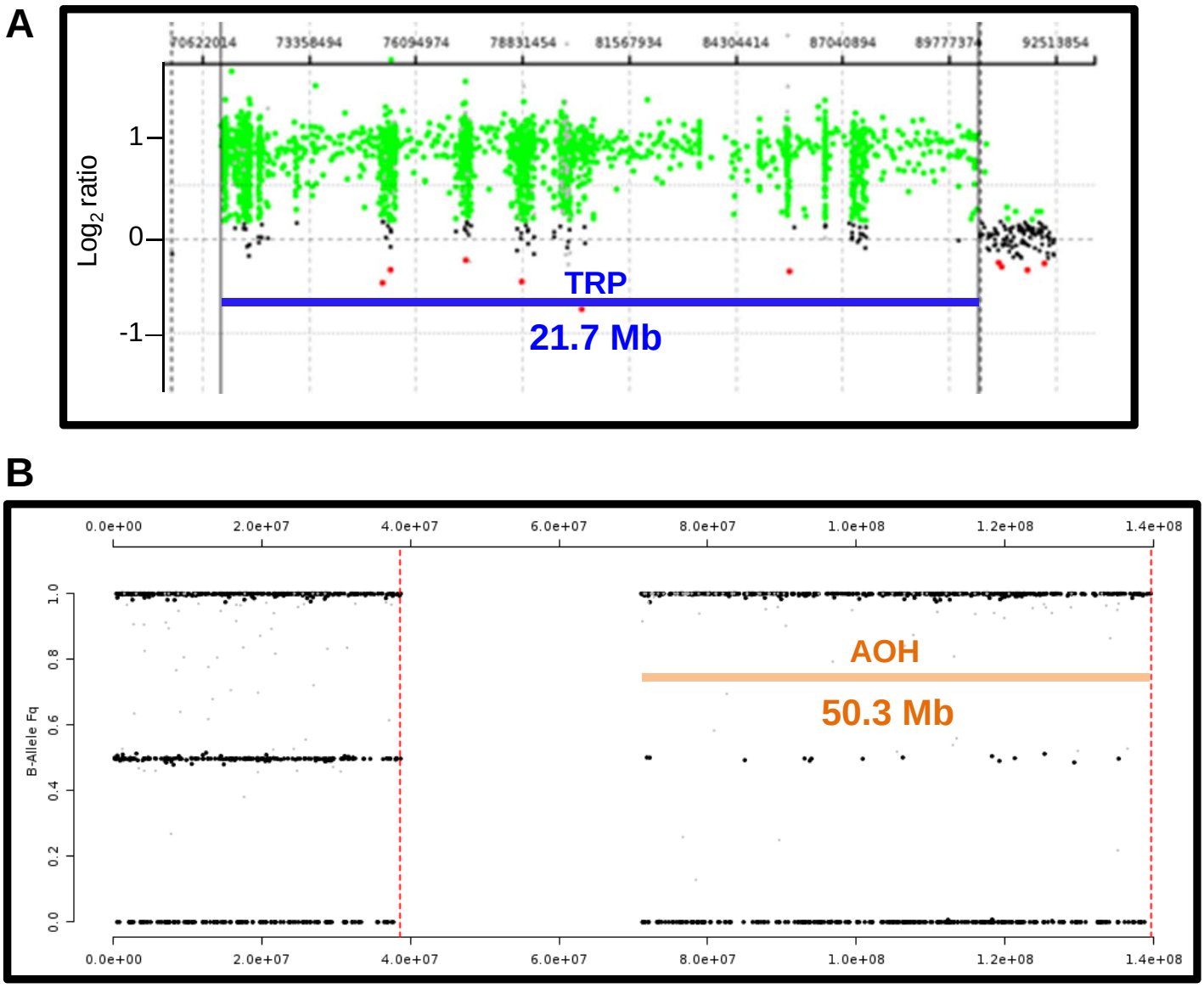

Figure S1 – Additional results for individual BAB4539. (**A**) Result of Agilent customized aCGH (MGL V9.1) of chromosome 9q region harboring a 21.7 Mb triplication (log_2 ratio ~1) . This array includes SNP probes to detect regions with absence of heterozygosity (AOH). (**B**) SNP array data (B-allele frequency) for chromosome 9 showing 50.3 Mb of AOH from end of triplicated segment to the telomere.

