

**Table S1. *Cis* and *trans* configurations of coding variants in 14 different populations**

Ancestry group <sup>1</sup>	Population <sup>1</sup>	No. phased genomes	<i>Cis</i> configs dam mut <sup>2</sup> (%) <sup>4</sup>	<i>Trans</i> configs dam mut <sup>2</sup> (%) <sup>5</sup>	<i>Cis</i> configs AA exch <sup>3</sup> (%) <sup>4</sup>	<i>Trans</i> configs AA exch <sup>3</sup> (%) <sup>5</sup>	<i>Cis</i> configs sSNPs (%) <sup>4</sup>	<i>Trans</i> configs sSNPs (%) <sup>5</sup>
EUR	GBR	89	61.3	38.7	60.6	39.4	63.0	37.0
	FIN	93	61.4	38.6	60.8	39.2	62.9	37.1
	IBS	14	61.6	38.4	60.9	39.1	63.1	36.9
	CEU	85	61.6	38.4	61.0	39.0	62.9	37.1
	TSI	98	60.8	39.2	60.4	39.6	62.7	37.3
EAS	CHS	100	59.6	40.4	60.2	39.8	62.8	37.2
	CHB	97	59.5	40.5	60.5	39.5	62.7	37.3
	JPT	89	59.6	40.4	59.8	40.2	63.0	37.0
AMR	PUR	55	60.1	39.9	59.7	40.3	62.3	37.7
	CLM	60	60.1	39.9	59.4	40.6	62.0	38.0
	MXL	66	60.0	40.0	59.6	40.4	62.7	37.3
AFR	YRI	88	54.4	45.6	53.2	46.8	54.7	45.3
	LWK	97	54.0	46.0	53.0	47.0	54.6	45.4
	ASW	61	56.3	43.7	55.6	44.4	56.8	43.2

<sup>1</sup> Description of ancestry groups and population samples in Abecasis et al. (2012), statistically haplotype-resolved genomes from the 1000 Genomes (1000G) Project.

<sup>2</sup> 'Potentially damaging', or 'damaging', mutations from 1000G annotation database; also described as 'protein-altering' in text.

<sup>3</sup> Amino acid (AA) exchanges from 1000G database.

<sup>4</sup> Values represent the median of *cis* fractions per genome (%); these were assessed as the number of *cis* configurations observed across all autosomal protein-coding genes (RefSeq hg19 from UCSC table browser), divided by total configuration count.

<sup>5</sup> Analogous to <sup>4</sup>.

Dam mut, damaging mutations; AA exch, amino acid exchanges; sSNPs, synonymous SNPs.

