Data S2: Figure series related to Figure 5

Inversion polymorphisms affecting SD architectures with suggested implications for morbid CNV formation

Table of Contents

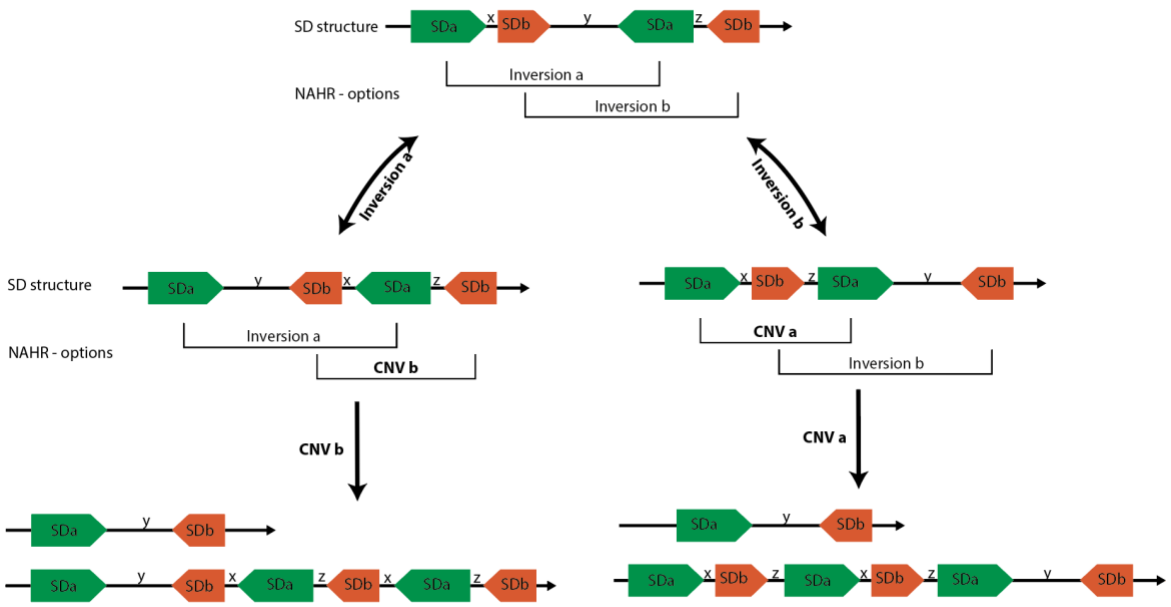

Figure i: Illustration of how inversions may alter SD architectures to result in putative pre-mutational states for morbid CNV formation.

Conceptual drawing of a hypothetical locus with two overlapping pairs of identical SDs (designated SDa and SDb). In this example, starting from a pre-inverted locus (top row), a NAHR-mediated inversion, using the SD architecture, would change the orientation of an SD pair (middle row) to now potentially allow subsequent morbid CNV formation by NAHR (bottom row).

