Population Genetics of Southeast Asian Peoples

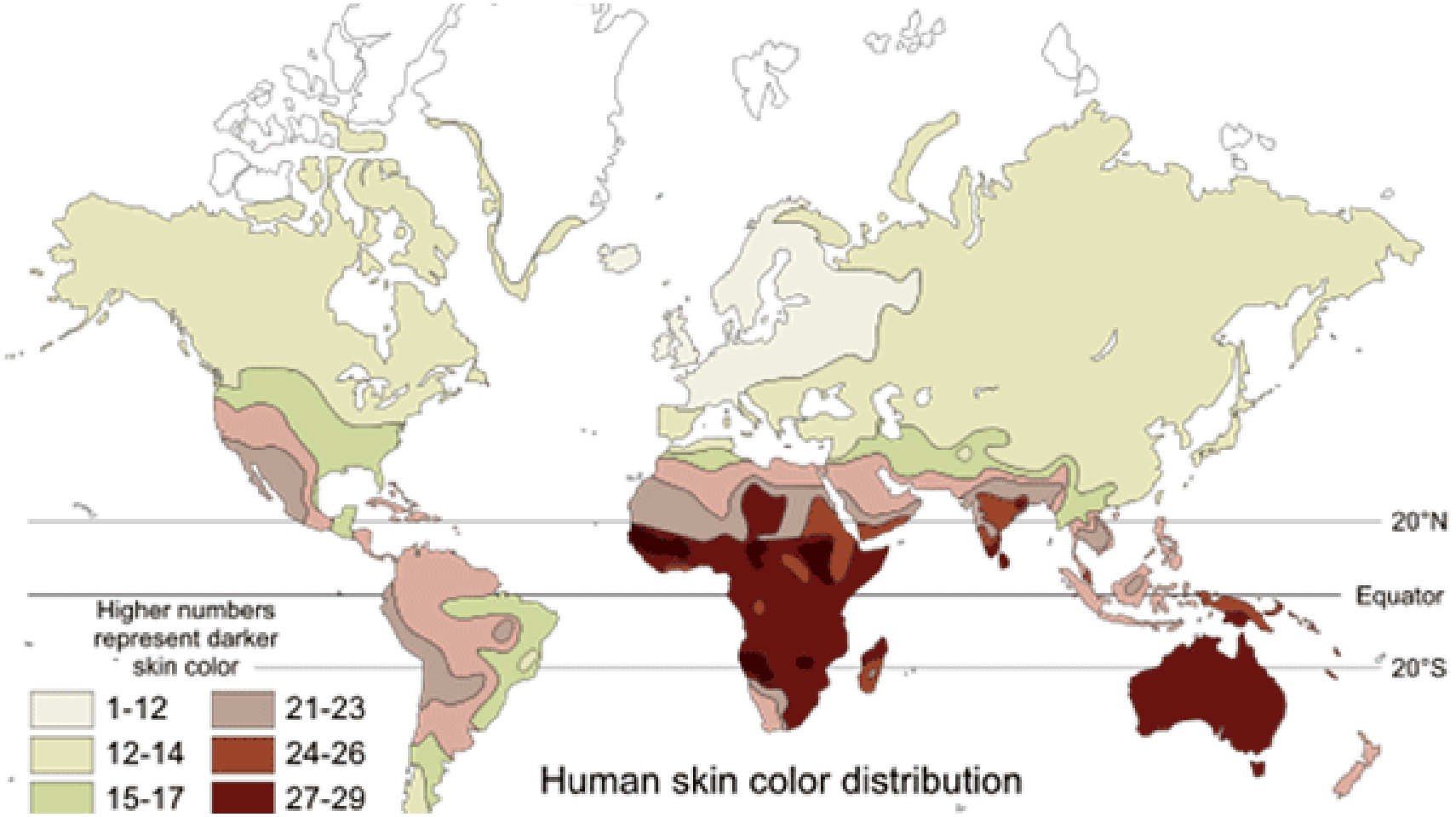
Motivation

- Genetic tells us different stories of human history from languages and ethnicity (where you come from)

Genes control many things

- Common diseases, e.g., diabetes, cancer, hypertension etc.
- Behaviors and personalities
- Physical characteristics, e.g., height, weight, body shape/figure etc.