**Table S2A. *Cis* and *trans* configurations in relation to number of protein-altering mutations**

No. mutations <sup>1</sup> per gene	Sample <sup>2</sup> No. genomes <sup>3</sup>	Total configs <sup>4</sup>		<i>Cis</i> configs <sup>5</sup>		<i>Trans</i> configs <sup>5</sup>		<i>Cis/Trans</i> ratio		
		No.	%	No.	%	No.	%	<i>Cis</i> (%) <sup>6</sup>	<i>Trans</i> (%) <sup>6</sup>	
2	<b>1000G</b> 1,092	375,456	66.7	247,220	74.6	128,236	55.4	65.8	34.2	
3		104,687	18.6	52,316	15.8	52,371	22.6	50.0	50.0	
4		40,826	7.3	19,033	5.7	21,793	9.4	46.6	53.4	
5		17,433	3.1	6,919	2.1	10,514	4.5	39.7	60.3	
...		...	...	...	...	...	...	...	...	...
Total		562,698	100	331,250	100	231,448	100	58.9	41.1	
2	<b>EUR</b> 379	123,119	67.1	83,432	74.3	39,687	55.8	67.8	32.2	
3		33,595	18.3	17,622	15.7	15,973	22.4	52.5	47.5	
4		13,366	7.3	6,818	6.1	6,548	9.2	51.0	49.0	
5		5,612	3.1	2,384	2.1	3,228	4.5	42.5	57.5	
...		...	...	...	...	...	...	...	...	...
Total		183,536	100	112,362	100	71,174	100	61.2	38.8	
2	<b>EAS</b> 286	86,830	67.7	57,947	75.8	28,883	55.8	66.7	33.3	
3		22,432	17.5	11,163	14.6	11,269	21.8	49.8	50.2	
4		9,005	7.0	4,201	5.5	4,804	9.3	46.7	53.3	
5		4,092	3.2	1,865	2.4	2,227	4.3	45.6	54.4	
...		...	...	...	...	...	...	...	...	...
Total		128,212	100	76,436	100	51,776	100	59.6	40.4	
2	<b>AMR</b> 181	61,418	66.8	40,873	74.0	20,545	56.0	66.5	33.5	
3		16,977	18.5	8,810	16.0	8,167	22.2	51.9	48.1	
4		6,680	7.3	3,407	6.2	3,273	8.9	51.0	49.0	
5		2,824	3.1	1,108	2.0	1,716	4.7	39.2	60.8	
...		...	...	...	...	...	...	...	...	...
Total		91,925	100	55,213	100	36,712	100	60.1	39.9	
2	<b>AFR</b> 246	104,089	65.5	64,968	74.5	39,121	54.5	62.4	37.6	
3		31,683	19.9	14,721	16.9	16,962	23.6	46.5	53.5	
4		11,775	7.4	4,607	5.3	7,168	10.0	39.1	60.9	
5		4,905	3.1	1,562	1.8	3,343	4.7	31.8	68.2	
...		...	...	...	...	...	...	...	...	...
Total		159,025	100	87,239	100	71,786	100	54.9	45.1	
2	<b>PGP</b> 184	45,313	64.8	30,264	71.6	15,049	54.4	66.8	33.2	
3		13,886	19.9	7,520	17.8	6,366	23.0	54.2	45.8	
4		4,716	6.7	2,018	4.8	2,698	9.8	42.8	57.2	
5		2,347	3.4	986	2.3	1,361	4.9	42.0	58.0	
...		...	...	...	...	...	...	...	...	...
Total		69,936	100	42,265	100	27,671	100	60.4	39.6	

<sup>1</sup> Protein-altering, equivalent to (potentially) damaging mutations from 1000 Genomes (1000G) annotation database (Abecasis et al. 2012), or annotated by PolyPhen-2 (Adzhubei et al. 2010) and SIFT (Ng and Henikoff 2001) as well as GERP conservation scores (Cooper et al. 2005) in the PGP genomes. The number '2' of protein-altering mutations per gene for instance means that all genes that have precisely 2 mutations in a gene have been pooled and analysed.

<sup>2</sup> Description of 1000G samples and the four different ancestry groups (EUR, EAS, AMR, AFR) in Abecasis et al. (2012); description of PGP samples in Mao et al. (2016).

<sup>3</sup> 1,092 genomes (1000G) statistically haplotype-resolved; 184 genomes (PGP) experimentally haplotype-resolved.

<sup>4</sup> The total number of configurations (equivalent to the number of genes with  $\geq 2$  mutations) assessed across specified numbers of genomes from all genes which have given number of protein-altering mutations; expressed in addition as fractions (%) relative to the total number of configurations counted for all genes across all genomes.

<sup>5</sup> The number of *cis* and *trans* configurations, respectively, among the total number of configurations scored in specified samples for genes with given number of mutations; the respective fractions (%) of *cis* and *trans* configurations relative to their total numbers assessed across all genes and genomes are given in addition. The numbers of *cis* and *trans* configurations for any given number of mutations add up to the corresponding totals in <sup>4</sup>.

<sup>6</sup> *Cis* fraction (%) calculated for genes with given number of mutations, dividing the number of *cis* configurations in <sup>5</sup> by the total number of configurations in <sup>4</sup>; *trans* fractions (%) are calculated accordingly. Configs, configurations.

