学　位　論　文　（要約）

Evolutionary Study on Deletion Polymorphism of the *GSTM1* Gene in Humans

（ヒト *GSTM1* 欠失多型の進化学的研究）

平成 28 年博士（理学）申請

東京大学　大学院理学系研究科

生物科学専攻

齊藤真理恵

Doctoral Thesis

Evolutionary Study on Deletion Polymorphism of the *GSTM1* Gene in Humans

Marie Saitou

Department of Biological Sciences

Graduate School of Science

The University of Tokyo

2016

Abstract

The glutathione S-transferase (GST) gene superfamily comprises phase II detoxification enzymes that catalyze conjugation of glutathione (GSH) to xenobiotics (Sheehan et al., 2001). GSTs play key roles in cellular protection against xenobiotics (McIlwain et al., 2006). Genetic variants of GSTs have been studied with respect to disease susceptibility and drug resistance. The GST-μ (GSTM) family is encoded by a tandem gene cluster on chromosome one; a whole gene deletion of the *GSTM1* has been found (Xue et al., 1998). The *GSTM1* enzyme impairment is thought to result in inefficient detoxification, which leads to genetic damages and increased disease risks (Sheehan et al., 2001; McIlwain et al., 2006) and response rates to some chemotherapy (Hayes and Pulford, 1995).

The *GSTM1* deletion homozygous genotypes have extensively been studied in various human populations from the viewpoint of epidemiology. The *GSTM1* deletion homozygous genotypes can be observed in various human population commonly (Garte et al., 2001; Gaspar et al., 2002; Buchard et al., 2007; Saadat, 2007; Fujihara et al., 2009; Piacentini et al., 2011). For example, the prevalence of the *GSTM1* deletion homozygous genotype in Europeans, Asians and Africans was 47~57%, 42~54% and 16~36%, respectively (Garte et al., 2001). Such high frequency and these differences in the frequencies of the *GSTM1* deletion homozygous genotypes among human populations may have been because of an evolutionary advantage; however, the reasons (1) why has this deletion been maintained in the human populations at very high frequencies and (2) why does frequency of this deletion vary among populations remain unknown. The aim of this study is to reveal evolutionary force which shaped the distribution of the human

GSTM1 deletion from the viewpoint of biological anthropology.

In chapter 1, I investigated geographical distribution of the *GSTM1* deletion homozygous genotype distribution. The *GSTM1* deletion homozygous individuals showed higher sensitivity to UVB than in individuals with *GSTM1* wild-type allele. It is thus speculated that UV light irradiation was the selective pressure which facilitated relatively low frequencies of the *GSTM1* deletion homozygous genotype in Africans (Dandara et al., 2002). Dandara et al (2002) postulated that tropical populations adaptively maintained the *GSTM1* gene to protect their cells against oxidative stresses caused by strong UV irradiation. Meanwhile, little was known about the *GSTM1* genotype frequency among other populations residing in the tropics such as Southeast Asian and Oceanic populations. I collected the *GSTM1* deletion homozygous genotype data from 19 populations in Southeast Asia and Oceania, which were lacking in the previous studies and incorporated the data from experiments with published *GSTM1* deletion genotype data from 81 human populations. Comprehensive analysis of frequency of the *GSTM1* deletion genotype revealed the geographic distribution of this polymorphism of 81 populations in the previous studies and 19 populations in Southeast Asia and Oceania by the present study. It revealed that most Southeast Asian and Oceanic populations showed high frequencies of the *GSTM1* deletion homozygous genotype. There was no correlation between the *GSTM1* null allele frequency and the absolute latitude in the worldwide populations contrary to the previously raised expectation. This non-latitudinal geographical pattern of the *GSTM1* deletion is thought to be attributed to human migration, genetic drift or adaptations, but is not due to adaptation to UV irradiation.