Common trait like skin color

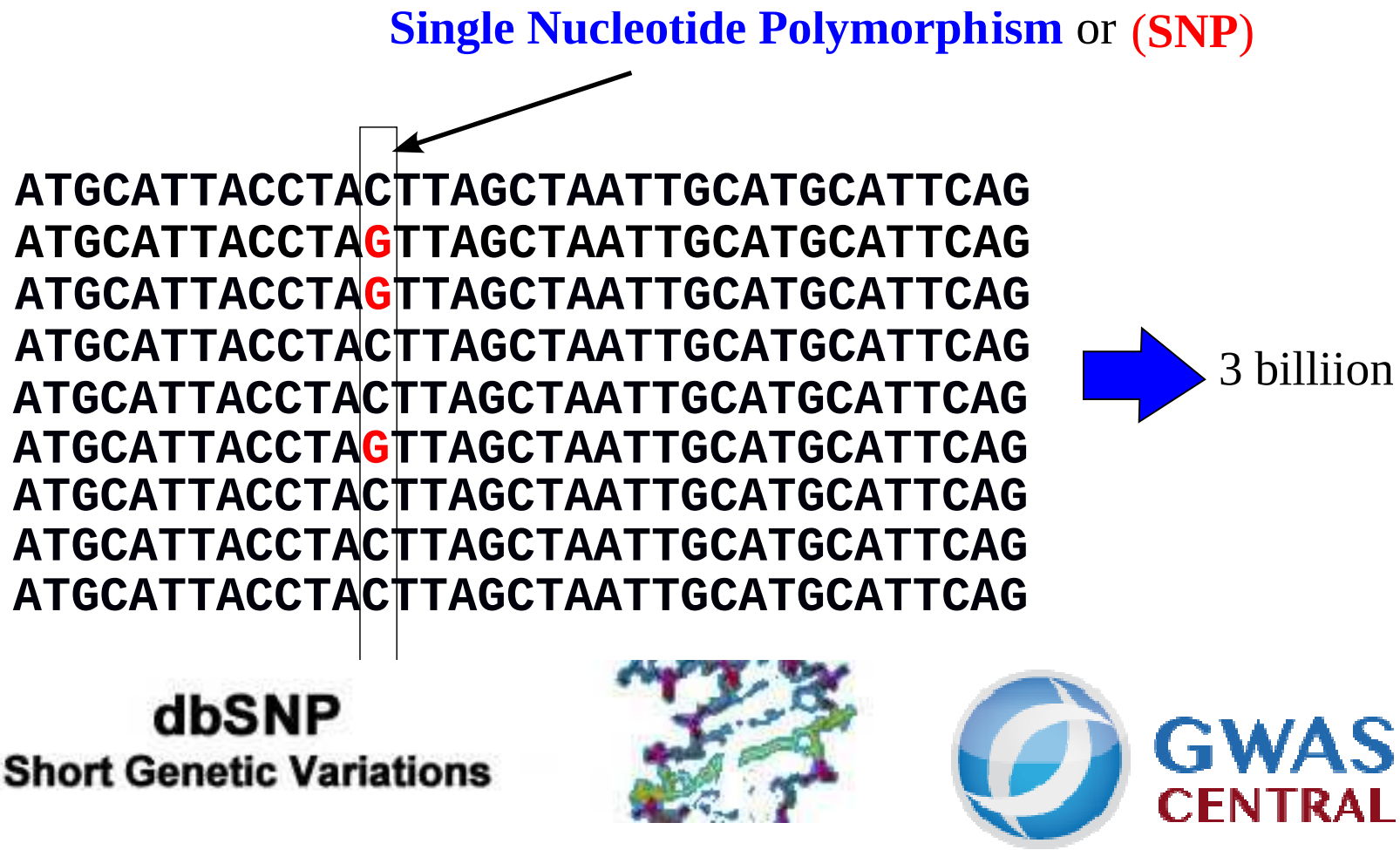


http://www.gbhealthwatch.com/Trait-Skin-Color.php

Sickle-cell trait

http://discovermagazine.com/2005/mar/human-study-thyself#.UVlDjL_hZTs

- Make up 90% of all variations
- Occur every 100-300 bases along 3 billion bp



How to get SNP data

- DNA chip up to ~2 million SNPs
- Capture the *pattern* of all SNPs
- Identify SNP patterns among group of individuals (genetic affiliation)
- Why are these patterns of genetic affiliation important?

Genetic affiliation of SEA people

- How different are SEA people from each other in the genetic sense?
- We can place boundaries according to ethnicity, geography and languages
- Medical benefit when knowing your genetic affiliation