**Table S2B. *Cis* and *trans* configurations in relation to number of AA exchanges and sSNPs**

Type of variant <sup>1</sup>	No. variants per gene <sup>2</sup>	Total configs <sup>3</sup>		No. <i>cis</i> configs <sup>4</sup>	No. <i>trans</i> configs <sup>4</sup>	<i>Cis/Trans</i> ratio	
		No.	%			<i>Cis</i> (%) <sup>5</sup>	<i>Trans</i> (%) <sup>5</sup>
AA exchanges	2	869,855	57.7	590,886	278,969	67.9	32.1
	3	323,655	21.5	171,127	152,528	52.9	47.1
	4	140,027	9.3	63,395	76,632	45.3	54.7
	5	67,527	4.5	27,505	40,022	40.7	59.3
	...	...	...	...	...	...	...
	Total	1,506,764	100	880,549	626,215	58.4	41.6
sSNPs	2	1,037,140	59.6	725,071	312,069	69.9	30.1
	3	381,027	21.9	204,269	176,758	53.6	46.4
	4	156,911	9.0	70,346	86,565	44.8	55.2
	5	72,263	4.2	28,506	43,757	39.4	60.6
	...	...	...	...	...	...	...
	Total	1,740,093	100	1,053,883	686,210	60.6	39.4
AA exchanges and sSNPs combined	2	1,610,775	47.3	1,106,736	504,039	68.7	31.3
	3	777,497	22.8	416,154	361,343	53.5	46.5
	4	406,970	12.0	181,054	225,916	44.5	55.5
	5	224,911	6.6	87,761	137,150	39.0	61.0
	...	...	...	...	...	...	...
	Total	3,403,910	100	1,900,122	1,503,788	55.8/58.2 <sup>6</sup>	44.2/41.8 <sup>6</sup>

<sup>1</sup> Amino acid (AA) exchanges from 1000 Genomes (1000G) Project database (Abecasis et al. 2012); synonymous SNPs (sSNPs).

<sup>2</sup> The number '2' for instance of any type/combination of coding variants per gene specified means that all genes that have precisely 2 variants in a gene have been pooled and analysed in the 1,092 genomes (1000G).

<sup>3</sup> The total number of configurations (equivalent to the number of genes with  $\geq 2$  coding variants) assessed across 1,092 genomes from all genes which have specified number of coding variants, expressed in addition as fractions (%) relative to the total number of configurations counted for all genes across all genomes.

<sup>4</sup> The number of *cis* and *trans* configurations, respectively, among the total number of configurations scored in the 1,092 genomes for genes with specified number of mutations.

<sup>5</sup> *Cis* fraction (%) calculated for genes with given number of coding variants, dividing the number of *cis* configurations in <sup>4</sup> by the total number of configurations in <sup>3</sup>; *trans* fractions (%) are calculated accordingly.

<sup>6</sup> The median values, which are used in main text to be consistent with the other *cis/trans* ratios, are presented in addition. All other values are average values.

Configs, configurations.

**Table S3. Simulation of expected *cis/trans* ratios in relation to number of mutations**

Type of variant	No. variants per gene <sup>2</sup>	Total configs <sup>1</sup>		No. <i>cis</i> configs <sup>2</sup>	No. <i>trans</i> configs <sup>2</sup>	<i>Cis/Trans</i> ratio	
		No.	%			Cis (%)	Trans (%)
Protein-altering	2	436,960	66.4	217,789	219,171	49.8	50.2
	3	132,308	20.1	33,327	98,981	25.2	74.8
	4	45,285	6.9	5,669	39,616	12.5	87.5
	5	17,931	2.7	1,112	16,819	6.2	93.8
	...	...	...	...	...	...	...
	Total	657,875	100	258,267	399,608	39.3	60.7
AA exchanges	2	951,313	59.0	475,137	476,176	49.9	50.1
	3	355,549	22.0	89,295	266,254	25.1	74.9
	4	144,799	9.0	18,029	126,770	12.5	87.5
	5	65,041	4.0	4,056	60,985	6.2	93.8
	...	...	...	...	...	...	...
	Total	1,612,714	100	588,494	1,024,220	36.5	63.5
sSNPs	2	1,098,368	60.3	548,709	549,659	50.0	50.0
	3	407,820	22.4	101,771	306,049	25.0	75.0
	4	162,939	8.9	20,387	142,552	12.5	87.5
	5	70,648	3.9	4,448	66,200	6.3	93.7
	...	...	...	...	...	...	...
	Total	1,821,931	100	676,751	1,145,181	37.1	62.9
AA exchanges and sSNPs	2	1,803,140	50.1	900,986	902,154	50.0	50.0
	3	848,382	23.6	212,572	635,810	25.1	74.9
	4	415,336	11.5	51,629	363,707	12.4	87.6
	5	211,775	5.9	13,293	198,482	6.3	93.7
	...	...	...	...	...	...	...
	Total	3,599,219	100	1,183,593	2,415,626	32.9	67.1