It has been technically difficult to test neutrality of whole gene deletion

polymorphisms by comparing sequences of the wild-type allele with those of mutant alleles. In chapter 2, I calculated statistics for neutrality tests and analyzed haplotypes using the flanking regions of the *GSTM1* deletion, following the methods of Eaaswarkhanth et al. (2016), using the 1000 genome datasets. Using 1000 Genome datasets, I observed that LD between the *GSTM1* deletion and SNVs was moderately conserved only in CHB. The decay of LD beyond the *GSTM1* deletion suggests that gene conversions and recombinations broke the LD between the flanking SNVs and the deletion. However, this result does not disprove the possibility of recurrent deletion. For the Tajima's D, no difference were observed between the target regions which locate on the *GSTM1* deletion flanking regions and control regions. It suggests that the *GSTM1* deletion is neutral or LD between the deletion and target regions were too weak to catch signature of natural selection on the deletion. Haplotypes of the target regions were highly differentiated between East Asia and Africa. The East Asian-dominant SNVs on target regions significantly change other *GSTM* genes according to the GTEx-portal data. It is thus possible that in East Asia, the *GSTM1* deletion hitchhiked the East-Asian dominant haplotypes which experienced non-neutral evolution.

The *GSTM1* deletion allele is thought to have been generated by homologous recombination of two SDs (segmental duplications) (Xu et al. 1998). SDs have been defined as long (≥1–5 kb) and highly similar (≥90% similarity) sequences appear to have arisen by duplication (Bailey and Eichler, 2006). SDs account for 5% of the human genome (Bailey et al., 2002). Since SDs can generate gene copy number variations by initiating non-allelic homologous recombination, they have been contributed gene family evolution in the primate genomes, in particular, in the ape genomes (Bailey and Eichler, 2006). SDs and gene copy number variations thus have had considerable impact on the

primate genomic evolution. Recently developed genomic datasets make it possible to conduct comparative genomic analysis of various species to reveal genomic evolution. In chapter 3, I conducted *in silico* comparative genomic analyses among primates to reveal evolutionary history of the *GSTM* genes and SDs. Comparative genomic analyses revealed the evolutionary history of the primate *GSTM* genes. The chimpanzee *GSTM5*, gorilla *GSTM4* and rhesus macaque *GSTM1* were pseudogenized. The nonhuman primate *GSTM1* genes were flanked by the two SD sequences as well as humans. It is thus suggested that the SDs are thought to be generated in the early stage of primate evolution. The nonhuman primate SDs may cause the *GSTM1* gene deletion.

In chapter 4, experimental comparative studies of the *GSTM1* region among primates were conducted. Chimpanzees have the *GSTM1* deletion allele as well as humans at polymorphic state. Sliding window analyses and phylogenetic analyses demonstrated that the human *GSTM1* deletion allele and chimpanzee *GSTM1* deletion allele were generated independently. Sliding window analyses also revealed break points of SD fusion and gene conversions in the human and the chimpanzee. A linage-specific sequence-swapping in chimpanzee lineage was also detected.

Studies on this thesis revealed the complexity of the evolutionary history of a gene deletion polymorphism which has biomedical importance: recurrent deletions in humans and chimpanzee, frequent gene conversions, and recombinations in the flanking SDs. The framework of this study can be expanded to studies on other copy number variations.

Contents

General Introduction

Studies with recently developed whole-genome sequencing and array-based method indicated the heterogeneity of copy number variation in the human genome (Conrad et al. 2010; Handsacker et al. 2011). Several copy number variations have been associated with inter-individual differences in drug response (Feuk et al. 2006; Gamazon et al. 2011).

The glutathione S-transferase (GST) gene superfamily comprises phase II detoxification enzymes that catalyze conjugation of glutathione (GSH) to xenobiotics (Sheehan et al., 2001). GSTs are expressed in response to a variety of stresses and play key roles in cellular protection against xenobiotics (McIlwain et al., 2006). GSTs are involved in the metabolic detoxification of products generated by oxidative stress, electrophilic compounds, carcinogens, environmental toxins and therapeutic drugs (McIlwain et al., 2006). GSTs have been classified into three families, cytosolic, mitochondrial and membrane-bound microsomal, by their cellular localization (reviewed in Frova, 2006). The human cytosolic GST family comprises seven main classes according to chromosomal localization of the genes: α, μ, ω, π, σ, θ and ζ (Hayes et al., 2005). Genetic variants of GSTs have been studied with respect to disease susceptibility and drug resistance. The GST-μ (GSTM) family is encoded by a tandem gene cluster on chromosome one as 5'-*GSTM4*-*GSTM2*-*GSTM1*-*GSTM5*-*GSTM3*–3'; a whole gene deletion of the *GSTM1* has been found (Xue et al., 1998). The *GSTM1* enzyme impairment is thought to result in inefficient detoxification, which leads to genetic damages and increased disease risks. In fact, the *GSTM1* deletion homozygous genotype is associated with various types of cancer (Sheehan et al., 2001; McIlwain et al., 2006), asthma (Minelli et al., 2010), diabetes (Yalin et al., 2007) and

response rates to some chemotherapy (Hayes and Pulford, 1995).

