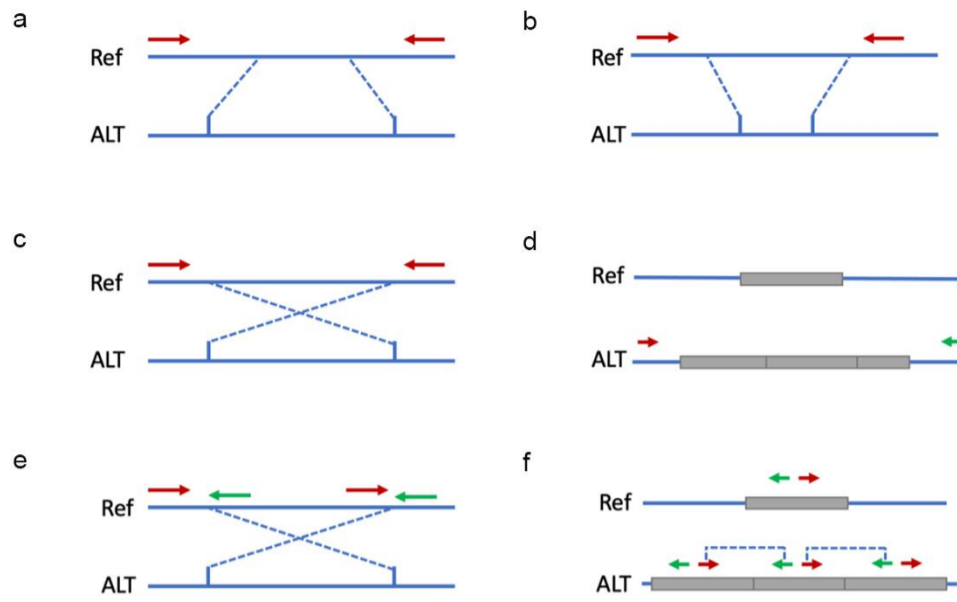


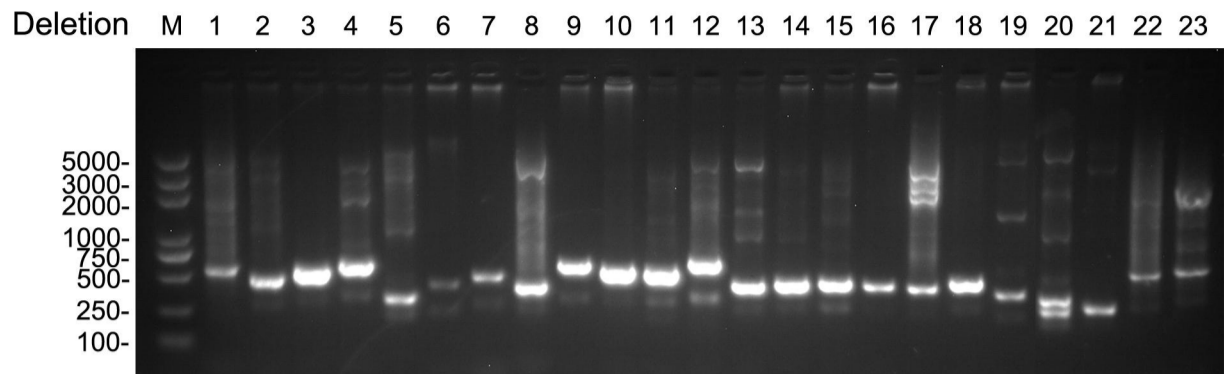
1 **Supplementary Information**



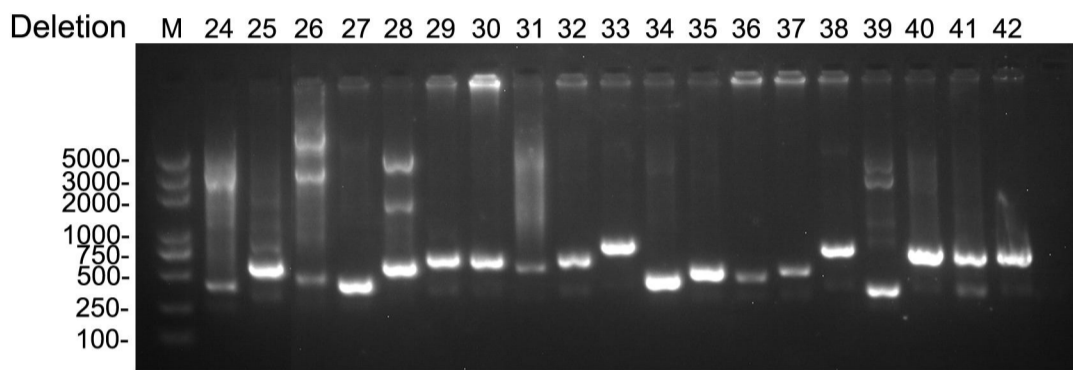
2

3 **Fig. S1:** The scheme of structural variants validation using a PCR-based approach. Primers were designed in the
4 regions outside of breakpoints for insertion (a), deletion (b), inversion (c), and duplication (d) calls. In addition, we
5 also used both forward or both reverse primers for inversions (e) and tail-to-tail primers for duplications (f).