Simulations of phased genomes were performed as described in Methods. To simulate *cis/trans* ratios which can be expected under conditions of random distribution of variants between the two homologues, in principle, each variant had a 50:50 chance to reside on either one of these.

<sup>1</sup> The total number of configurations assessed across 1,092 simulated genomes from all genes which have indicated number of protein-altering mutations, amino acid (AA) exchanges, or synonymous SNPs (sSNPs); expressed in addition as fractions (%) relative to the total number of configurations counted across all genes.

<sup>2</sup> The number of *cis* and *trans* configurations, respectively, among the total number of configurations obtained from the 1,092 simulated genomes from genes with given number of mutations; the respective fractions (%) of *cis* and *trans* configurations relative to their total numbers assessed across all genes with defined number of variants are given in addition. The numbers of *cis* and *trans* configurations for a given number of mutations add up to the corresponding totals in <sup>1</sup>.

**Table S4. Inter-mutation genome distances in *cis*- versus *trans* configurations**

Population Samples <sup>1</sup>	No. phased genomes	<i>Cis</i> distance dam mut <sup>2</sup> (bp)	<i>Trans</i> distance dam mut <sup>2</sup> (bp)	<i>Cis</i> distance AA exch <sup>3</sup> (bp)	<i>Trans</i> distance AA exch <sup>3</sup> (bp)	<i>Cis</i> distance sSNPs <sup>4</sup> (bp)	<i>Trans</i> distance sSNPs <sup>4</sup> (bp)
Total	1,092	1,607	5,125	2,130	6,275	3,665	7,761
EUR	379	1,570	5,290	2,248	6,532	3,638	7,982
AFR	246	1,830	4,771	2,189	5,546	3,888	7,228
PGP	184	2,584	5,984	2,058	5,920	3,562	8,280

<sup>1</sup> Description of 1000 Genomes (1000G) samples and ancestry groups EUR and AFR in Abecasis et al. (2012); description of PGP samples in Mao et al. (2016).

<sup>2</sup> Distance in bp between pairs of (potentially) damaging mutations, equivalent to 'protein-altering' mutations, that reside in *cis* configurations compared to pairs in *trans* configurations; annotation from 1000G annotation database, or annotated by PolyPhen-2 (Adzhubei et al. 2010) and SIFT (Ng and Henikoff 2001) as well as GERP conservation scores (Cooper et al. 2005) in the PGP genomes.

<sup>3</sup> Distance in bp between pairs of amino acid (AA) exchanges that reside in *cis* configurations compared to those in *trans* configurations.

<sup>4</sup> Distance in bp between pairs of synonymous SNPs (sSNPs) that reside in *cis* configurations compared to those in *trans* configurations.

**Table S5. Number of phase-sensitive genes and their *cis* and *trans* configurations****(A) Genes with configurations of protein-altering mutations<sup>1</sup>**

Samples	No. phased genomes	No. phase-sensitive genes <sup>2</sup> per genome <sup>3,4</sup>	Min – Max	No. genes with <i>cis</i> configs <sup>3</sup> per genome <sup>3</sup>	Min – Max	No. genes with <i>trans</i> configs <sup>3</sup> per genome <sup>3</sup>	Min – Max
Total	1,092	487	393 - 710	297	221 - 397	193	132 - 342
EUR	379	484	428 - 554	296	258 - 353	187	145 - 228
EAS	286	449	393 - 505	269	221 - 307	180	147 - 223
AMR	181	510	393 - 613	306	224 - 374	201	132 - 287
AFR	246	647	436 - 710	354	275 - 397	292	161 - 342

<sup>1</sup> Protein-altering mutations: equivalent to (potentially) damaging mutations from the 1000 Genomes annotation database (Abecasis et al. 2012).