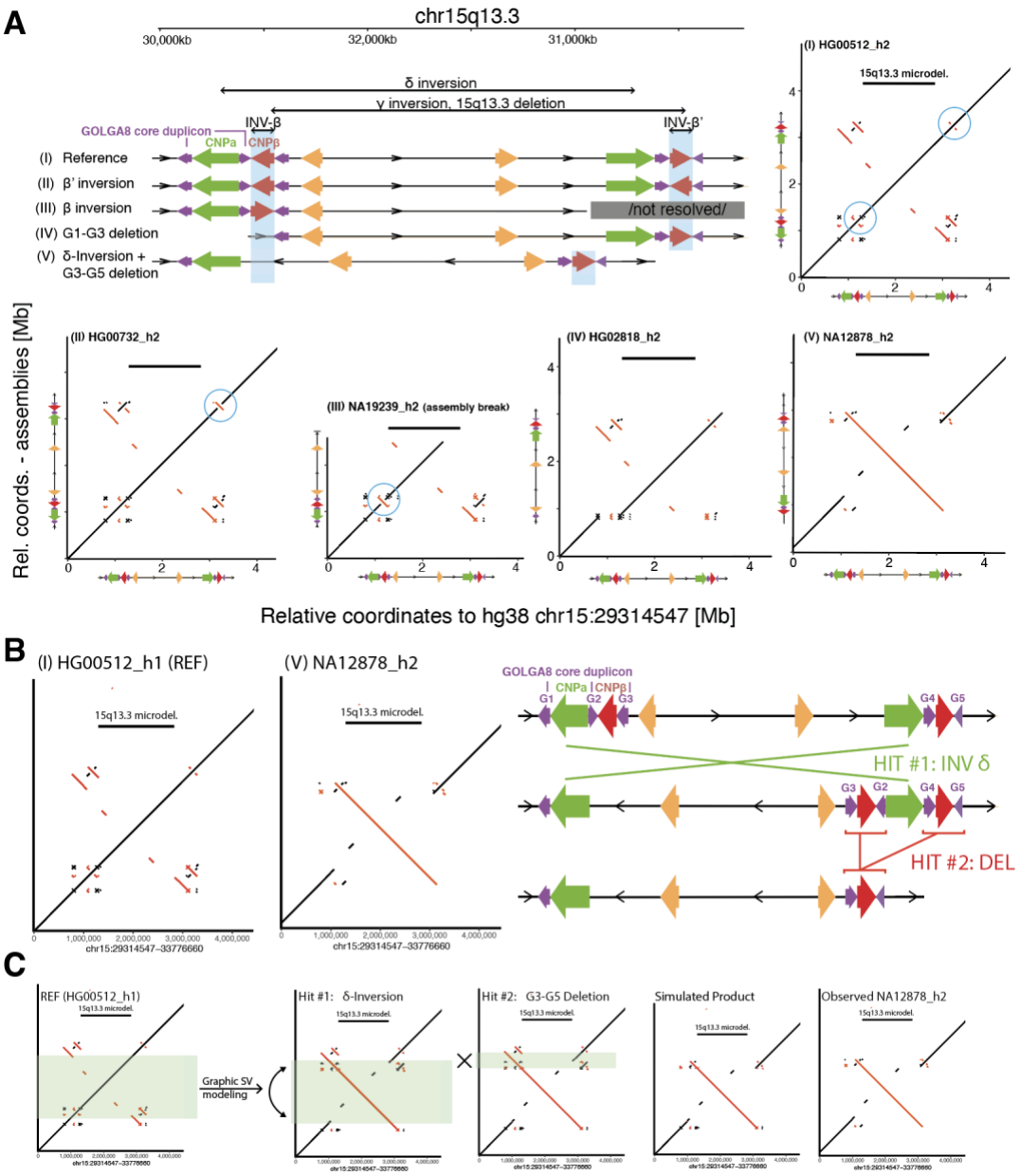

Figure ii: Detailed view of the 15q13.3 inversion/microdeletion region.

A) Repeat structure and SVs observed in the 15q13.3 microdeletion region. SD structures of different haplotype assemblies are depicted as cartoons (top left) and as dot plots (bottom, right). We observe inversions of both copies of CNPβ (red arrows, blue circles, haplotypes II and III), implicated in formation of the recurrent inversion γ and the 15q13.3 microdeletion. We additionally observe one haplotype structure (IV) carrying a deletion of CNPβ and CNPα, as well as a haplotype (V) carrying inversion δ in conjunction with a deletion of CNPβ and CNPα. We hypothesise that these deletions act as a protective allele both against the 15q13.3 microdeletion as well as the γ and δ inversions, thereby preventing further inversion toggling. We note that this model may explain why analysis of INV-β alone has not yielded a significant correlation with 15q13.3 morbid CNV formation (Antonacci et al., 2014). **B)** A two-hit model that may explain the formation of haplotype (V) as a series of (1) NAHR-mediated inversion δ, followed by (2) NAHR-mediated deletion using the directly oriented copies of CNPβ. **C)** A graphical examination shows the plausibility of the two rearrangements that may have led to haplotype (V).