The *GSTM1* deletion homozygous genotypes have extensively been studied in various human populations from the viewpoint of epidemiology. The *GSTM1* deletion homozygous genotypes can be observed in various human population commonly (Garte et al., 2001; Gaspar et al., 2002; Buchard et al., 2007; Saadat, 2007; Fujihara et al., 2009; Piacentini et al., 2011). For example, the prevalence of the *GSTM1* deletion homozygous genotype in Europeans, Asians and Africans was 47~57%, 42~54% and 16~36%, respectively (Garte et al., 2001). Such high frequency and these differences in the frequencies of the *GSTM1* deletion homozygous genotypes among human populations may have been because of an evolutionary advantage; however, the reason why this deletion has been maintained in the human populations at very high frequencies remains unknown. Recently, evolution driven by gene loss was discussed (Albalat and Cañestro, 2016). An experimental study showed that loss-of-function mutations, especially of metabolizing genes, can be adaptive in bacteria, by modifying gene expression and the flow of metabolites to better fit new environments (Hottes, et al., 2013). The aim of this study is to reveal evolutionary background of the common human *GSTM1* deletion from the viewpoint of biological anthropology.

In chapter 1, I investigated geographical distribution of the *GSTM1* deletion homozygous genotype distribution. I collected the *GSTM1* deletion homozygous genotype data from 19 populations in Southeast Asia and Oceania, which were lacking in the previous studies and incorporated the data from experiments with published *GSTM1* deletion genotype data from 81 human populations.

It has been technically difficult to test neutrality of whole gene deletion polymorphisms by comparing sequences of the wild-type

allele with those of mutant alleles. In chapter 2, I calculated statistics for neutrality tests and analyzed haplotypes using the flanking regions of the *GSTM1* deletion, following the methods of Eaaswarkhanth et al. (2016), using the 1000 genome datasets.

The *GSTM1* deletion allele is thought to have been generated by homologous recombination of two SDs (segmental duplications) (Xu et al. 1998). SDs have been defined as long (≥1–5 kb) and highly similar (≥90% similarity) sequences appear to have arisen by duplication (Bailey and Eichler, 2006). SDs account for 5% of the human genome (Bailey et al., 2002). Since SDs can generate gene copy number variations by initiating non-allelic homologous recombination, they have been contributed gene family evolution in the primate genomes, in particular, in the ape genomes (Bailey and Eichler, 2006). SDs and gene copy number variations thus have had considerable impact on the primate genomic evolution. Recently developed genomic datasets make it possible to conduct comparative genomic analysis of various species to reveal genomic evolution. In chapter 3, I conducted *in silico* comparative genomic analyses among primates to reveal evolutionary history of the *GSTM* genes and SDs. Experimental comparative studies of the *GSTM1* region among primates were conducted in chapter 4.

Chapter 1. Geographical distribution of the human *GSTM1* gene deletion polymorphism

1.1. Introduction

The *GSTM1* deletion homozygous genotype frequency among African was lower than that of European (Garte et al. 2001; Dandara et al., 2002; Roodi et al., 2004). People live in low-latitudinal area are exposed to high dose of UV. UVB irradiation correlates with latitude and positive natural selection has operated upon many of these candidate genes to result in a latitudinal cline in human skin pigmentation (Jablonski and Chaplin 2010). Latitudinal clines of frequencies of the polymorphisms on the human vitamin D receptor and skin color genes were reported, suggesting adaptation to UVB irradiation (Tiosano et al., 2016). UV light irradiation induces oxidative stresses in cells (McIlwain et al., 2006) and GSTs detoxify reactive metabolites generated by oxidative stresses (Hayes et al., 2005); Smith et al. (2011) reported that the *GSTM1* genotypes were associated with sensitivity to UVB. The *GSTM1* deletion homozygous individuals showed higher sensitivity to UVB than in individuals with *GSTM1* wild-type allele. It is thus speculated that UV light irradiation was the selective pressure which facilitated relatively low frequencies of the *GSTM1* deletion homozygous genotype in Africans (Dandara et al., 2002). Dandara et al (2002) postulated that tropical populations adaptively maintained the *GSTM1* gene to protect their cells against oxidative stresses caused by strong UV irradiation. Meanwhile, little was known about the *GSTM1* genotype frequency among other populations residing in the tropics such as Southeast Asian and Oceanic populations. Because there has been no systematic review of geographic distribution of the *GSTM1* deletion, studies about anthropological background of this gene has been