6

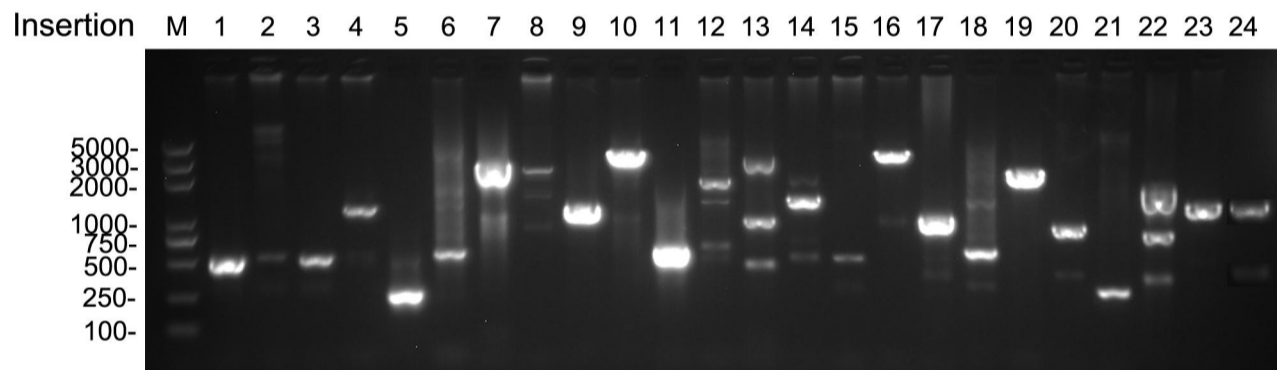


Sanger sequencing + + + + + + + + + + + + + + + + +

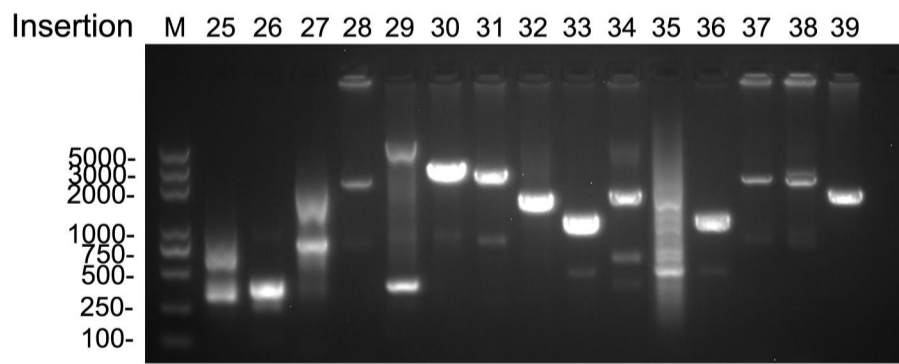


Sanger sequencing + + + + + + +

7

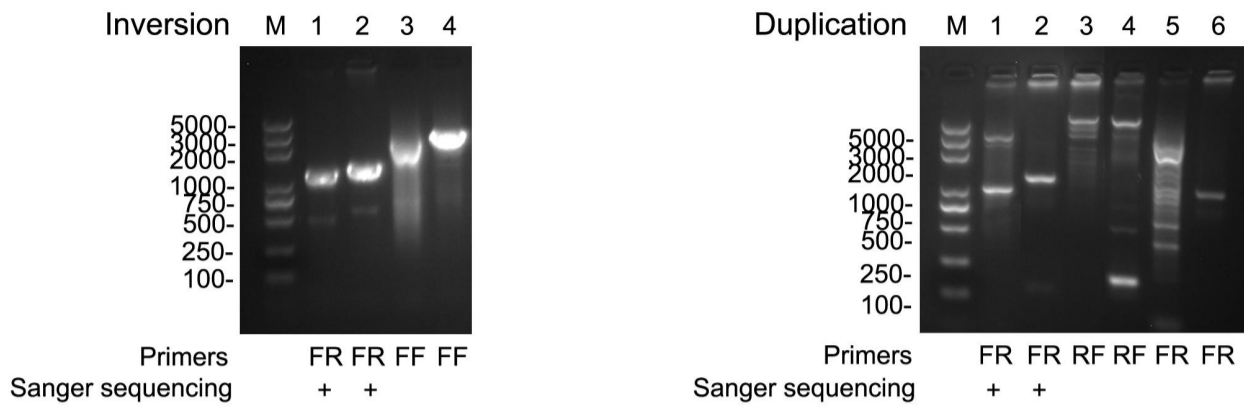


Sanger sequencing +



Sanger sequencing + + + + + + + +

8

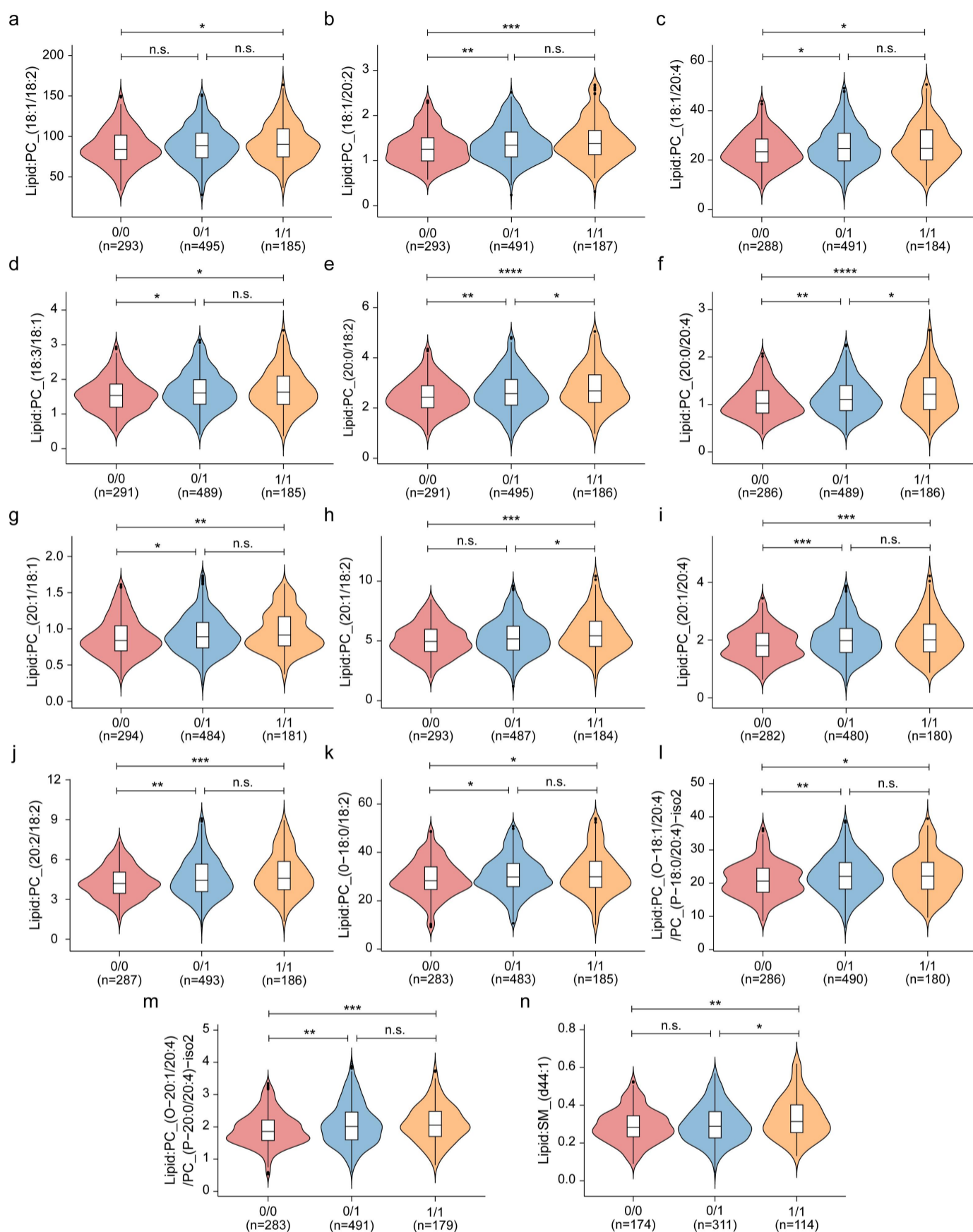


9

10 **Fig. S2:** Agarose gel electrophoresis (1.3%) of PCR products of 91 SVs. We verified a total of 91 SVs including 42
 11 deletions, 39 insertions, 4 inversions, and 6 duplications. “+” indicates the PCR product was sequenced using Sanger
 12 sequencing. M: Size marker. Besides the forward-reverse (FR) primers, we also used forward-forward (FF) or tail-to-
 13 tail (RF) primers for the verifications of inversions and duplications.

