

cWebsite for getting allele frequencies: [https://useast.ensembl.org/Homo\\_sapiens/Info/Index](https://useast.ensembl.org/Homo_sapiens/Info/Index)  
Website for finding SNPs: <https://www.snpedia.com/index.php/Genotype>,  
*a few studies to get them from too:*

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0079522> (good for MSEA populations), or <https://pmc.ncbi.nlm.nih.gov/articles/PMC2775604/#B37> singapore thing,  
<https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1010036>

More studies:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC12221367/#:~:text=Southeast%20Asia%E2%80%93ome%20to%20nearly,both%20within%20and%20across%20borders.>,

<https://pmc.ncbi.nlm.nih.gov/articles/PMC4480199/#:~:text=After%20data%20filtering%2C%20463%2C265%20overlapping,are%20presented%20in%20Supplement%201.6>.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC3626844/#:~:text=Figure%206.&text=Phylogenetic%20network%20of%20haplotypes%20of,D%20indicates%20%2D%2DSEA%20allele>.

<https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms-SNPs>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC3525085/>

<https://pubmed.ncbi.nlm.nih.gov/40369069/>

[https://reich.hms.harvard.edu/sites/reich.hms.harvard.edu/files/inline-files/Chapter25\\_OHAL\\_AustronesianArchaeolinguistics.pdf](https://reich.hms.harvard.edu/sites/reich.hms.harvard.edu/files/inline-files/Chapter25_OHAL_AustronesianArchaeolinguistics.pdf)

<https://pmc.ncbi.nlm.nih.gov/articles/PMC3437904/>

<https://www.ecns.cn/news/2023-09-05/detail-ihcscvcpq5395055.shtml>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC7268520/#:~:text=Conclusion,genotypes%20for%20non%2Dmodel%20organisms>.

[https://www.researchgate.net/figure/A-HPLC-profile-of-the-retained-material-obtained-by-immunoaffinity-purification-using\\_fig3\\_263585805](https://www.researchgate.net/figure/A-HPLC-profile-of-the-retained-material-obtained-by-immunoaffinity-purification-using_fig3_263585805)

[https://resources.qiagenbioinformatics.com/manuals/clcsequenceviewer/current/index.php?manual=Scale\\_bar.html](https://resources.qiagenbioinformatics.com/manuals/clcsequenceviewer/current/index.php?manual=Scale_bar.html)

<https://pmc.ncbi.nlm.nih.gov/articles/PMC3126721/>

<https://pmc.ncbi.nlm.nih.gov/articles/PMC7475039/#sec9>

<https://www.digitalatlasofancientlife.org/learn/systematics/phylogenetics/reading-trees/#:~:text=Trees%20may%20also%20have%20branch,went%20extinct%2C%20as%20shown%20below>.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC8970429/>

Create phylogenetic tree in POPTREE2 ([med.kagawa-uc.ac.jp](http://med.kagawa-uc.ac.jp)) by Naoko Takezaki, Masatoshi Nei, and Koichiro Tamura

Legend:

CDX = Dai from Xishuangbanna - Tai Kadai people are an ethnic group in Southwestern China, Laos, Thailand, parts of Cambodia, Myanmar, Vietnam, and India

KHV = Kinh from Ho Chi Minh City, Vietnam - The Kinh people are the primary ethnic group of Vietnam, making up over 80% of the country's population, also residing in parts of Cambodia and the deep south of China

BEB = Bengali in Bangladesh - Major Indo-Aryan ethnic group making up most of Bangladesh and West Bengal in India

STU = Sri Lankan Tamil - Dravidian group residing in Sri Lanka, Southern India, Malaysia and Singapore

CHS = Southern Han Chinese - Cantonese speaking Han Chinese ethnic group of provinces such as Yunnan, Guangdong, etc, along with Hong Kong