<sup>2</sup> Defined by presence of  $\geq 2$  protein-altering mutations.

<sup>3</sup> Data represent median values.

<sup>4</sup> Numbers are not equivalent to the sum of genes with *cis* and *trans* configurations, because median values are used. Configs, configurations.

**(B) Genes with configurations of amino acid (AA) exchanges<sup>1</sup>**

Samples	No. phased genomes	No. phase-sensitive genes <sup>2</sup> per genome <sup>3,4</sup>	Min – Max	No. genes with <i>cis</i> configs <sup>3</sup> per genome <sup>3</sup>	Min – Max	No. genes with <i>trans</i> configs <sup>3</sup> per genome <sup>3</sup>	Min – Max
Total	1,092	1,317	1,025 - 1,858	796	607 - 1,021	521	402 - 885
EUR	379	1,309	1,195 - 1,437	795	734 - 886	514	452 - 594
EAS	286	1,201	1,044 - 1,298	722	610 - 789	476	419 - 560
AMR	181	1,363	1,025 - 1,664	817	607 - 963	551	402 - 756
AFR	246	1,711	1,210 - 1,858	918	765 - 1,021	793	445 - 885

<sup>1</sup> AA exchanges from the 1000 Genomes annotation database (Abecasis et al. 2012).

<sup>2</sup> Defined by presence of  $\geq 2$  AA exchanges.

<sup>3</sup> Data represent median values.

<sup>4</sup> Numbers are not necessarily equivalent to the sum of genes with *cis* and *trans* configurations, because median values are used. Configs, configurations.

**Table S6. From mutations to genes to phase configurations****(A) From protein-altering mutations to genes to phase configurations**

Population samples	No. phased genomes	(a) No. mutations <sup>1</sup> per genome <sup>2</sup>	(b) No. genes with $\geq 1$ mutation per genome	Quot (b)/(a)	(c) No. genes with $=1$ mutation per genome	Quot (c)/(a) <sup>3</sup>	(d) No. genes with $\geq 2$ mutations <sup>4,5</sup> per genome	Quot (d)/(a)	(e) No. genes with <i>cis</i> configs per genome	Quot (e)/(a)	(f) No. genes with <i>trans</i> configs per genome	Quot (f)/(a)
Total	1,092	2,869	2,023	0.71	1,541	0.54	487	0.17	297	0.10	193	0.07
EUR	379	2,847	2,015	0.71	1,532	0.54	484	0.17	296	0.10	187	0.07
EAS	286	2,656	1,872	0.71	1,422	0.54	449	0.17	269	0.10	180	0.07
AMR	181	2,956	2,088	0.71	1,576	0.53	510	0.17	306	0.10	201	0.07
AFR	246	3,672	2,557	0.70	1,907	0.52	647	0.18	354	0.10	292	0.08

<sup>1</sup> Protein-altering, used synonymously for (potentially) damaging, mutations from 1000 Genomes annotation database (Abecasis et al. 2012).

<sup>2</sup> Mutations within exonic sequences of the autosomal protein-coding genes; calculated from a total of 3,280,032 mutations in 1,092 genomes, 1,082,993 in EUR, 759,570 in EAS, 534,596 in AMR, and 902,873 in AFR.

<sup>3</sup> Equivalent to the fraction of mutations that are contained in this gene category, see Supplemental Table S2B.

<sup>4</sup> Numbers are not equivalent to the sum of genes with *cis* and *trans* configurations, because median values are used. Essentially all data represent median values.

<sup>5</sup> Genes with  $\geq 2$  mutations also defined as 'phase-sensitive' genes.

Quot, quotient; configs, configurations.