14

15

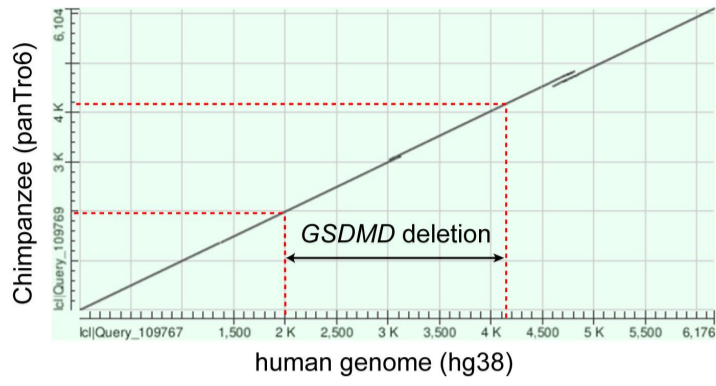


16

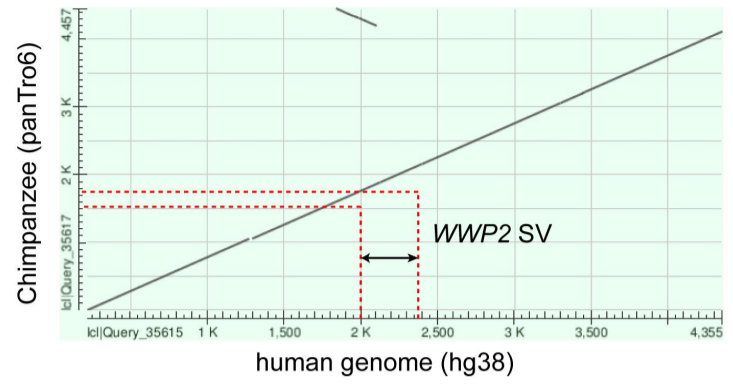
17 **Fig. S3:** Significant increases of 13 phosphatidylcholines (**a**, Lipid: PC_(18:1/18:2); **b**, Lipid: PC_(18:1/20:2). **c**,
 18 Lipid: PC_(18:1/20:4). **d**, Lipid: PC_(18:3/18:1). **e**, Lipid: PC_(20:0/18:2). **f**, Lipid: PC_(20:0/20:4). **g**, Lipid:
 19 PC_(20:1/18:1). **h**, Lipid: PC_(20:1/18:2). **i**, Lipid: PC_(20:1/20:4). **j**, Lipid: PC_(20:2/18:2). **k**, Lipid:
 20 PC_(0-18:0/18:2). **l**, Lipid: PC_(O-18:1/20:4)/PC_(P-18:0/20:4)-iso2 and **m**, Lipid:
 21 PC_(O-20:1/20:4)/PC_(P-20:0/20:4)-iso2 and sphingomyelin (**n**, Lipid: SM_(d44:1)) were observed in SV carriers
 22 than the in non-carriers. X-axes indicate the samples of different genotypes and Y-axes indicate the levels of 13
 23 phosphatidylcholines and sphingomyelin. *P* values were calculated by Kruskal-Wallis test for a-n. **p* <0.05, ***p*
 24 <0.01, ****p* <0.001, *****p* <0.0001, n.s. non-significant.