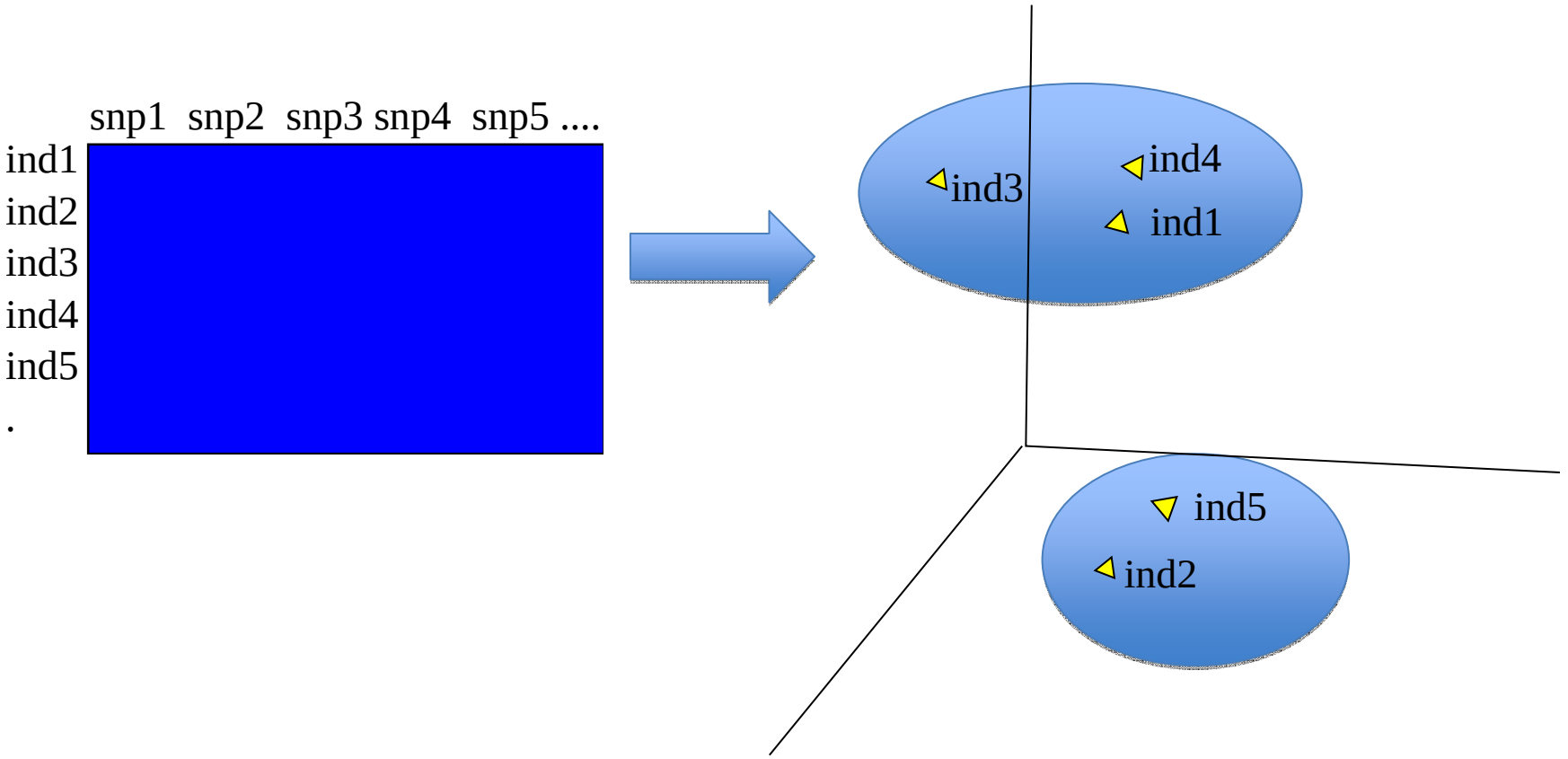
Where the data come from?

- SEA data from Vietnamese, Cambodian, Chinese, Japanese, Thais, Indonesian from Xing et al. dataset
- Thalassemia patients minor and major from Thailand
- Depressive disorder patients from Thailand