**(B) Distribution of protein-altering mutations among gene categories**

Population samples	No. phased genomes	No. mutations <sup>1</sup> per genome <sup>2</sup>	Mutations in genes with 1 mutation (%) <sup>3</sup>	Mutations in genes with $\geq 2$ mutations (%) <sup>4</sup>	No. mutations per gene with $\geq 2$ mutations <sup>5</sup>	Mutations in genes with <i>cis</i> configs (%) <sup>6,8</sup>	Mutations in genes with <i>trans</i> configs (%) <sup>7,8</sup>
Total	1,092	2,869	53.24	46.76	2.73	24.81	21.95
EUR	379	2,847	53.96	46.04	2.72	25.66	20.38
EAS	286	2,656	53.8	46.2	2.75	24.66	21.54
AMR	181	2,956	53.07	46.93	2.71	25.55	21.38
AFR	246	3,672	52.02	47.98	2.73	23.47	24.51

<sup>1</sup> Protein-altering, used synonymously for (potentially) damaging, mutations from 1000 Genomes annotation database (Abecasis et al. 2012).

<sup>2</sup> Mutations within exonic sequences of the autosomal protein-coding genes.

<sup>3</sup> Number of genes with one mutation (equivalent to the number of mutations contained in this gene category) divided by total number of mutations per genome.

<sup>4</sup> Number of mutations contained in genes with  $\geq 2$  mutations divided by total number of mutations per genome.

<sup>5</sup> Number of mutations contained in genes with  $\geq 2$  mutations divided by number of genes with  $\geq 2$  mutations per genome (see Supplemental Table S6A).

<sup>6</sup> Number of mutations contained in *cis* configurations divided by number of mutations per genome (see Supplemental Table S6A).

<sup>7</sup> Number of mutations contained in *trans* configurations divided by number of mutations per genome (see Supplemental Table S6A).

<sup>8</sup> Numbers are not necessarily equivalent to the sum of mutations contained in genes with  $\geq 2$  mutations, because median values are used.

### (C) From amino acid (AA) exchanges to genes to phase configurations

Population samples	No. phased genomes	(a) No. mutations <sup>1</sup> per genome <sup>2</sup>	(b) No. genes with $\geq 1$ mutation per genome	Quot (b)/(a) <sup>3</sup>	(c) No. genes with =1 mutation per genome	Quot (c)/(a) <sup>3</sup>	(d) No. genes with $\geq 2$ mutations per genome <sup>4, 5</sup>	Quot (d)/(a)	(e) No. genes with <i>cis</i> configs per genome	Quot (e)/(a)	(f) No. genes with <i>trans</i> configs per genome	Quot (f)/(a)
Total	1,092	6,456	3,771	0.58	2,456	0.38	1,317	0.20	796	0.12	521	0.08
EUR	379	6,429	3,759	0.59	2,447	0.38	1,309	0.20	795	0.12	514	0.08
EAS	286	5,972	3,483	0.58	2,283	0.38	1,201	0.20	722	0.12	476	0.08
AMR	181	6,676	3,876	0.58	2,511	0.38	1,363	0.20	817	0.12	551	0.08
AFR	246	8,229	4,689	0.57	2,974	0.36	1,711	0.21	918	0.11	793	0.10

<sup>1</sup> Amino acid (AA) exchanges from 1000 Genomes annotation database (Abecasis et al. 2012).

<sup>2</sup> AA exchanges within exonic sequences of the autosomal protein-coding genes.

<sup>3</sup> Equivalent to the fraction of AA exchanges that are contained in this gene category, see Supplemental Table S6D.

<sup>4</sup> Numbers are not equivalent to the sum of genes with *cis* and *trans* configurations, because median values are used. Essentially all data represent median values.

<sup>5</sup> Genes with  $\geq 2$  AA exchanges, also defined as 'phase-sensitive' genes.

Quot, quotient; configs, configurations.