25

a



b

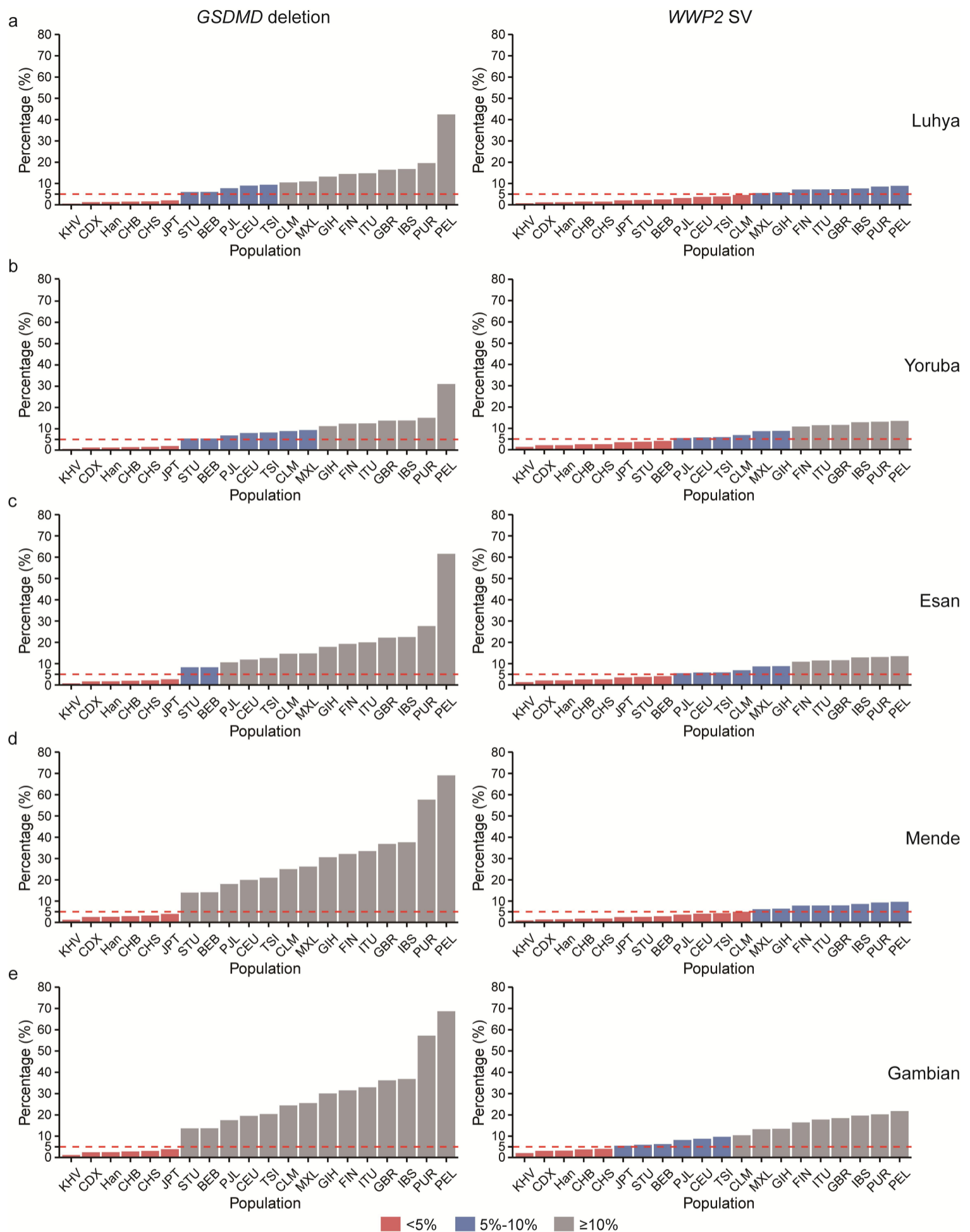


26

27 **Fig. S4:** Dot plots showing the alignment of the human and chimpanzee reference genomes at the *GSDMD* (a) and
28 *WWP2* (b) loci using BLASTN. The marked regions are chr8:143,551,896–143,554,071 (GRCh38) at *GSDMD* locus
29 and chr16:69,820,428–69,820,783 (GRCh38) at *WWP2* locus, respectively.

30

31



32

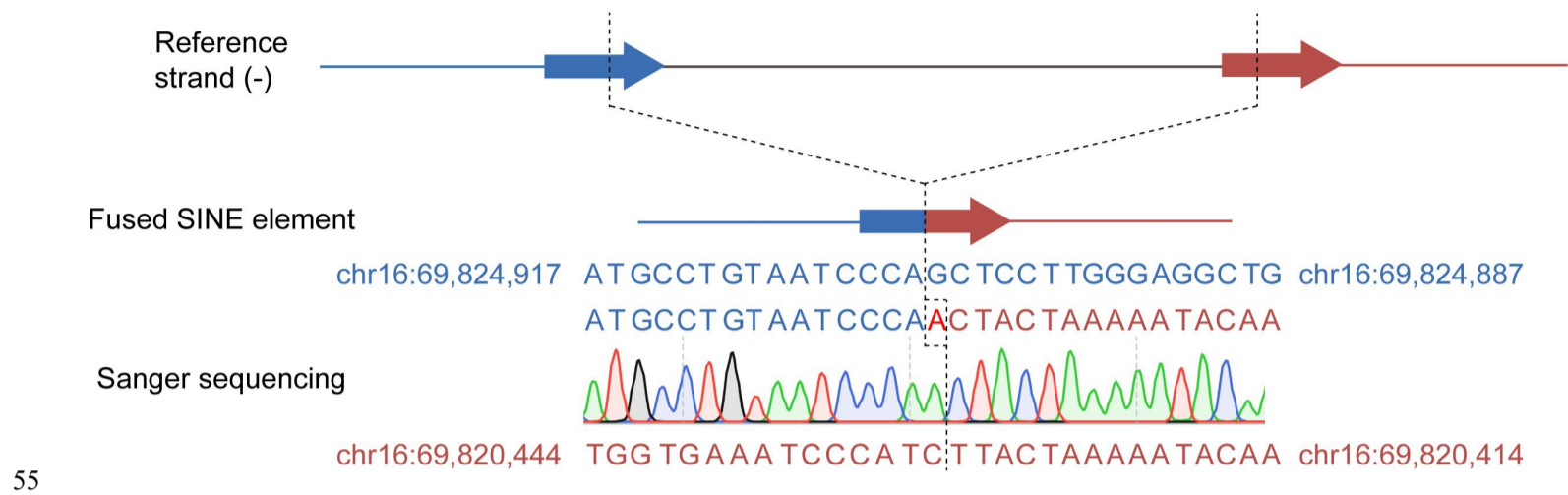
33 **Fig. S5:** Of the 102,119 structural variants (SVs) analyzed, the SVs at the *GSDMD* and *WWP2* loci exhibit among
 34 the highest levels of genetic differentiation, measured by F_{ST} (the fixation index), between African and non-African
 35 populations. F_{ST} was calculated between five African populations and 19 non-African populations from the 1000
 36 Genomes Project, as well as an independent Han Chinese cohort (referred as “Han”) from the present study. For the
 37 calculations, 50 samples were randomly selected to represent each population. The X-axis shows the names of the

38 non-African populations analyzed, while the Y-axis shows the ranking percentile of the *GSDMD* and *WWP2*
39 structural variants (SVs) based on genetic differentiation (F_{ST}) between African and non-African groups across
40 102,119 total SVs. Red, blue and gray bars denote percentile ranks below 5%, between 5-10%, and greater than or
41 equal to 10%, respectively. The red dashed line indicates the F_{ST} ranking cutoff at the 5th percentile.