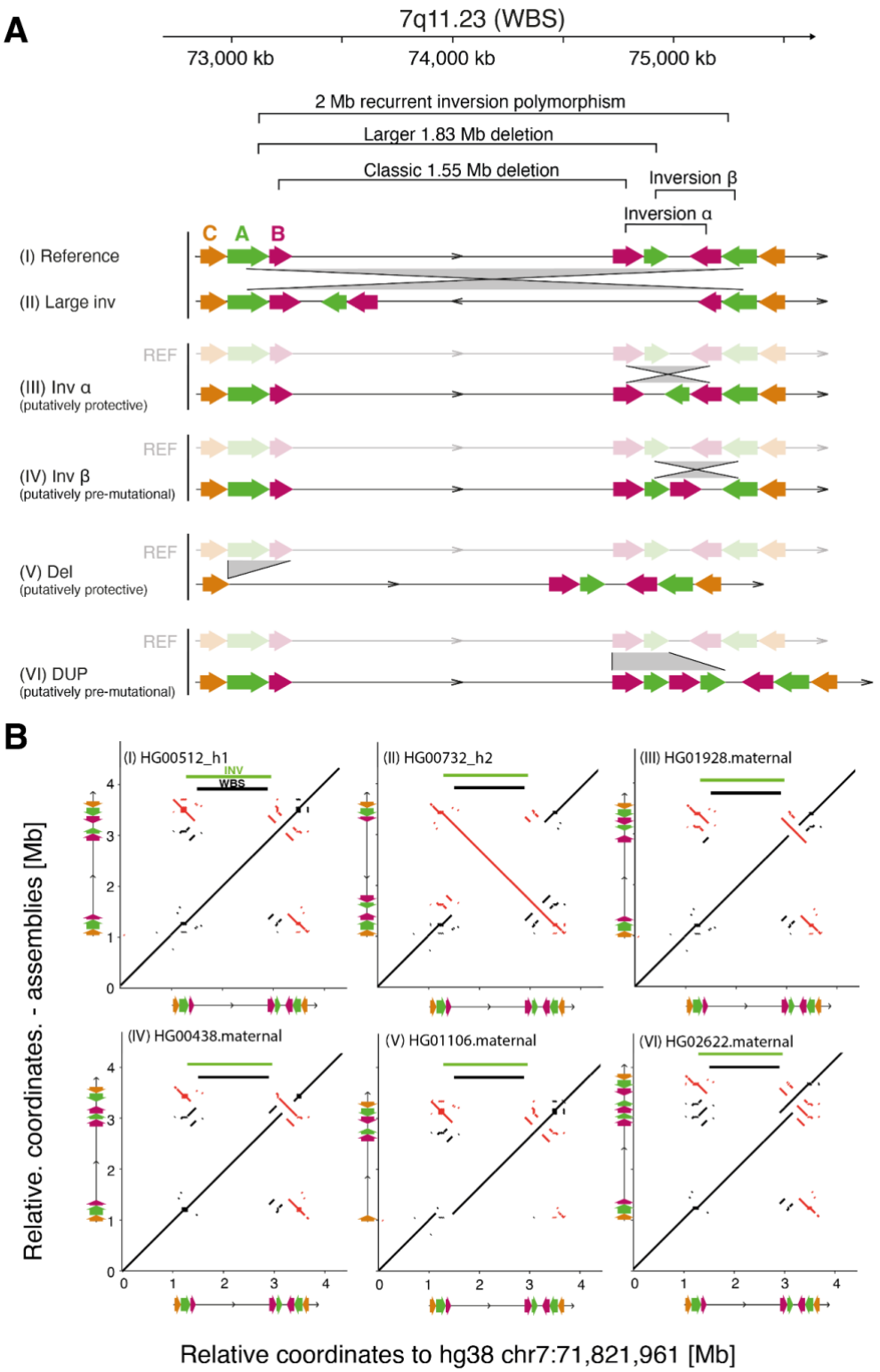

Figure iii: Detailed view of haplotypes in the 7q11.23 WBS region.
A) Overview of genomic rearrangements described in the region and organisation of SDs in six samples, derived from phased whole-genome assemblies. We predict the nested inversions α and β (chr7:74,869,916-75,074,184 and chr7:74,984,488-75,217,887) affect the SD architecture of this region, each producing potentially protective or pre-mutation structural haplotypes with respect to subsequent inversions and CNVs (haplotypes III and IV). We find further haplotypes with a likely protective deletion (haplotype V) and a potentially CNV-enabling duplication (haplotype VI). **B)** Dotplots of the region in the samples described in (A). Inferred loci and SD structures are indicated along the axes. Sample-specific sequence data underlying cases (III-VI) correspond to open access-phased assemblies provided by the Human Pangenome Reference Consortium (HPRC) (**Data and Software Availability**).