CHB = Han Chinese in Beijing - Han Chinese of northern China

JPT = Japanese in Tokyo, Japan - Yamato ethnic group of Japan

HGDP Cambodian = Cambodian in Cambodia

Outlier group to amplify Asia-Pacific cluster - YRI = Yoruba in Ibadan, Nigeria

Locus 1 Done SNP = rs1128503 - Alleles: A, G, Ancestral: G

CDX = A: 62% G: 38%, pop: 186, KHV = A: 58%, G: 42%, pop: 198, BEB = A: 63% , G: 37%, pop: 172, STU = A: 62%, G: 38%, pop: 204, CHS = A: 63%, G: 37%, pop: 210, CHB = A: 70%, G: 30%, pop: 206 JPT = A: 60%, G: 40%, pop: 208, HGDP = A: 58%, G: 42%, pop: 22, YRI = A: 14%, G: 86%, pop: 216

2 Done SNP = rs17822931 - Alleles: C, T, Ancestral: C

CDX = C: 46%, T: 54%, pop: 186, KHV = C: 36%, T: 64%, pop: 198, BEB = C: 45%, T: 55%, pop: 172, STU = C: 47%, T: 53%, pop: 194, CHS = C: 16%, T: 84%, pop: 210, CHB = C: 3%, T: 97%, pop: 206, JPT = C: 12%, T: 88%, pop: 208, HGDP = C: 40%, T: 60%, YRI = C: 100%, T: 0%, pop: 216

3 Done SNP = rs1426654 - Alleles: A, G, Ancestral: G

CDX = A: 0% G: 100% pop: 186, KHV = A: 1%, G: 99% pop: 198, BEB = A: 53%, G: 47% pop: 172, STU = A: 49%, G: 51%, pop: 204, CHS = A: 2%, G: 98%, pop: 210, CHB = A: 3%, G: 97%, pop: 206, JPT = A: 0%, G: 100%, pop: 208, HGDP = A: 5%, G: 95%, YRI = A: 1%, G: 99%, pop: 216

4 Done SNP = rs16891982 - Alleles: C, G,

CDX = C: 100% G: 0%, pop: 186, KHV = C: 99%, G: 1%, BEB = C: 96%, G: 4%, pop: 172, STU = C: 94%, G: 6%, pop: 204, CHS = C: 100%, G: 0%, pop: 210, CHB = C: 99%, G: 1%, pop: 206, JPT = C: 100%, G: 0%, pop: 208, HGDP = C: 100%, G: 0%, YRI = C: 100%, G: 0%, pop: 216

6 Done SNP = rs7754840 - Alleles: C, G, Ancestral:

CDX = C: 63%, G: 37%, pop: 186, KHV = C: 57%, G: 43%, pop: 198, BEB = C: 65%, G: 35%, pop: 172, STU = C: 79%, G: 21%, pop: 204, CHS = C: 62%, G: 38%, pop: 210, CHB = C: 59%, G: 41%, pop: 206, JPT = C: 59%, G: 41%, pop: 208, HGDP: C: 70%, G: 30%, YRI = C: 33%, G: 67%, pop: 216

7 Done SNP = rs4778220 - Alleles: A, G, Ancestral: G

CDX = A: 99% G: 1% pop: 186, KHV = A: 100%, G: 0% pop: 198, BEB = A: 59%, G: 41% pop: 172, STU = A: 52%, G: 48%, pop: 204, CHS = A: 100%, G: 0%, pop: 210, CHB = A: 99%, G: 1%, pop: 206, JPT = A: 99%, G: 1%, pop: 208, HGDP = A: 95%, G: 5%, YRI = A: 47%, G: 53%, pop: 216

8 Done SNP = rs987870 - Alleles: A, G, Ancestral: C

CDX = A: 92%, G: 8%, pop: 186, KHV = A: 90%, G: 10%, pop: 198, BEB = A: 73%, G: 27%, pop: 172, STU = A: 71%, G: 29%, pop: 204, CHS = A: 91%, G: 9%, pop: 210, CHB = A: 90%, G: 10%, pop: 206, JPT = A: 87%, G: 13%, pop: 208, HGDP: A: 65%, G: 35%, YRI = A: 51%, G: 49%, pop: 216

9 Done SNP = rs3805322 - Alleles: A, G, Ancestral:

CDX = A: 73%, G: 27%, pop: 186, KHV = A: 68%, G: 32%, pop: 198, BEB = A: 100%, G: 0%, pop: 172, STU = A: 100%, G: 0%, pop: 204, CHS = A: 57%, G: 43%, pop: 210, CHB = A: 52%, G: 48%, pop: 206, JPT = A: 38%, G: 62%, pop: 208, HGDP = A: 90%, G: 10%, YRI = A: 100%, G: 0%, pop: 216