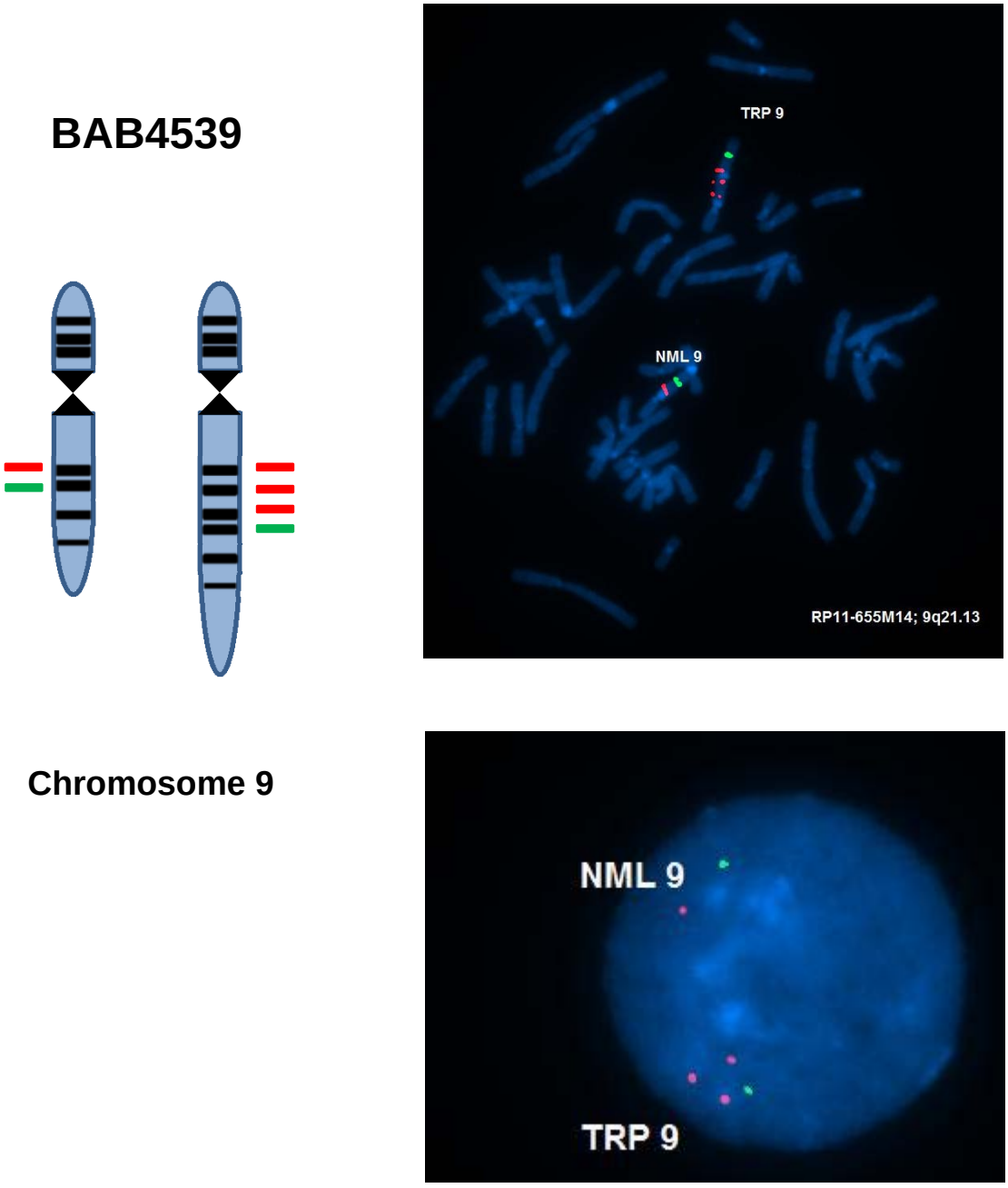

Figure S2. FISH result for subject BAB4539 indicates intrachromosomal triplication: three evenly spaced fluorescent signals for the triplicated segments are arranged along 9q21.13. Metaphase FISH was performed using BAC probes RP11-655M14 (test, red) and RP11-338N12 (control, green). Approximate locations of probes in 9q arm are schematically shown.