42 **a**, F_{ST} ranking of SVs at *GSDMD* and *WWP2* loci between Luhya and each of non-African populations. **b**, F_{ST}
43 ranking of SVs at *GSDMD* and *WWP2* loci between Yoruba and each of non-African populations. **c**, F_{ST} ranking of
44 SVs at *GSDMD* and *WWP2* loci between Esan and each of non-African populations. **d**, F_{ST} ranking of SVs at
45 *GSDMD* and *WWP2* loci between Mende and each of non-African populations. **e**, F_{ST} ranking of SVs at *GSDMD* and
46 *WWP2* loci between Gambian and each of non-African populations.

47 IBS: Iberian populations in Spain; GBR: British in England and Scotland; TSI: Toscani in Italy; FIN: Finnish in
48 Finland; PJI: Punjabi in Lahore, Pakistan; GIH: Gujarati Indians in Houston, TX; ITU: Indian Telugu in the UK;
49 STU: Sri Lankan Tamil in the UK; BEB: Bengali in Bangladesh; CDX: Chinese Dai in Xishuangbanna, China; CHS:
50 Han Chinese South; CHB: Han Chinese in Beijing, China; JPT: Japanese in Tokyo, Japan; KHV: Kinh in Ho Chi
51 Minh City, Vietnam; CEU: Utah residents (CEPH) with Northern and Western European; MXL: Mexican Ancestry
52 in Los Angeles, California; PUR: Puerto Rican in Puerto Rico; CLM: Colombian in Medellin, Colombia; PEL:
53 Peruvian in Lima, Peru.

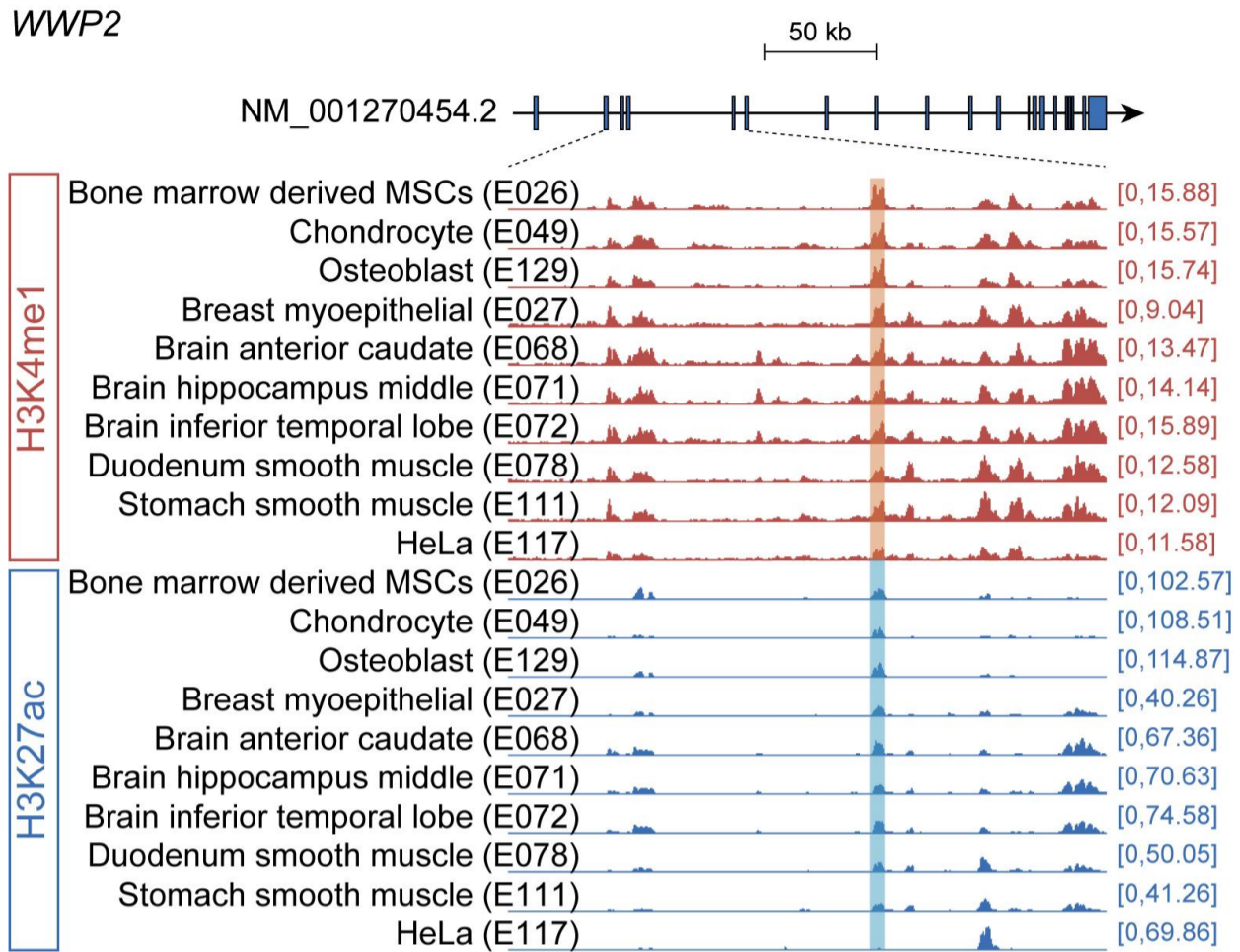
54



55

56 **Fig. S6:** Schematic of the SINE fusion event at the *WWP2* locus.