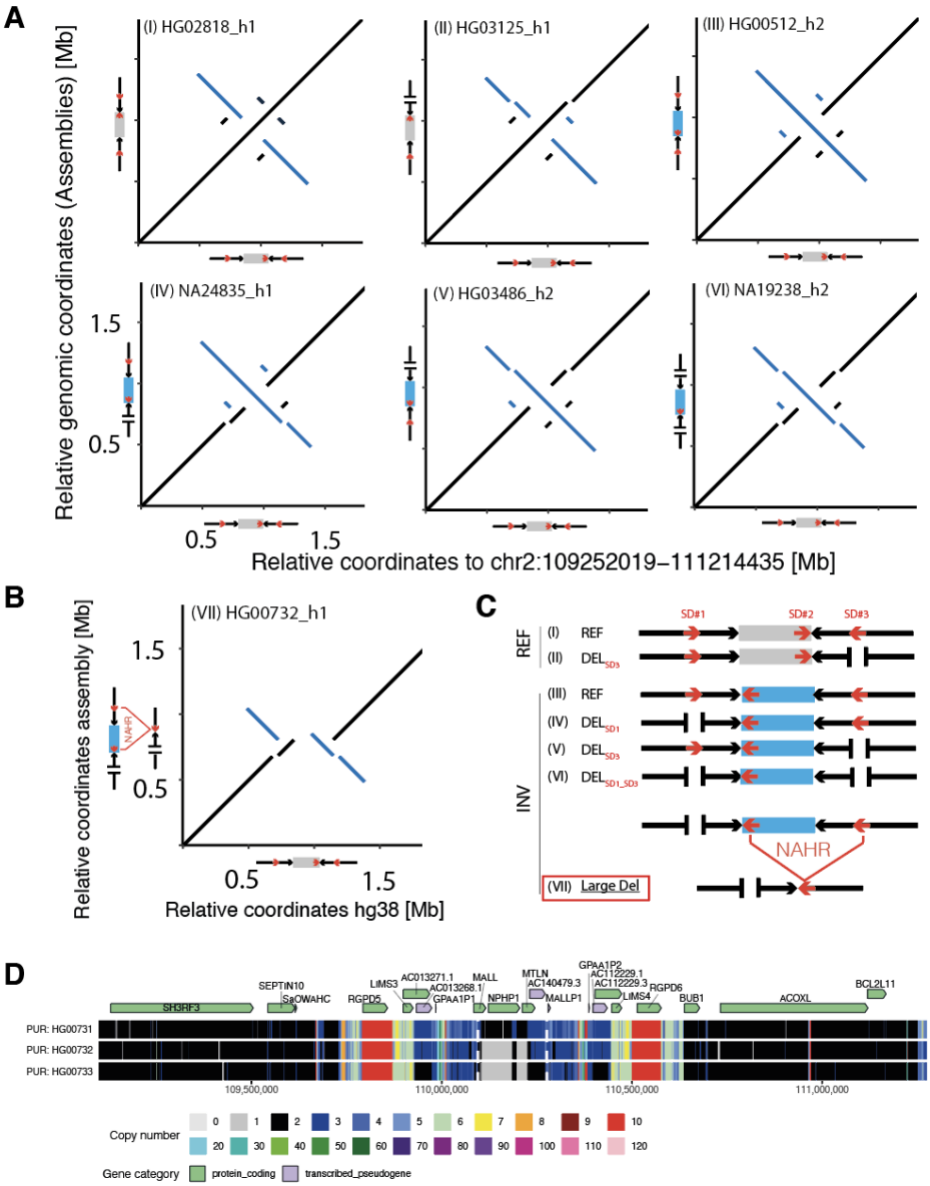

Figure iv: Detailed view of the 2q13 balanced inversion region.

A) Dot plots reflecting all represented haplotype structures found across samples used in this study. Inferred structures are added as cartoons along the axes, with the inversion locus in gray (reference) or blue (inverted), and SDs in orange. Deletions of SD1 and SD2 are observed both in presence and absence of the 2q13 inversion, which may possibly be due to inversion recurrence. **B)** The assembly of sample HG00732_h1 displays a large (~400 kbp) deletion (confirmed by orthogonal Strand-Seq genotyping) encompassing the disease-relevant gene *NPHP1* explicable by NAHR on the inverted haplotype. **C)** Schematic plot of SDs, deletions, and inversions observed across different samples. Based on the SD architecture, we predict that a large deletion can form via NAHR (SDs in tandem orientation) in inverted and reference states (I, III, V and VI), with deletions of individual SDs potentially acting as protective (cases II, IV, VII, VIII). **D)** Short-read-based copy number tracks of the region in the Puerto Rican (PUR) family trio confirm the likely NAHR-mediated heterozygous deletion (copy number=1) in sample HG00732 (mother), which appears to be inherited in HG00733 (child).