****

**Figure S1.** Numbers of substitutions that give rise to the human amino acid variant (white) or the mean number of substitutions to a particular amino acid (black) at a site for sites with at least one substitution in Metazoan dataset.



**Figure S2.** Numbers of substitutions that give rise to the human amino acid variant (white) or the mean number of substitutions to a particular amino acid (black) at a site for sites with at least one substitution in opisthokonts dataset.



**Figure S3.** Ratios of the phylogenetic distances between the human branch and substitutions to the human reference allele vs. to other amino acids, for metazoan dataset. Ratios <1 imply that the considered allele arises independently closer at the phylogeny to humans than other alleles. The bar height and the error bars represent respectively the median and the 95% confidence intervals obtained from 1,000 bootstrap replicates, and asterisks show the significance of difference from the one-to-one ratio (\*, P<0.05; \*\*, P<0.01;\*\*\*, P<0.001). ref, human reference allele; access, human reference alleles from accessible amino acid pairs (see text); rand, a random non-human amino acid among those that were present in the site; sim, human allele in simulated data.

****

**Figure S4.** Higher fraction of homoplasic substitutions to the human reference amino acid, compared with random amino acids that had independently originated at this site (H/D ratio), in species closely related to human, for metazoan dataset. Horizontal axis, distance between branches carrying the substitutions and the human branch, measured in numbers of amino acid substitutions per site, split into bins by log2(distance). Vertical axis, H/D ratios for substitutions at this distance. Black line, mean; grey confidence band, 95% confidence interval obtained from 1000 bootstrapping replicates. The red line shows the expected H/D ratio of 1. Arrows represent the distance between human and *Drosophila.*

****

**Figure S5.** Ratios of the phylogenetic distances between the human branch and substitutions to the human non-reference allele vs. to other amino acids, for metazoan dataset. Notations same as in Figure S1.

****

**Figure S6.** Higher fraction of homoplasic substitutions to the human non-reference amino acid, compared with random amino acids that had independently originated at this site (H/D ratio), in species closely related to human, for metazoan dataset. Notations same as in Figure S2.

****

**Figure S7.** Ratios of the phylogenetic distances between the human branch and substitutions to the human pathogenic allele vs. to other amino acids, for metazoan dataset. Notations same as in Figure S1.

****

**Figure S8.** Higher fraction of homoplasic substitutions to the human pathogenic amino acid, compared with random amino acids that had independently originated at this site (H/D ratio), in species closely related to human, for metazoan dataset. Notations same as in Figure S2.



**Figure S9.** Substitutions giving rise to the human pathogenic allele (red circles) and to other amino acids (green circles) on the 4350-species phylogeny. a, site 156 of ATP6; b, site 217 of ATP6; c, site 220 of ATP6; d, site 278 of CYTB; e, site 35 of CYTB; f,site 40 of CYTB. Blue star, *H. sapiens*

**Table S1.** Homoplasies giving rise to the human reference allele.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gene** | **Species** | **Sites** | **Analyzed sites** | **Substitutions to the human amino acid per site** | **Substitutions to the human amino acid per site in simulation** |
| ATP6 | Metazoans | 186 | 131 | 25.3 | 16.7 |
| COX1 | Metazoans | 404 | 146 | 12.1 | 13.9 |
| COX2 | Metazoans | 165 | 110 | 22.5 | 22.3 |
| COX3 | Metazoans | 198 | 128 | 22.4 | 16.2 |
| CYTB | Metazoans | 327 | 183 | 32.2 | 29.7 |
| ND1 | Metazoans | 253 | 160 | 15.5 | 12.1 |
| ND2 | Metazoans | 299 | 220 | 36.3 | 34.0 |
| ND3 | Metazoans | 94 | 67 | 26.6 | 20.0 |
| ND4 | Metazoans | 392 | 263 | 19.9 | 17.3 |
| ND4L | Metazoans | 82 | 70 | 18.4 | 19.6 |
| ND5 | Metazoans | 516 | 302 | 8.7 | 7.3 |
| ND6 | Metazoans | 119 | 91 | 11.4 | 10.1 |
| ATP6+COX1+COX2+COX3+CYTB | Opisthokonts | 1524 | 964 | 30.2 | 24.4 |

**Table S2.** Homoplasies giving rise to the human non-reference allele.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gene** | **Species** | **Polymorphic sites** | **Analyzed sites** | **Substitutions to the human non-reference allele per non-reference allele** | **Mean no. of non-reference alleles per polymorphic site** |
| ATP6 | Metazoans | 142 | 12 | 21.1 | 1.9 |
| COX1 | Metazoans | 131 | 10 | 21.0 | 1.4 |
| COX2 | Metazoans | 81 | 72 | 28.3 | 1.4 |
| COX3 | Metazoans | 108 | 93 | 24.8 | 1.6 |
| CYTB | Metazoans | 199 | 18 | 36.0 | 1.7 |
| ND1 | Metazoans | 101 | 88 | 16.8 | 1.5 |
| ND2 | Metazoans | 143 | 64 | 49.1 | 1.6 |
| ND3 | Metazoans | 34 | 32 | 30.3 | 1.6 |
| ND4 | Metazoans | 136 | 10 | 20.9 | 1.2 |
| ND4L | Metazoans | 29 | 26 | 17.3 | 1.5 |
| ND5 | Metazoans | 223 | 17 | 9.9 | 1.6 |
| ND6 | Metazoans | 55 | 45 | 11.1 | 1.5 |
| ATP6+COX1+COX2+COX3+CYTB | Opisthokonts | 775 | 516 | 32.2 | 1.6 |

**Table S3.** Homoplasies giving rise to the human pathogenic variants.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gene** | **Species** | **Sites with pathogenic alleles** | **Analyzed sites** | **Substitutions to****the human pathogenic amino acid per pathogenic amino acid**  | **Mean no. of pathogenic amino acids per site with pathogenic alleles** |
| ATP6 | Metazoans | 15 | 11 | 9.7 | 1.1 |
| COX1 | Metazoans | 20 | 15 | 41.0 | 1.0 |
| COX2 | Metazoans | 10 | 9 | 44.8 | 1.0 |
| COX3 | Metazoans | 9 | 7 | 14.8 | 1.0 |
| CYTB | Metazoans | 20 | 16 | 15.5 | 1.1 |
| ND1 | Metazoans | 25 | 17 | 6.3 | 1.1 |
| ND2 | Metazoans | 10 | 8 | 29.5 | 1.0 |
| ND3 | Metazoans | 6 | 5 | 3.5 | 1.0 |
| ND4 | Metazoans | 7 | 4 | 15.1 | 1.2 |
| ND4L | Metazoans | 3 | 3 | 6.0 | 1.0 |
| ND5 | Metazoans | 24 | 8 | 5.1 | 1.0 |
| ND6 | Metazoans | 13 | 10 | 8.0 | 1.1 |
| ATP6+COX1+COX2+ COX3+CYTB | Opisthokonts  | 89 | 72 | 25.0 | 1.0 |