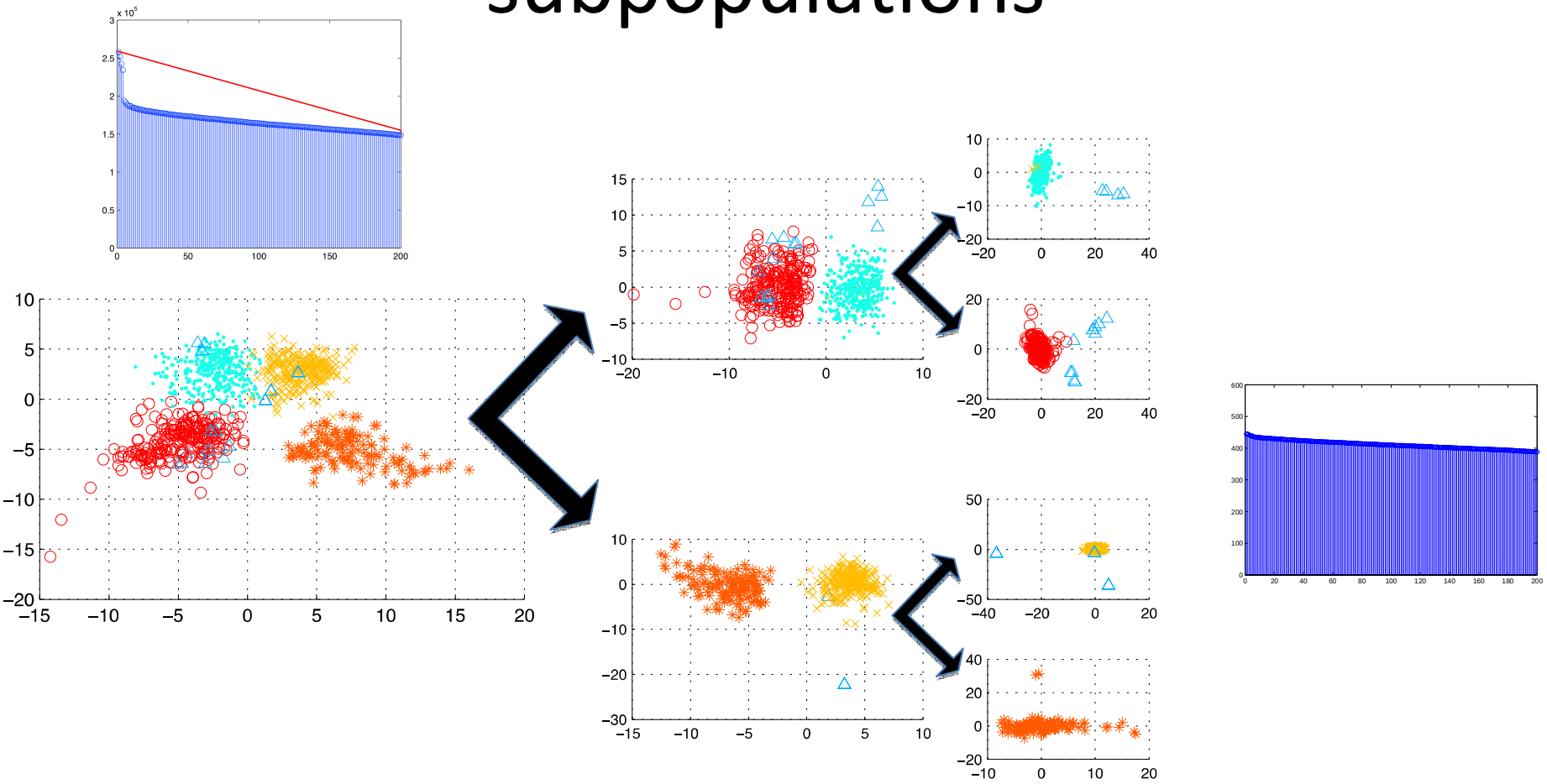
Grouping Individuals

- Based on SNP patterns
- Using PCA-based technique to group them
- Subpopulation defines a group of *unrelated* individuals in which the variation among those individuals contains nothing more than what would be expected from random