### (D) Distribution of AA exchanges among gene categories

Population samples	No. phased genomes	No. mutations <sup>1</sup> per genome <sup>2</sup>	Mutations in genes with 1 mutation (%) <sup>3</sup>	Mutations in genes with $\geq 2$ mutations (%) <sup>4</sup>	No. mutations per gene with $\geq 2$ mutations <sup>5</sup>	Mutations in genes with <i>cis</i> configs (%) <sup>6,8</sup>	Mutations in genes with <i>trans</i> configs (%) <sup>7,8</sup>
<b>Total</b>	1,092	6,456	37.57	62.43	3.20	31.32	31.11
<b>EUR</b>	379	6,429	38.13	61.87	3.04	32.60	29.27
<b>EAS</b>	286	5,972	38.34	61.66	3.07	31.78	29.88
<b>AMR</b>	181	6,676	37.58	62.42	3.06	32.06	30.36
<b>AFR</b>	246	8,229	36.25	63.75	3.06	28.93	34.82

<sup>1</sup> AA exchanges from 1000 Genomes annotation database (Abecasis et al. 2012).

<sup>2</sup> Mutations within exonic sequences of autosomal protein-coding genes.

<sup>3</sup> Number of genes with one AA exchange (equivalent to the number of AA exchanges contained in this gene category) divided by total number of AA exchanges per genome.

<sup>4</sup> Number of AA exchanges contained in genes with  $\geq 2$  AA exchanges, divided by total number of AA exchanges per genome.

<sup>5</sup> Number of AA exchanges contained in genes with  $\geq 2$  AA exchanges divided by number of genes with  $\geq 2$  AA exchanges per genome (see Supplemental Table S6A).

<sup>6</sup> Number of AA exchanges contained in *cis* configurations divided by number of AA exchanges per genome (see Supplemental Table S6A).

<sup>7</sup> Number of AA exchanges contained in *trans* configurations divided by number of AA exchanges per genome (see Supplemental Table S6A).

<sup>8</sup> Numbers are not necessarily equivalent to the sum of mutations contained in genes with  $\geq 2$  AA exchanges, because median values are used.



**Table S7. Simulation of proportionality constants (quotients)**

	No. phased genomes	(a) No. mutations per genome <sup>1</sup>	(b) No. genes with $\geq 1$ mutation per genome	Quot (b)/(a)	(c) No. genes with $= 1$ mutation per genome	Quot (c)/(a)	(d) No. genes with $\geq 2$ mutations per genome <sup>4,5</sup>	Quot (d)/(a)	(e) No. genes with <i>cis</i> configs per genome	Quot (e)/(a)	(f) No. genes with <i>trans</i> configs per genome	Quot (f)/(a)
<b>SIMULATED DATA</b>												
Protein-altering	1,092	3,102	2,119	0.68	1,519	0.49	597	0.19	237	0.08	363	0.12
AA exchanges	1,092	6,855	4,023	0.59	2,543	0.37	1,478	0.22	540	0.08	936	0.14
<b>REAL DATA</b>												
Protein-altering	1,092	2,869	2,023	0.71	1,541	0.54	487.0	0.17	297.0	0.10	193.0	0.07
AA exchanges	1,092	6,456	3,771	0.58	2,456	0.38	1,317	0.2	796	0.12	521	0.08

The generation of virtual sets of 1,092 phased genomes, the analyses of which are presented under 'SIMULATED DATA' is described in Methods; the data presented in the section 'REAL DATA' are excerpts from Supplemental Tables S6A, C. See also legends of these tables for further detail. See also Supplemental Note for approximation of proportionality constants.

<sup>1</sup> For each virtual genome, random numbers of protein-altering mutations, equivalent to (potentially) damaging mutations, were drawn in the range observed in the 1,092 genomes dataset (~2,500–3,500) from the 1000 Genomes (1000G) database (Abecasis et al. 2012). These mutations were sampled from the total of ~300,000 protein-altering mutations annotated in this data set. To generate a second virtual set of 1,092 phased genomes from the entirety of amino acid (AA) exchanges, between ~5,500 and ~7,500 nsSNPs were drawn randomly from the entire pool of ~1.5 Mio nsSNPs annotated in the 1,092 'real' genomes. Importantly, phase was simulated assigning to every single mutation drawn a 50:50 chance to exist on either homologue 1 or 2 (see also Methods). With the exception of the quotients, all numbers of mutations per genome in both the simulated and 'real data' represent median values. Quot, quotient; configs, configurations.