57

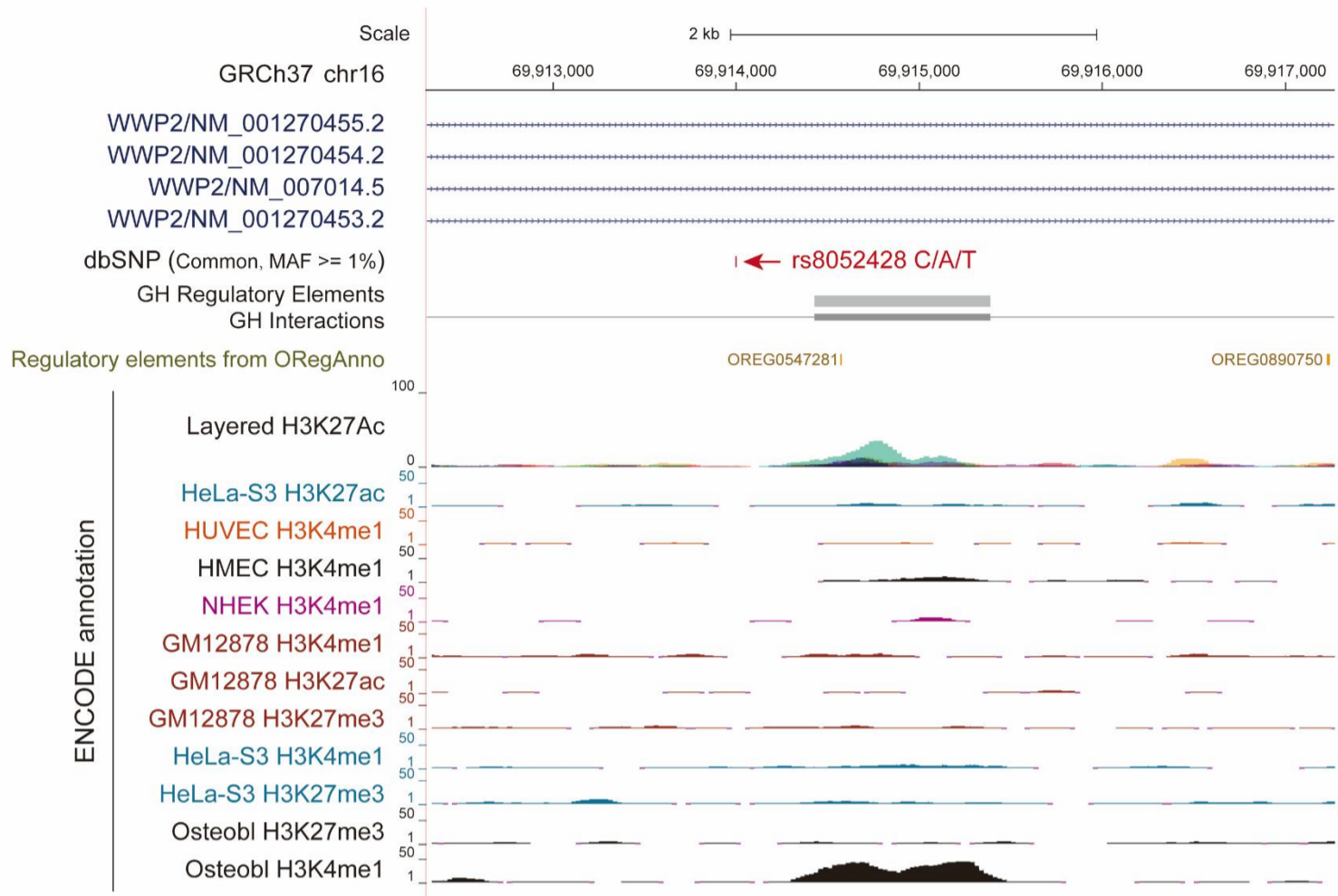


59

60 **Fig. S7:** The deletion in the fourth intron of *WWP2* (NM_001270454.2) eliminates a region that is enriched of
 61 histone modification signals. H3K4me1 and H3K27ac ChIP-seq tracks for Bone marrow derived MSCs (E026),
 62 Mesenchymal stem cell derived chondrocyte (E049), Osteoblast (E129), Breast myoepithelial (E027), Brain anterior
 63 caudate (E068), Brain hippocampus middle (E071), Brain inferior temporal lobe (E072), Duodenum smooth muscle
 64 (E078), Stomach smooth muscle (E111), Hela-S3 cervical carcinoma (E117) from the Roadmap Epigenomics dataset
 65 were visualized in the UCSC Genome Browser.