Figure S3 - Human Genome Structural Variation Project (HGSV) discordant end alignment track at the UCSC website (hg18) for region chr9:89,705,673-89,944,931, corresponding to hg19 region chr9:90,515,853-90,755,111, suggests the presence of a polymorphic inversion between inverted repeats IR1 and IR2 (each represented by the orange rectangles at the segmental dups track). Seven out nine fosmid libraries have clones that map to the potential inversion breakpoint as indicated by the green arrows on the left. Each fosmid library was constructed using whole genome from individuals of distinct ethnicity[1].

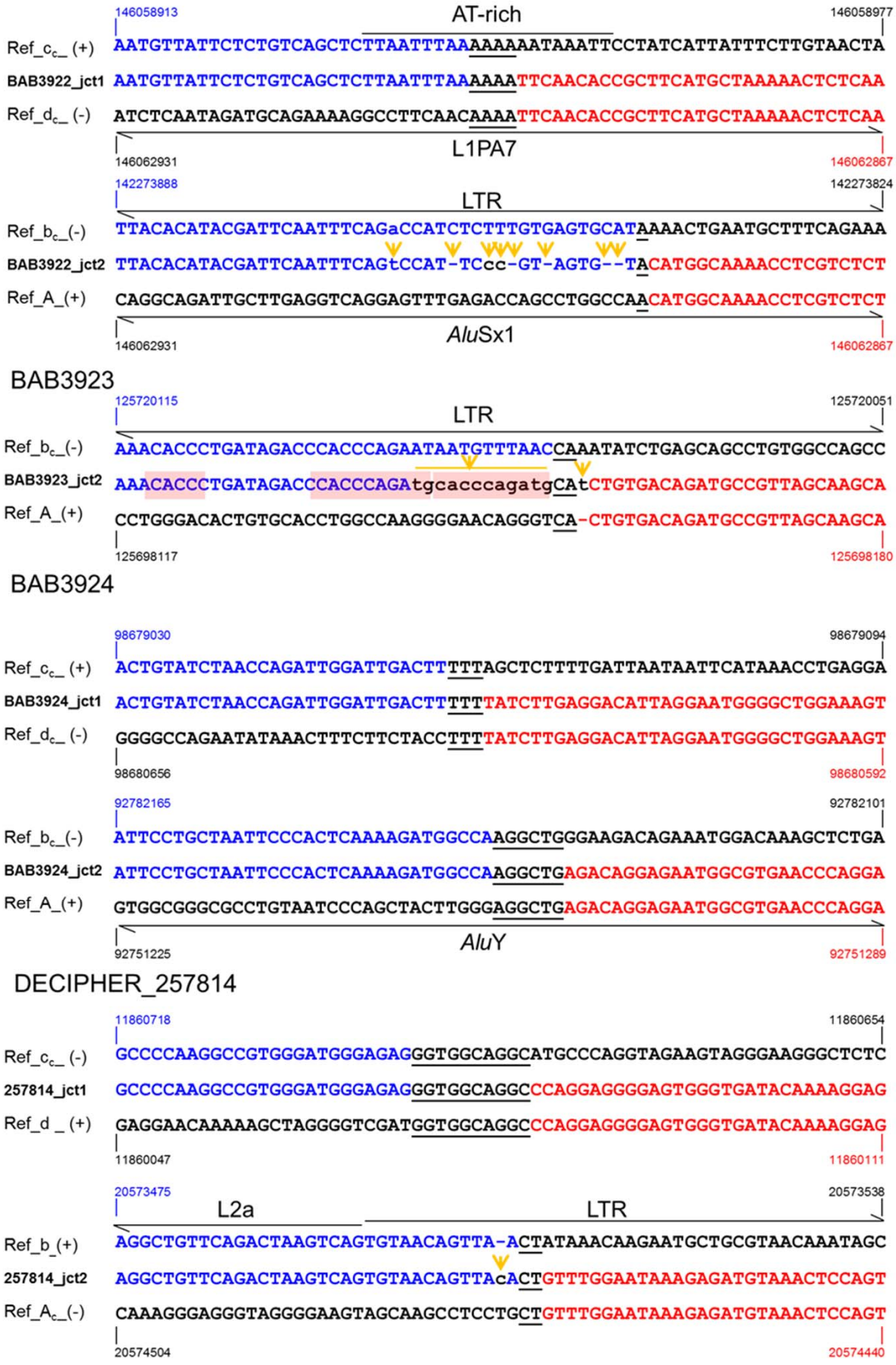
BAB3922
146058913
AT-rich
146058977
Ref_c_ (+)
AATGTTATTCTCTGTCAGCTCTTAATTTAAAAAAAATAAATTCCTATCATTATTTCTTGTAACTA
BAB3922_jct1
AATGTTATTCTCTGTCAGCTCTTAATTTAAAAATTCAACACCGCTTCATGCTAAAAACTCTCAA
Ref_d_ (-)
ATCTCAATAGATGCAGAAAAGGCCTTCAACAAAATTCAACACCGCTTCATGCTAAAAACTCTCAA
L1PA7
146062931
146062867
142273888
142273824
LTR
Ref_b_(-)
TTACACATACGATTCAATTTCAGaCCATCTCTTTGTGAGTGCATAAAACTGAATGCTTTCAGAAA
BAB3922_jct2
TTACACATACGATTCAATTTCAGtCCAT-TCcc-GT-AGTG--TACATGGCAAAACCTCGTCTCT
Ref_A_(+)
CAGGCAGATTGCTTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGCAAAACCTCGTCTCT
AluSx1
146062931
146062867
BAB3923
125720115
LTR
125720051
Ref_b_(-)
AAACACCCTGATAGACCCACCCAGAATAATGTTTAACCAAATATCTGAGCAGCCTGTGGCCAGCC