10 Done SNP = rs6053171 - Alleles: G, T, Ancestral: G

CDX = G: 3%, T: 97%, pop: 186, KHV = G: 2%, T: 98%, pop: 198, BEB = G: 35%, T: 65%, pop: 172, STU = G: 40%, T: 60%, pop: 204, CHS = G: 1%, T: 99%, pop: 210, CHB = G: 3%, T: 97%, pop: 216, JPT = G: 1%, T: 99%, pop: 208, HGDP = G: 2%, T: 98%, YRI = G: 62%, T: 38%

11 Done SNP = rs671 - Alleles: G, A, Ancestral: C

CDX = G: 96%, A: 4%, pop: 186, KHV = G: 86%, A: 14%, pop: 198, BEB = G: 100%, A: 0%, pop: 172, STU = G: 100%, A: 0%, pop: 204, CHS = G: 73%, A: 27%, pop: 210, CHB = G: 84%, A: 16%, pop: 216, JPT = G: 76%, A: 24%, pop: 208, HGDP = G: 75%, A: 25%, YRI = G: 100%, A: 0%, pop: 216

12 Done SNP = rs2273379 - Alleles: T, C, Ancestral: C

CDX = T: 91%, C: 9%, pop: 186, KHV = T: 91%, C: 9%, pop: 198, BEB = T: 96%, C: 4%, pop: 172, STU = T: 98%, C: 2%, pop: 204, CHS = T: 92%, C: 8%, pop: 210, CHB = T: 93%, C: 7%, pop: 206, JPT = T: 91%, C: 9%, pop: 208, HGDP = T: 91%, C: 9%, YRI = T: 82%, C: 18%, pop: 216

13 Done SNP = rs9940585 - Alleles: C, T, Ancestral: C

CDX = C: 35%, T: 65%, pop: 186, KHV = C: 31%, T: 69%, pop: 198, BEB = C: 26%, T: 74%, pop: 172, STU = C: 27%, T: 73%, pop: 204, CHS = C: 37%, T: 63%, pop: 210, CHB = C: 33%, T: 67%, pop: 206, JPT = C: 31%, T: 69%, pop: 208, HGDP = C: 31%, T: 69%, YRI = C: 86%, T: 14%, pop: 216

14 Done SNP = rs214247 - Alleles: T, C, Ancestral: C

CDX = T: 85%, C: 15%, pop: 186, KHV = T: 93%, C: 7%, pop: 198, BEB = T: 72%, C: 28%, pop: 172, STU = T: 62%, C: 38%, pop: 204, CHS = T: 81%, C: 19%, pop: 210, CHB = T: 79%, C: 21%, pop: 206, JPT = T: 82%, C: 18%, HGDP = T: 100%, C: 0%, pop: 208, YRI = T: 24%, C: 76%, pop: 216

15 Done SNP = rs77308790 - Alleles: C, T, Ancestral: C

CDX = C: 87%, T: 13%, pop: 186, KHV = C: 91%, T: 9%, pop: 198, BEB = C: 99%, T: 1%, pop: 172, STU = C: 100%, T: 0%, pop: 204, CHS = C: 90%, T: 10%, pop: 210, CHB = C: 94%, T: 6%, pop: 206, JPT = C: 94%, T: 6%, pop: 208, HGDP = C: 96%, T: 4%, YRI = C: 100%, T: 0%, pop: 216

16 Done SNP = rs3760053 - Alleles: T, G, Ancestral:

CDX = T: 84%, G: 16%, pop: 186, KHV = T: 91%, G: 9%, pop: 198, BEB = T: 99%, G: 1%, pop: 172, STU = T: 100%, G: 0%, pop: 204, CHS = T: 90%, G: 10%, pop: 210, CHB = T: 93%, G: 7%, pop: 206, JPT = T: 94%, G: 6%, pop: 208, HGDP = T: 91%, G: 9%, YRI = T: 100%, G: 0%, pop: 216