**Table S14. *Cis* configurations per autosome in 1,092 genomes and different ancestry-based groups  
(A) *Cis* configurations of protein-altering mutations**

	<i>Cis</i> configurations (%) <sup>1</sup>				
	1,092	EUR	EAS	AMR	AFR
chr1	63.1	65.6	64.3	63.6	55.1
chr2	58.6	59.1	60.8	58.1	56.3
chr3	60.0	61.9	61.1	62.1	53.8
chr4	62.5	64.7	66.7	62.5	52.4
chr5	60.9	65.0	60.0	61.9	55.2
chr6	53.1	56.0	52.6	54.2	48.8
chr7	60.0	62.5	60.0	60.0	55.7
chr8	52.9	53.8	53.8	53.3	49.0
chr9	55.9	59.1	52.5	59.1	52.8
chr10	52.6	54.5	50.0	54.2	52.0
chr11	60.6	60.0	62.2	60.9	60.0
chr12	54.5	57.1	50.0	52.9	54.5
chr13	62.5	66.7	66.7	66.7	57.1
chr14	68.8	71.4	73.5	68.8	61.5
chr15	56.3	55.6	61.5	57.1	50.9
chr16	61.9	66.7	60.0	65.0	55.4
chr17	57.6	59.4	58.5	57.6	55.9
chr18	55.6	50.0	57.1	55.6	55.6
chr19	60.6	61.9	61.1	62.0	58.3
chr20	64.3	68.8	66.7	64.3	56.8
chr21	60.0	66.7	60.0	62.5	50.0
chr22	66.7	71.4	62.5	70.0	50.0
Median	60.0	61.9	60.4	61.4	55.2
Min	52.6	50.0	50.0	52.9	48.8
Max	68.8	71.4	73.5	70.0	61.5

<sup>1</sup> Values represent the median of *cis* fractions (%) per autosome per genome; these were assessed as the number of *cis* configurations observed across all autosomal protein-coding genes contained in designated autosome, divided by total configuration count per autosome.

**(B) Cis configurations of AA exchanges**

	<b>Cis configurations (%)<sup>1</sup></b>				
	1,092	EUR	EAS	AMR	AFR
chr1	58.5	60.2	59.8	58.5	52.8
chr2	56.9	57.7	58.1	57.3	53.5
chr3	63.1	65.8	64.9	64.4	54.6
chr4	60.0	61.7	60.0	60.0	56.2
chr5	59.3	61.9	60.0	59.6	52.2
chr6	54.5	56.4	54.8	55.2	50.5
chr7	58.3	60.6	59.2	58.9	52.7
chr8	57.9	58.5	63.6	57.9	50.8
chr9	54.8	56.5	54.3	56.9	51.8
chr10	57.1	59.2	57.1	58.0	53.5
chr11	60.3	60.5	62.0	60.6	57.7
chr12	57.1	59.3	57.1	57.4	53.2
chr13	59.1	62.5	58.3	60.0	53.8
chr14	60.2	61.7	61.8	61.0	55.5
chr15	55.8	57.6	57.9	56.8	47.8
chr16	56.4	58.9	54.8	58.0	52.9
chr17	58.9	59.4	62.1	58.5	54.2
chr18	47.8	47.8	47.4	47.6	50.0
chr19	67.3	68.7	69.6	67.7	57.9
chr20	63.3	66.7	63.8	63.3	57.1
chr21	53.3	56.3	55.0	56.3	47.2
chr22	60.0	64.3	60.0	64.3	49.0
Median	58.4	59.8	59.5	58.5	53.1
Min	47.8	47.8	47.4	47.6	47.2
Max	67.3	68.7	69.6	67.7	57.9

<sup>1</sup> Values represent the median of *cis* fractions (%) per autosome per genome; these were assessed as the number of *cis* configurations observed across all autosomal protein-coding genes contained in designated autosome, divided by total configuration count per autosome.