BAB3923_jct2
AAACACCCTGATAGACCCACCCAGAtgcacccagatgCAtCTGTGACAGATGCCGTTAGCAAGCA
Ref_A_(+)
CCTGGGACACTGTGCACCTGGCCAAGGGGAACAGGGTCA-CTGTGACAGATGCCGTTAGCAAGCA
125698117
125698180
BAB3924
98679030
98679094
Ref_c_ (+)
ACTGTATCTAACCAGATTGGATTGACTTTTTAGCTCTTTTGATTAATAATTCATAAACCTGAGGA
BAB3924_jct1
ACTGTATCTAACCAGATTGGATTGACTTTTTATCTTGAGGACATTAGGAATGGGGCTGGAAAGT
Ref_d_ (-)
GGGGCCAGAATATAAACTTTCTTCTACCTTTTATCTTGAGGACATTAGGAATGGGGCTGGAAAGT
98680656
98680592
92782165
92782101
Ref_b_(-)
ATTCCTGCTAATTCCCACTCAAAAGATGGCCAAGGCTGGGAAGACAGAAATGGACAAAGCTCTGA
BAB3924_jct2
ATTCCTGCTAATTCCCACTCAAAAGATGGCCAAGGCTGAGACAGGAGAATGGCGTGAACCCAGGA
Ref_A_(+)
GTGGCGGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGACAGGAGAATGGCGTGAACCCAGGA
AluY
92751225
92751289
DECIPHER_257814
11860718
11860654
Ref_c_ (-)
GCCCCAAGGCCGTGGGATGGGAGAGGGTGGCAGGCATGCCCAGGTAGAAGTAGGGAAGGGCTCTC
257814_jct1
GCCCCAAGGCCGTGGGATGGGAGAGGGTGGCAGGCCCAGGAGGGGAGTGGGTGATACAAAAGGAG
Ref_d _ (+)
GAGGAACAAAAAGCTAGGGGTCGATGGTGGCAGGCCCAGGAGGGGAGTGGGTGATACAAAAGGAG
11860047
11860111
20573475
20573538
L2a
LTR
Ref_b_(+)
AGGCTGTTCAGACTAAGTCAGTGTAACAGTTA-ACTATAAACAAGAATGCTGCGTAACAAATAGC
257814_jct2
AGGCTGTTCAGACTAAGTCAGTGTAACAGTTAcACTGTTTGGAATAAAGAGATGTAAACTCCAGT
Ref_A_(-)
CAAAGGGAGGGTAGGGGAAGTAGCAAGCCTCCTGCTGTTTGGAATAAAGAGATGTAAACTCCAGT
20574504
20574440