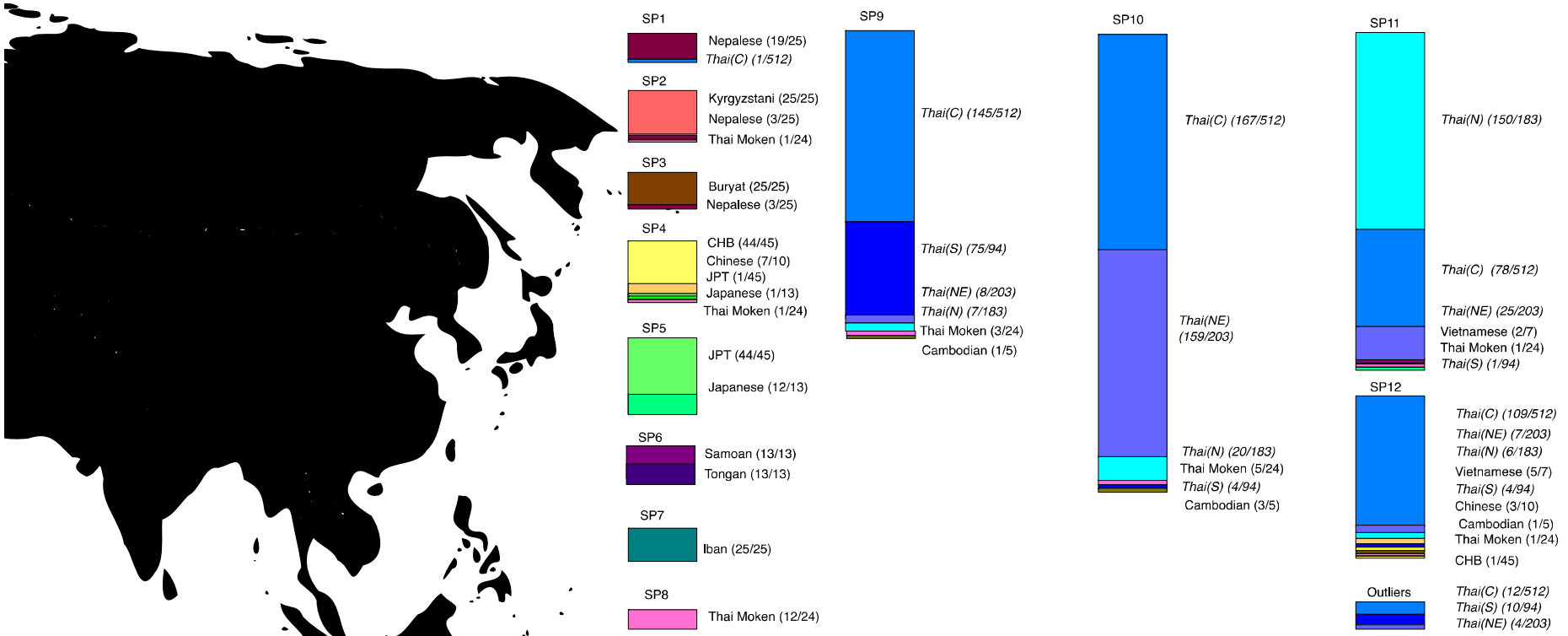
How we identify the subpopulations



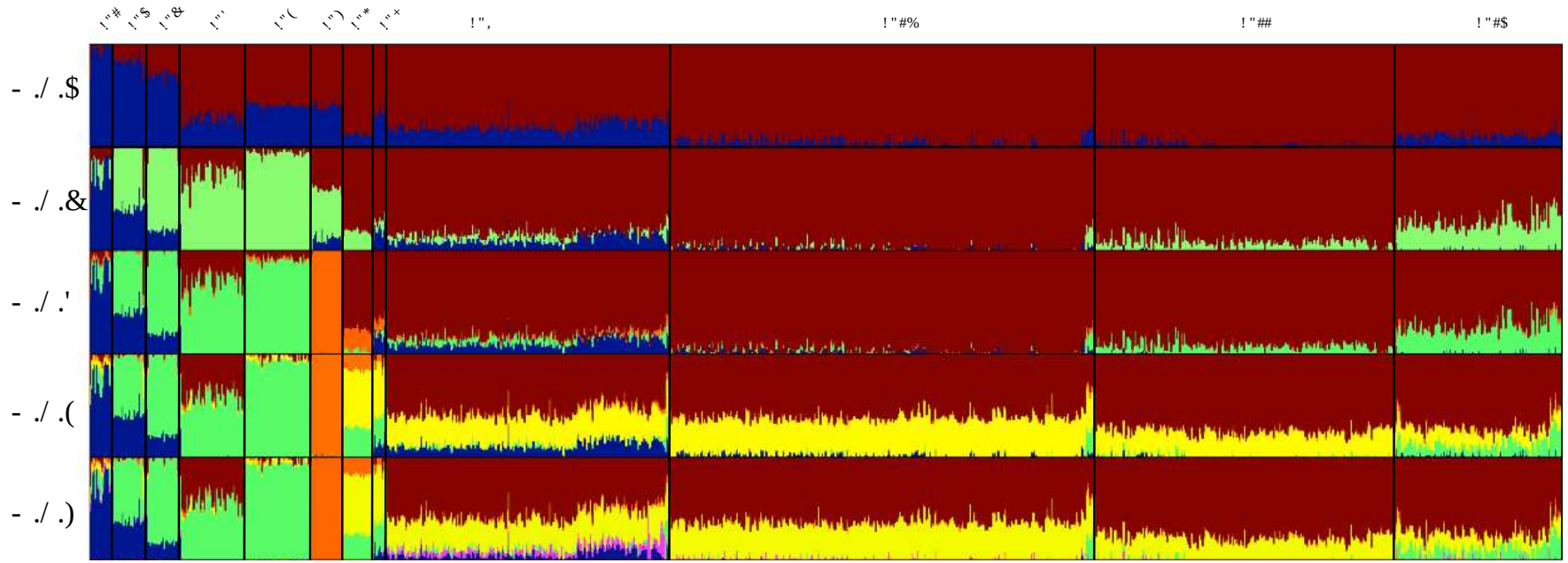
ipPCA method for resolving subpopulations



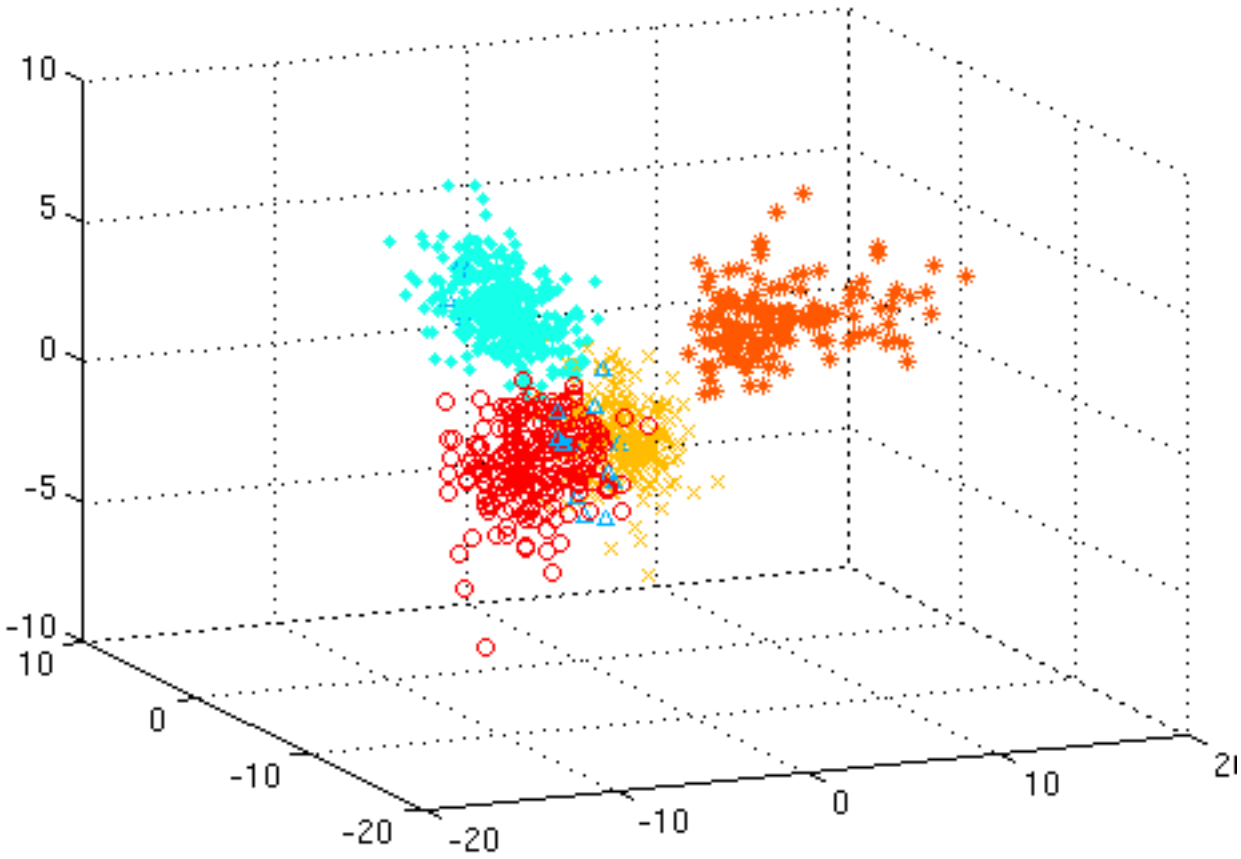
Study of large and highly stratified population datasets by combining iterative pruning principal component analysis and structure. Limpiti T, Intarapanich A, Assawamakin A, Shaw PJ, Wangkumhang P, Piriyapongsa J, Ngamphiw C, Tongsima S. ***BMC Bioinformatics***. 2011 Jun 23;12:255.