17 Done SNP = rs1211375 - Alleles: A, C, Ancestral: G

CDX = A: 60% C: 40% pop: 186, KHV = A: 48%, C: 52% pop: 198, BEB = A: 34%, C: 66% pop: 172, STU = A: 41%, C: 59%, pop: 204, CHS = A: 50%, C: 50%, pop: 210, CHB = A: 46%, C: 54%, pop: 206, JPT = A: 45%, C: 55%, pop: 208, HGDP = A: 50%, C: 50%, YRI = A: 31%, C: 69%, pop: 216

18 Done SNP = rs3918352 - Alleles: A, G, Ancestral: G

CDX = A: 72% G: 28% pop: 186, KHV = A: 68%, G: 32% pop: 198, BEB = A: 62%, G: 38% pop: 172, STU = A: 68%, G: 32%, pop: 204, CHS = A: 67%, G: 33%, pop: 210, CHB = A: 64%, G: 36%, pop: 206, JPT = A: 66%, G: 34%, pop: 208, HGDP = A: 70%, G: 30%, YRI = A: 82%, G: 18%, pop: 216

19 Done SNP = rs1203974 - Alleles: G, A, Ancestral: C

CDX = G: 50%, A: 50%, pop: 186, KHV = G: 44%, A: 56%, pop: 198, BEB = G: 33%, A: 67%, pop: 172, STU = G: 29%, A: 71%, pop: 204, CHS = G: 46%, A: 54%, pop: 210, CHB = G: 48%, A: 52%, pop: 216, JPT = G: 41%, A: 59%, pop: 208, HGDP = G: 30%, A: 70%, YRI = G: 42%, A: 58%, pop: 216

20 Done SNP = rs11248914 - Alleles: T, C, Ancestral: C

CDX = T: 40%, C: 60%, pop: 186, KHV = T: 42%, C: 58%, pop: 198, BEB = T: 49%, C: 51%, pop: 172, STU = T: 59%, C: 41%, pop: 204, CHS = T: 40%, C: 60%, pop: 210, CHB = T: 39%, C: 61%, pop: 206, JPT = T: 38%, C: 62%, pop: 208, HGDP = T: 60%, C: 40%, YRI = T: 68%, C: 32%, pop: 216

21 Done SNP = rs8997 - Alleles: A, G, Ancestral: G

CDX = A: 6% G: 94% pop: 186, KHV = A: 8%, G: 92% pop: 198, BEB = A: 2%, G: 98% pop: 172, STU = A: 2%, G: 98%, pop: 204, CHS = A: 12%, G: 88%, pop: 210, CHB = A: 5%, G: 95%, pop: 206, JPT = A: 5%, G: 95%, pop: 208, HGDP = A: 5%, G: 95%, YRI = A: 55%, G: 45%, pop: 216

22 Done SNP = rs9572312 - Alleles: C, T, Ancestral: C

CDX = C: 70%, T: 30%, pop: 186, KHV = C: 72%, T: 28%, pop: 198, BEB = C: 59%, T: 41%, pop: 172, STU = C: 55%, T: 45%, pop: 204, CHS = C: 71%, T: 29%, pop: 210, CHB = C: 74%, T: 26%, pop: 206, JPT = C: 69%, T: 31%, pop: 208, HGDP = C: 70%, T: 30%, YRI = C: 81%, T: 19%, pop: 216

23 Done SNP = rs186154 - Alleles: A, G, Ancestral: A

CDX = A: 10% G: 90% pop: 186, KHV = A: 16%, G: 84% pop: 198, BEB = A: 54%, G: 46% pop: 172, STU = A: 56%, G: 44%, pop: 204, CHS = A: 15%, G: 85%, pop: 210, CHB = A: 25%, G: 85%, pop: 206, JPT = A: 31%, G: 69%, pop: 208, HGDP = A: 35%, G: 65%, YRI = A: 84%, G: 16%, pop: 216

23 Done SNP = rs1447826 - Alleles: T, C, Ancestral: C

CDX = T: 73%, C: 27%, pop: 186, KHV = T: 71%, C: 29%, pop: 198, BEB = T: 60%, C: 40%, pop: 172, STU = T: 51%, C: 49%, pop: 204, CHS = T: 76%, C: 24%, pop: 210, CHB = T: 75%, C: 25%, pop: 206, JPT = T: 72%, C: 28%, pop: 208, YRI = T: 57%, C: 43%, pop: 216