Figure S4 – Color-matched sequence alignment of breakpoint junctions in rearrangements present in subjects BAB3922, BAB3923, BAB3924 and DECIPHER_257814. Orange arrows indicate location of *de novo* point mutations or indels that were likely generated concomitantly to the complex rearrangement (DUP-TRP/INV-DUP + AOH). Red boxes in BAB3923 sequencing data highlight the repetitive sequence CACCC and CACCCAGA; this latter sequence may have been copied *de novo* to jct2 along with the insertion of the dinucleotide TG. jct1: breakpoint junction 1; jct2: breakpoint junction 2. Genomic reference segments (Ref) are named according to the structure shown in Figure 1 (main text). Microhomology at the junctions are represented as bold underlined letters.

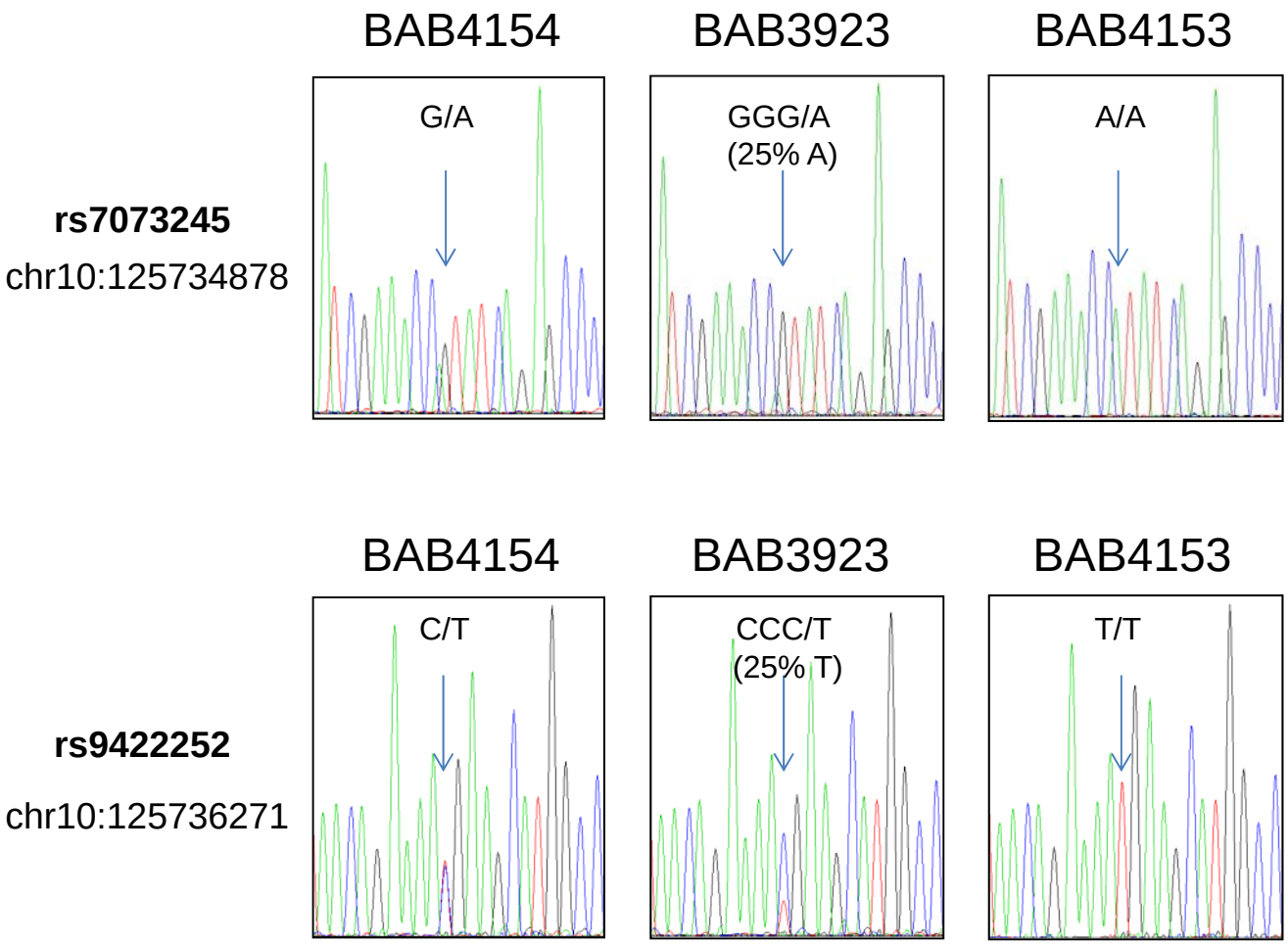

Figure S5- Sanger sequencing result of SNPs rs7073245 and rs9422252 from family trio (BAB3923, BAB4153-mother, BAB4154-father) both of which had shown inconsistent inheritance in the Affymetrix platform. Alleles inherited from father are consistently detected in higher quantity (G and C, respectively) indicating unequal allele dosage within part of the triplicated segment. These results support formation of jct3 in BAB3923 (see main text for details).