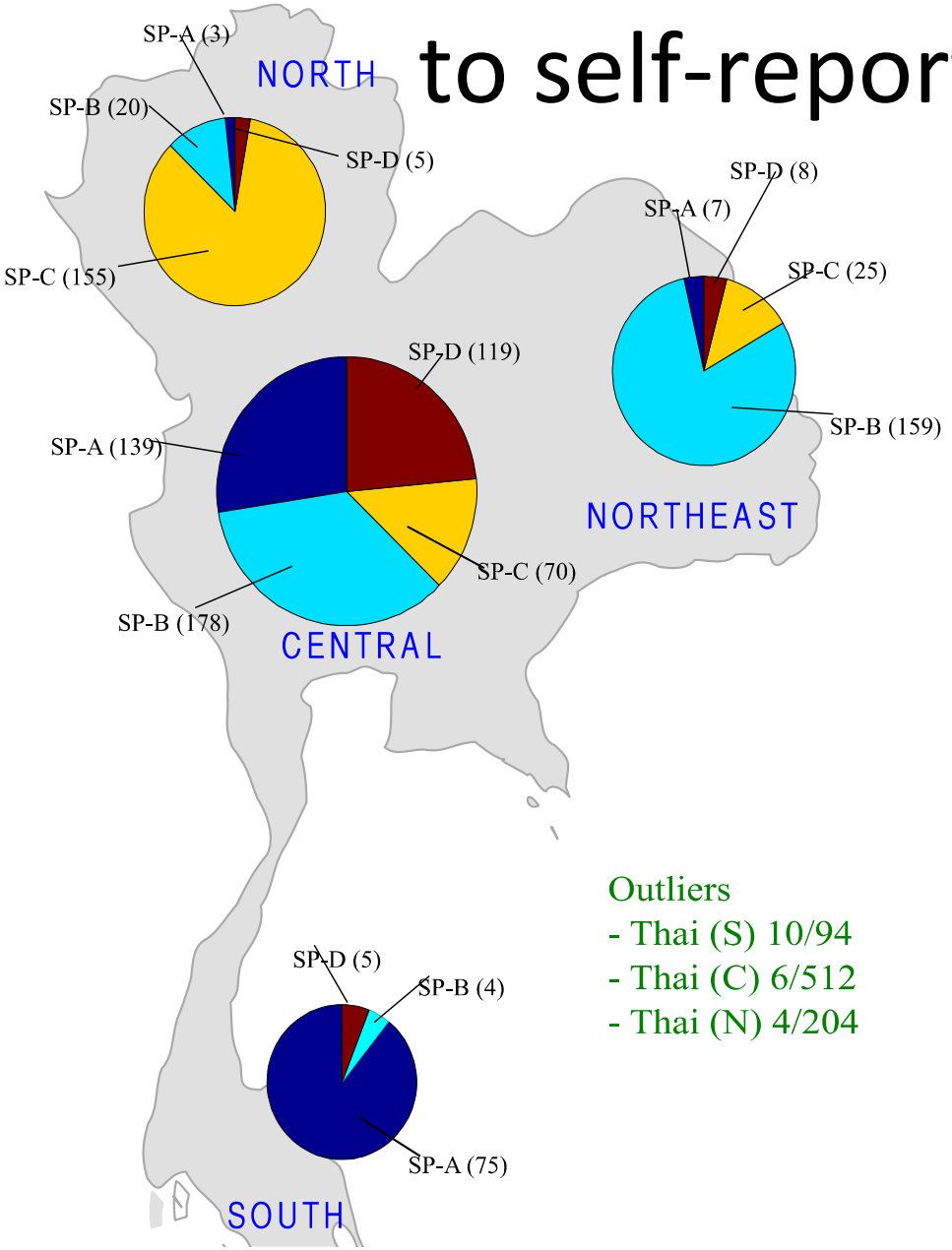
- Grouping according to genetic affiliation
- Some individuals were assigned to different countries.