limited.

Conventional PCR assays used in most of previous studies detect the deletion homozygous genotype but do not distinguish homozygous wild-type individuals from heterozygous wild/deletion individuals. Distinguishing three genotypes is important because gene expression level and corresponding enzyme activity as conjugators change with the copy number of the *GSTM1* gene (McCarrol et al. 2005; Moyer et al 2007; Smith et al. 2011; Arakawa et al. 2011). Several newly designed PCR methods are capable of identifying three genotypes, wild/wild, deletion/wild and deletion/deletion, of the *GSTM1* (Roodi et al., 2004; Buchard et al., 2007). By these methods or real-time quantitative PCR, the *GSTM1* genotypes and allele frequencies were identified in European (Buchard et al., 2007), Han Chinese-American (Moyer et al., 2007), African-American (Roodi et al. 2004, Moyer et al., 2007), European-American (Roodi et al. 2004; Moyer et al., 2007), Mexican-American (Moyer et al., 2007) and Japanese (Tatewaki et al., 2009).

To test the UV-adaptation hypothesis, I first reviewed previous studies and collected data of the *GSTM1* deletion homozygous frequency data in 81 human populations from previous studies. I also revealed the *GSTM1* genotypes for 1339 individuals from 19 populations in Southeast Asia and Oceania. I then analyzed geographic distribution of the *GSTM1* null polymorphism, incorporating data from the populations in Southeast Asia and Oceania with the data from previous studies.

1.2. Materials and Methods

1.2.1. Publication search

To perform a meta-analysis, publications were selected with the following protocol (Figure 1.1). I searched for studies comprising keywords "*GSTM1* null population genotype NOT meta" on PubMed up to December 2012, and then used the PubMed filters "Abstract available", "English", "Human", and "MEDLINE" for the further selection. References of related studies were manually searched and added.

1.2.2. Inclusion/exclusion criteria

I included publications (1) reporting frequencies of the *GSTM1* deletion homozygous genotypes for more than 50 healthy individuals; (2) using another gene as a internal control in the PCR for assurance that specimens are successfully amplified and detected or using the real-time PCR; (3) with a description of ethnic background of the subjects; (4) stating the location of the study population. We excluded publications (1) of meta-analyses and review; (2) of studies based upon families; (3) for the subjects with relatively recent migrations and/or genetic admixtures. Hence a large number of studies on populations such as in the United States, Brazil, Canada, Argentina, Australia, Mexico, Puerto Rico, Hong Kong, Taiwan, Singapore, Hawaii, Shanghai, Greenland and United Kingdom were excluded. When there were multiple publications for a given population, data for the largest sample size was adopted. Latitude of each location was obtained by Google search or published maps. Decimal system of latitude was adopted. When the location of the

subjected population was not clearly mentioned in the literature, I substituted the state capital for it (Table 1.1).

インターネット公表に関する共著者全員の同意が得られていないため、1.2.3-1.2.5 については、非公開

1.2.6. Statistical Analysis

Correlations between the absolute latitude and the prevalence of the *GSTM1* null genotype were tested by Spearman's rank correlation coefficient with R (version 2. 14. 1). Deviation from the Hardy–Weinberg's equilibrium was tested by the chi-square tests. A *p*-value less than 0.05 was considered to be significant.

1.3. Results

Starting from 400 publications in the PubMed, 63 publications for 81 populations were finally included in the study (Figure 1.1). These populations were located from 64.1°N to 23.5°S. Number and origin of the populations were as following; 14 from Africa, two from America, 46 from Asia and 19 from Europe. Table 1.1 shows the frequency of the *GSTM1* deletion homozygous genotype with the absolute latitude in each population. The frequency of the *GSTM1* deletion homozygous genotype ranged from 0.04 in Guarani (Brazil) (Gaspar et al., 2002) to 0.65 in Southern Thais (Kietthubthew, 2006).