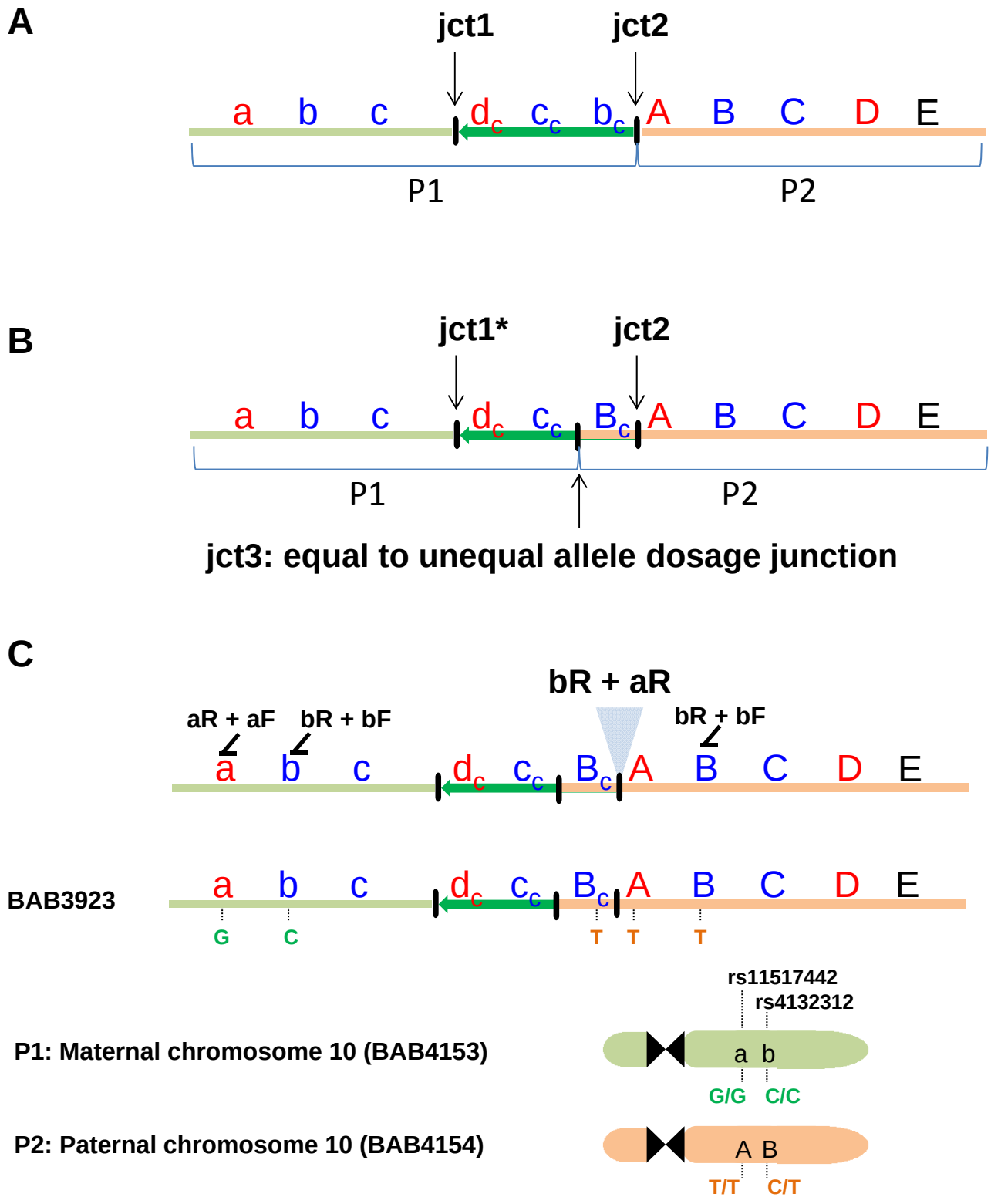

Figure S6- **(A)** DUP-TRP/INV-DUP model predicts formation of at least two breakpoint junctions in subject BAB3923 as a result of two template switches (refer to Figure 1, main text). **(B)** Analysis of allele peaks and B-allele frequency for BAB3923 subject from Affymetrix and Illumina SNP array platforms, respectively, shows evidence for at least one additional junction (jct3) within the triplicated segment. **(C)** Sanger sequencing of SNPs rs11517442 and rs4132312 in parents and index subjects shows that both segments that form BAB3923 jct2 were inherited exclusively from father, supporting the hypothesis that jct2 is intrachromosomal in this subject as opposed to be interchromosomal as expected from model in (**A**). aR + aF and bR + bF: PCR products of the ancestral segments involving in the formation of jct2; bR + aR: BAB3923 breakpoint junction PCR specific product. For BAB3923, jct1* was inferred from high-density custom array CGH. P1 and P2: inherited parental homologous chromosomes.