Thai subpopulations resolved by ipPCA



Subpopulation distributions according to self-report origins



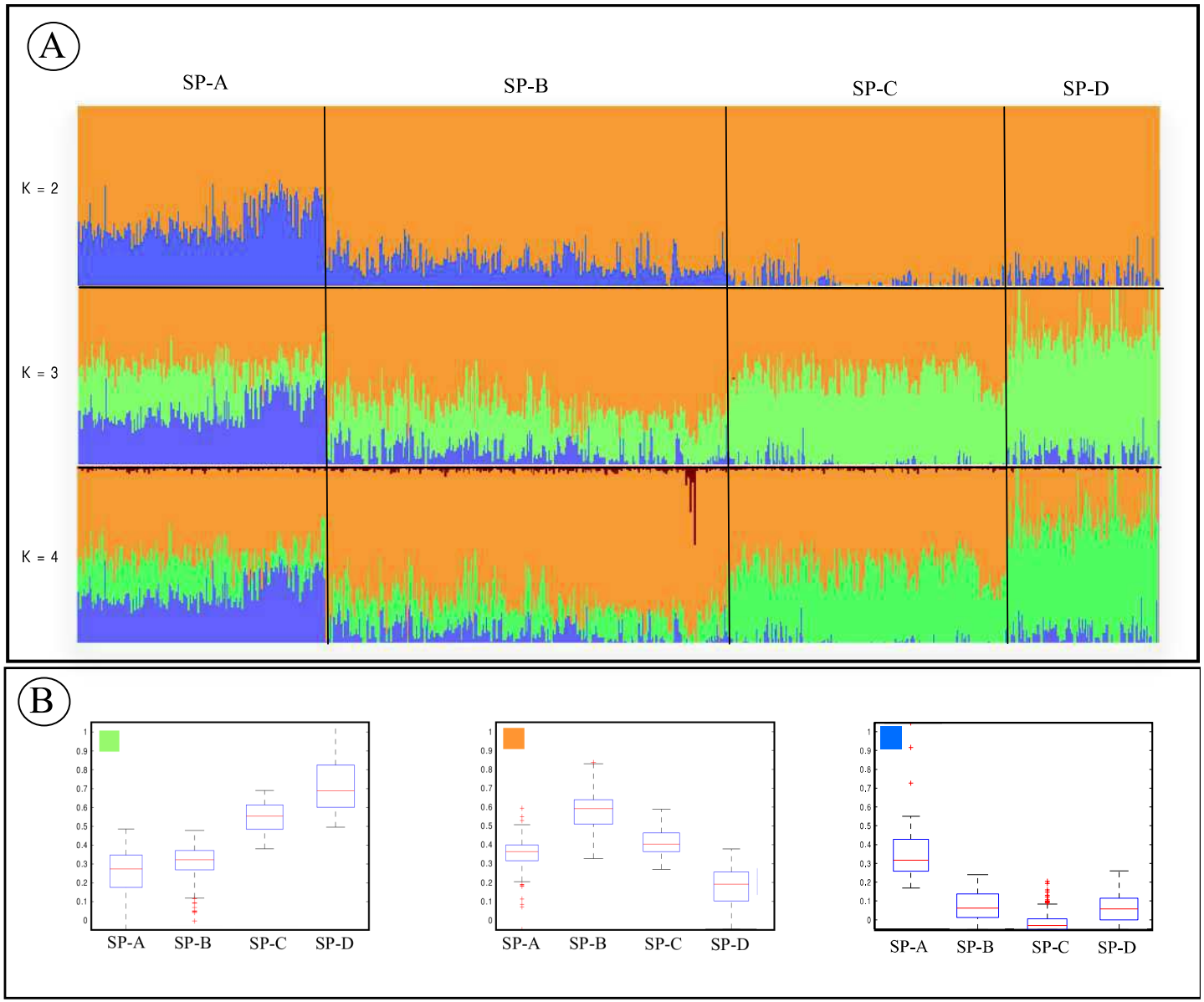
SP-C

SP-B

SP-D

SP-A

Outliers
- Thai (S) 10/94
- Thai (C) 6/512
- Thai (N) 4/204


A
SP-A
SP-B
SP-C
SP-D
K = 2
K = 3
K = 4
B
SP-A
SP-B
SP-C
SP-D
SP-A
SP-B
SP-C
SP-D
SP-A
SP-B
SP-C
SP-D

Discriminative SNPs

- Are there SNPs that can differentiate any two subpopulations, e.g., SP-A vs SP-B or -C or –D
- Phenotype-Genotype??

rs1426654 - SLC24A5

	AA	AB	BB
SPA	0.75	0.24	0.02
SPB	0.92	0.08	0.00
SPC	0.96	0.04	0.00
SPD	0.92	0.08	0.00

rs923336 - MTPN

	AA	AB	BB
SPA	0.37	0.46	0.17
SPB	0.50	0.43	0.07
SPC	0.35	0.48	0.17
SPD	0.23	0.53	0.23

rs3756464 - SLC45A2

	AA	AB	BB
SPA	0.45	0.46	0.08
SPB	0.56	0.35	0.09
SPC	0.72	0.24	0.03
SPD	0.56	0.39	0.05

rs671 - ALDH2

	AA	AB	BB
SPA	0.82	0.17	0.01
SPB	0.89	0.11	0.00
SPC	0.89	0.11	0.00
SPD	0.66	0.31	0.03

Science. 2005 Dec 16;310(5755):1782-6.

SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans.

Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Jurynec MJ, Mao X, Humphreville VR, Humbert JE, Sinha S, Moore JL,

ARTICLE

A Genomewide Association Study of Skin Pigmentation in a South Asian Population

Renee P. Stokowski, P. V. Krishna Pant, Tony Dadd, Amelia Fereday, David A. Hinds, Carl Jarman, Wendy Filsell, Rebecca S. Ginger, Martin R. Green, Frans J. van der Ouderaa, and David R. Cox

We have conducted a multistage genomewide association study, using 1,620,742 single-nucleotide polymorphisms to systematically investigate the genetic factors influencing intrinsic skin pigmentation in a population of South Asian descent. Polymorphisms in three genes—*SLC24A5*, *TYR*, and *SLC45A2*—yielded highly significant replicated associations with skin-reflectance measurements, an indirect measure of melanin content in the skin. The associations detected in these three genes, in an additive manner, collectively account for a large fraction of the natural variation of skin pigmentation in a South Asian population. Our study is the first to interrogate polymorphisms across the genome, to find genetic determinants of the natural variation of skin pigmentation within a human population.

Humans possess an impressive range of skin pigmentation, both within and between populations. This diversity is highly correlated with geographical location, indicating that environmental factors as well as genetics strongly influence skin color. The predominant environmental variable affecting skin pigmentation is sunlight, and it is certain that skin pigments play an important role in both protecting DNA from the effects of UV irradiation[1-4] and influencing the availability of UV radiation for the synthesis of necessary compounds, such as vitamin D.[5-7] Epidemiological studies in humans show that skin pigmentation is a polygenic quantitative trait with high herit-

mentation the same across different ethnic populations, and are there still-undiscovered pigmentation genes?

With the availability of the entire human genomic sequence in 2001,[27,28] the identification of millions of SNPs across the genome,[29-31] and the development of high-throughput genotyping technologies, the tools were available for investigation of the genetic components controlling human skin pigmentation with use of a high-density genomewide association study. In the present study, we applied a three-tiered methodology of quantitative pooled genotyping followed by individual genotyping of associated SNPs in original and replicate population

Int J Epidemiol. 2013 Feb;42(1):318-28. doi: 10.1093/ije/dys221. Epub 2012 Dec 14.

Is aldehyde dehydrogenase 2 a credible genetic instrument for alcohol use in Mendelian randomization analysis in Southern Chinese men?

Au Yeung SL, Jiang C, Cheng KK, Liu B, Zhang W, Lam TH, Leung GM, Schooling CM.

J Hum Hypertens. 2013 Mar;27(3):181-6. doi: 10.1038/jhh.2012.15. Epub 2012 May 3.

Association of a functional single-nucleotide polymorphism in the ALDH2 gene with essential hypertension depends on drinking behavior in a Chinese Han population.

Int J Cancer. 2013 Apr 15;132(8):1868-77. doi: 10.1002/ijc.27803. Epub 2012 Sep 28.

Single nucleotide polymorphisms of ADH1B, ADH1C and ALDH2 genes and esophageal cancer: a population-based case-control study in China.

Wu M, Chang SC, Kampman E, Yang J, Wang XS, Gu XP, Han RQ, Liu AM, Wallar G, Zhou JY, Kok FJ, Zhao JK, Zhang ZF.

Department of Chronic Disease Control, Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, Jiangsu, China.

Abstract

Alcohol drinking is a major risk factor for esophageal cancer (EC) and the metabolism of ethanol has been suggested to play an important in esophageal carcinogenesis. Epidemiologic studies, including genomewide association studies (GWAS), have identified single nucleotide polymorphisms (SNPs) in alcohol dehydrogenases (ADHs) and aldehyde dehydrogenases (ALDHs) to be associated with EC. Using a population-based case-control study with 858 EC cases and 1,081 controls conducted in Jiangsu Province, China, we aimed to provide fur information on the association of ADH1B (rs1229984), ADH1C (rs698) and ALDH2 (rs671) polymorphisms with EC in a Chinese population Results showed that ADH1B (rs1229984) was associated with EC with odds ratios (ORs) of 1.34 [95% confidence interval (CI): 1.08-1.66] G-allele carriers compared to A/A homozygotes. No heterogeneity was detected on this association across different strata of alcohol drinkir and tobacco smoking. Statistical interaction between ALDH2 (rs671) and alcohol drinking on EC susceptibility in both additive and

Conclusions

- Individuals can be grouped into different “genetic affiliated” groups
- Global SNP patterns can be used to cluster individuals
- Such unique patterns can hint us more on local selection/adaptation
- Complex traits involve both Genetics as well as environmental factors