インターネット公表に関する共著者全員の同意が得られていないため、1.3 の以降の部分及び 1.4 については、非公開

Table 1.1. *GSTM1* null genotype frequencies in 84 worldwide populations

Population (Location)	Latitude[1]	Number	Freq. of the *GSTM1* deletion homozygous	Reference
AFRICA				
Ibo (Abuja)	9.1	101	0.23	Ebeshi et al., 2011
Hausa (Abuja)	9.1	98	0.37	Ebeshi et al., 2011
Ethiopian (Addis Ababa)	9	153	0.44	Piacentini et al., 2011
Egyptian (Cairo)	30	200	0.56	Hamdy et al., 2003
Mandinka (Gambia)	13.5	114	0.28	Kirk et al., 2005
Fula (Gambia)	13.5	77	0.23	Kirk et al., 2005
Wollof (Gambia)	13.5	50	0.16	Kirk et al., 2005
Yoruba (Abuja)	9.1	101	0.31	Ebeshi et al., 2011
Sudanese (Khartoum)	15.5	114	0.39	Tiemersma et al., 2001
Tunisian (Mahdia)	35.5	182	0.54	Lakhdar et al., 2010
Somali (Mogadishu)	2	100	0.4	Buchard et al., 2007
Ovambo (Windhoek)	22.6	134	0.11	Fujihara et al., 2009
Cameroonian (Yaoundé)	3.8	126	0.28	Piacentini et al., 2011
Tunisians (Sousse)	35.8	186	0.63	Salem et al., 2011
AMERICA				
Guarani (Brazil)	23.2	51	0.04	Gaspar et al., 2002
Ache (Paraguay)	23.5	67	0.36	Gaspar et al., 2002

ASIA				
Bahrainis (Manama)	26.2	167	0.5	Salem et al., 2011
Thailander (Bangkok)	13.8	320	0.6	Pakakasama et al., 2005
Lebanese (Beirut)	33.9	141	0.53	Salem et al., 2011
Chinese (Beijing)	39.9	481	0.44	Li et al., 2012
Indian (Mumbai)	19	82	0.17	Nair et al., 1999
Chinese (Chengdu)	30.6	410	0.51	Jing et al., 2012
Indian (Delhi)	28.6	309	0.21	Singh et al., 2009
Chinese (Guangzhou)	23.1	412	0.47	Zhang et al., 2011
Vietnamese (Ha Nam)	20.5	100	0.42	Agusa et al., 2010
Chinese (Harbin)	45.8	226	0.46	Lu et al., 2011
Han (Henan)	33.9	212	0.51	Song et al., 2009
Pakistani (Islamabad)	33.7	162	0.36	Khan et al., 2010
Indonesian (Jakarta)	6.2	162	0.56	Amtha et al., 2009
Druze	31.8	159	0.6	Karban et al., 2011
Non-Ashkenazi Jews	31.8	172	0.55	Karban et al., 2011
Arab Moslem	31.8	101	0.56	Karban et al., 2011
Ashkenazi Jews	31.8	96	0.55	Karban et al., 2011
Chinese (Yangzhong)	32.1	419	0.51	Setiawan et al., 2000
Kabul, Pashtuns	34.5	257	0.42	Saify et al., 2012
Kabul, Tajiks	34.5	217	0.48	Saify et al., 2012