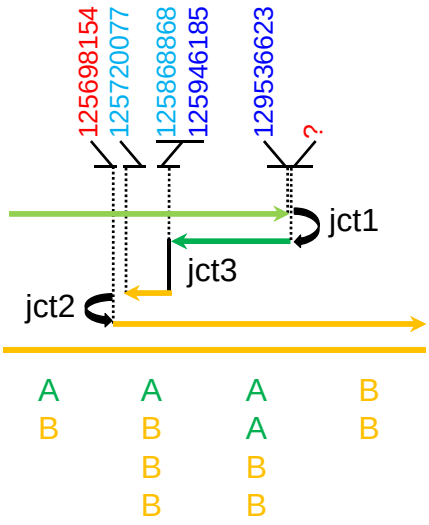

Figure S7: Color-matched schematic representation of alterations observed in chromosome 10 of subject BAB3923 using SNP array and aCGH platforms plus Sanger sequencing for jct2 (refer to the main text for details and Figure 1 for model). BAB3923 is hypothesized to have at least three breakpoint junctions generated by template-switching during a replication-based repair (Figure S6). Top: genomic coordinates (hg19) of breakpoint junctions in chromosome 10 inferred from techniques: jct2: sequencing data is shown in Figure S4; jct1 was inferred from high-density aCGH; jct3 was inferred from Illumina SNP array. Red represents duplicated segment; light blue represents triplicated segment harboring unequal allele dosage; dark blue represents triplicated segment harboring equal allele dosage; ?: duplication junction is hypothesized but genomic coordinates are unknown. Botton: expected allele distribution in distinct segments of the CGR in a SNP array platform.