66

67

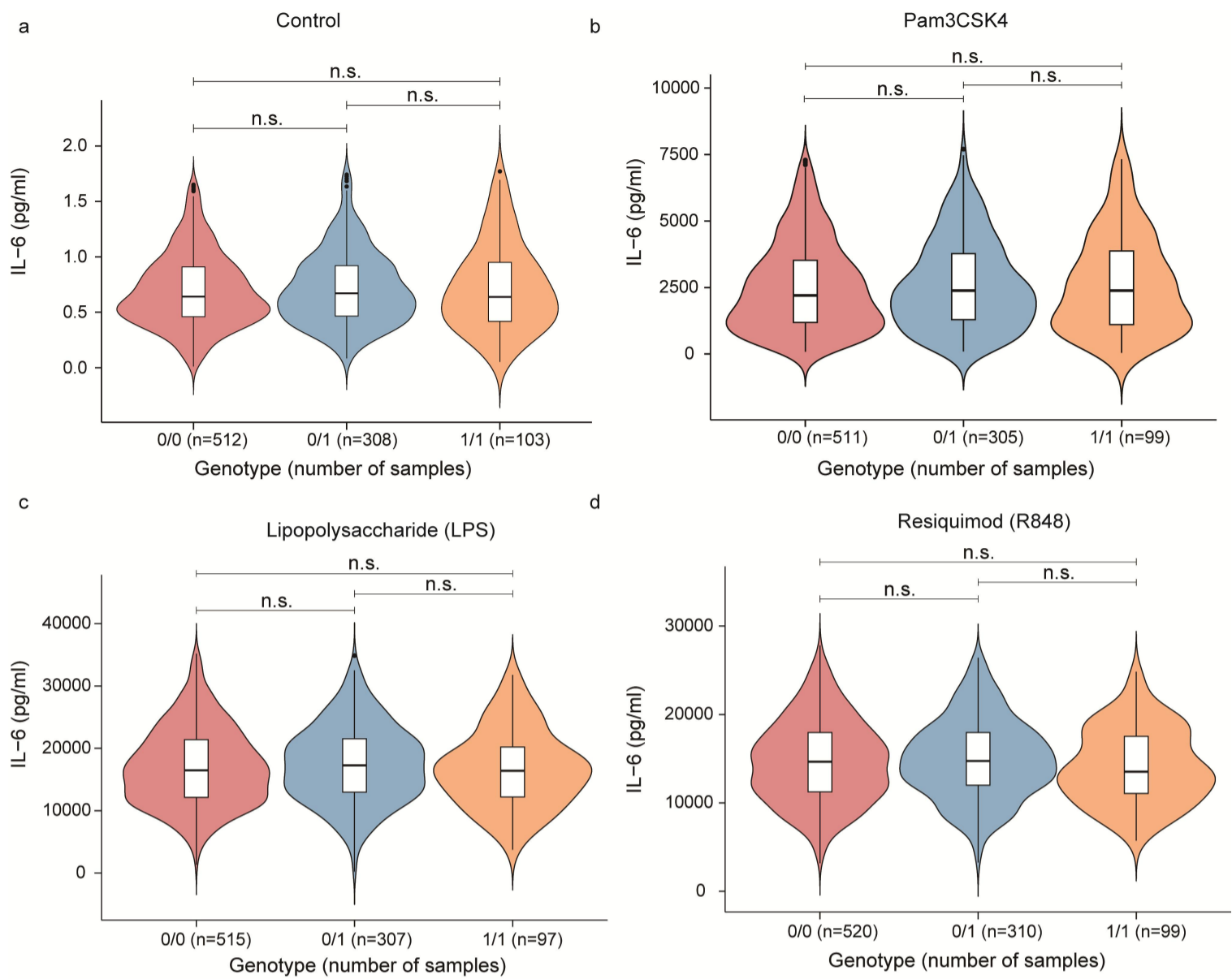


69

70 **Fig. S8:** There is no enhancer annotation surrounding rs8052428 (indicated by the red arrow), which is in linkage
 71 disequilibrium with the SV at the *WWP2* locus. The visualized annotation information includes regulatory elements
 72 from ORegAnno, enhancer annotations from the GeneHancer database, layered H3K27ac histone mark tracks across
 73 7 cell lines (GM12878, H1-hESC, HSMM, HUVEC, K562, NHEK and NHLF), as well as H3K4me1, H3K27me3,
 74 and H3K27ac ChIP-seq tracks for HUVEC, HeLa-S3, HMEC, NHEK, Osteoblasts, and GM12878 cell types from the
 75 ENCODE dataset. All annotations were visualized using the UCSC Genome Browser.

76

77



78

79 **Fig. S9:** No significant differences in IL-6 levels were observed between the carriers and non-carriers of the SV at
80 the *WWP2* locus when no stimulus (negative control, **a**) or stimulated with Pam3CSK4 (**b**), lipopolysaccharide (LPS,
81 **c**), and resiquimod (R848, **d**). X- and Y-axes indicate the samples of different genotypes and the level of IL-6,
82 respectively. *P* values were calculated by Kruskal-Wallis test for **a-d**. n.s. non-significant.

83

84 **Separate Excel file**
85 Tables S1 to S9
86