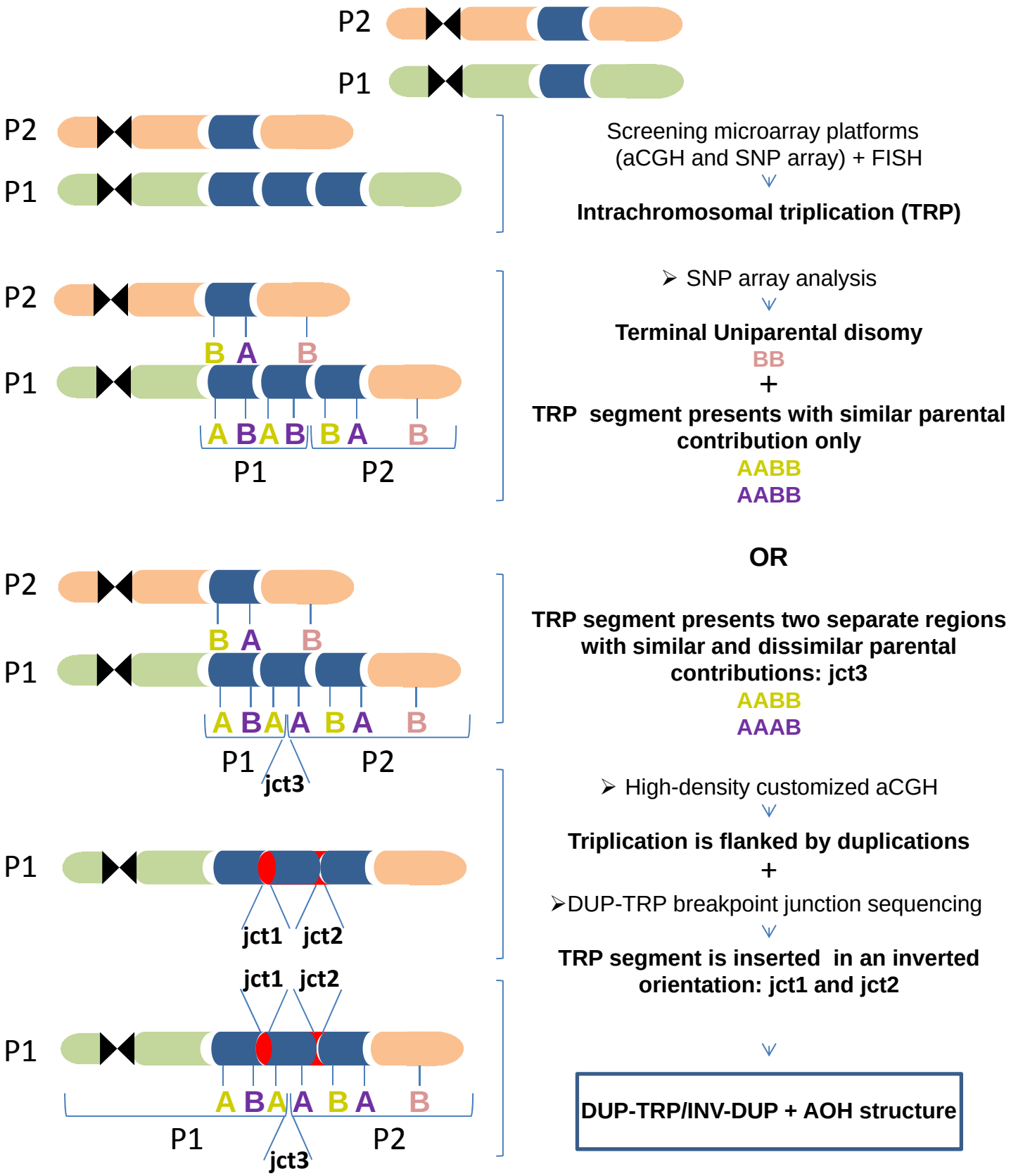

Figure S8- Schematic representation of techniques used to study triplications followed by absence of heterozygosity (AOH). Multiple layers of complexity were unveiled using diverse molecular tools as follows: screening SNP or CGH array platforms, FISH, high-resolution customized aCGH, long-range PCR followed by Sanger sequencing of CGR breakpoint junctions. P1, P2: parental homologous chromosomes; jct: breakpoint junction. Blue segment represents triplication; red segment represents duplication; A, B: SNP alleles. Jct3 were observed in two individuals only (BAB3923 and DECIPHER_257814)

Supplemental reference

1. Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F., et al. (2008). Mapping and sequencing of structural variation from eight human genomes. Nature *453*, 56-64.