Kabul, Hazaras	34.5	120	0.53	Saify et al., 2012
Kabul, Uzbeks	34.5	62	0.4	Saify et al., 2012
Kashmiri (Srinagar)	34.5	195	0.42	Malik et al., 2010
Indian (Kerala)	8.5	146	0.27	Sreeja et al., 2005
Thai (Khon Kaen)	16.4	94	0.6	Settheetham-Ishida et al., 2009
Japanese (Kitakyusyu)	33.8	126	0.44	Katoh et al., 1996
Tibetan (Lhasa)	29.4	86	0.61	Yan et al., 2006
Maharashtrian (Nagpur)	21.3	314	0.35	Devi et al., 2008
Bahrainis (Manama)	26.2	167	0.5	Salem et al., 2011
Filipino (Quezon)	14.7	127	0.59	Baclig et al., 2012
Chinese (Meizhou)	23.4	512	0.62	Pan et al., 2011
Mizos (Mizoram)	23.4	204	0.48	Malakar et al., 2012
Japanese (Nagoya)	35.2	320	0.58	Niwa et al., 2005
Chinese (Qingdao)	36.1	366	0.43	Jiang et al., 2011
Saudi (Riyadh)	24.7	513	0.55	Al-Dayel et al., 2008
Korean (Seoul)	37.5	549	0.51	Uhm et al., 2007
Iranian (Shiraz)	29.6	169	0.51	Moasser et al., 2012
Southern Thai (Songkhla)	7.2	164	0.65	Kietthubthew, 2006
Southern Punjab	30.1	111	0.45	Shaikh et al., 2010
Iranian (Tehran)	35.7	336	0.28	Safarinejad et al., 2011
Japanese (Tokyo)	35.4	203	0.5	Tamaki et al., 2011
Mongolian (Ulan Bator)	47.9	207	0.46	Fujihara et al., 2009

North Indian (Lucknow)	26.9	300	0.22	Singh et al., 2010
Han (Wenzhou)	28	152	0.48	Chen et al., 2012
Han (Xi'an)	34.3	763	0.52	Liu et al., 2009
Turkish (Ankara)	39.9	231	0.54	Ada et al., 2012
EUROPE				
Greek (Athens)	38	171	0.52	Dialyna et al., 2001
German (Heidelberg)	49.4	1251	0.51	Timofeeva et al., 2010
Mediterranean (Barcelona)	41.4	192	0.49	To-Figueras et al., 1997
Normanean (Basse-Normandie)	49.2	120	0.49	Abbas et al., 2004
Danish (Copenhagen)	55.7	200	0.53	Buchard et al., 2007
European (Covilha)	40.2	102	0.4	Ramalhinho et al., 2011
Scottish (Aberdeen)	57.1	383	0.58	Little et al., 2006
Finnish European (Helsinki)	60.3	478	0.42	Mitrunen et al., 2001
Ukrainian (Kiev)	50.5	253	0.51	Ebrahimi et al., 2004
Slovenian (Ljubljana)	46.1	116	0.54	Dolzan, et al., 2006
Polish (Lodz)	51.8	233	0.48	Kargas et al., 2003
Spanish (Madrid)	40.4	94	0.55	Piacentini et al., 2011
European (Martin)	49.1	220	0.48	Matakova et al., 2009
Czech (Bruno)	49.1	331	0.5	Holla et al., 2006

Norvegian (Oslo)	59.9	357	0.51	Skjelbred et al., 2011
Icelander (Reykjavik)	64.1	395	0.54	Gudmundsdottir et al., 2001
Italian (Rome)	41.9	143	0.53	Piacentini et al., 2012
Italian (Florence)	43.4	546	0.5	Palli et al., 2005
European (Vienna)	48.2	305	0.55	Gundacker et al., 2009

[1]absolute latitude

インターネット公表に関する共著者全員の同意が得られていないため、第二章、第三章、第四章については、非公開

Acknowledgements

I would like to express my gratitude to my advisor Prof. Takafumi Ishida for his continuous support and patience for more than five years since I entered the Department of Biology when I was senior. Without his guidance it would not be possible to complete my Ph.D. study.

My sincere thanks also goes to Prof. Omer Gokcumen, who provided me an opportunity to join his laboratory as visiting student for his insightful comments and encouragement. I thanks to his students as well.

I also would like to thank Prof Yoko Satta for energetic and stimulating discussions and support with molecular evolutionary theory and analysis.

I thank current and former members of Prof. Ishida's laboratory for their help, discussions and patience.

This research was partly supported by the JSPS research grant and GPLLI program.

References

Abbas, A. *et al.* susceptibility to esophageal cancer in a French population : Different pattern of squamous cell carcinoma and adenocarcinoma. *World J Gastroenterol* **10,** 3389–3393 (2004).

Ada, A. O. *et al.* Association between GSTM1, GSTT1, and GSTP1 polymorphisms and lung cancer risk in a Turkish population. *Mol. Biol. Rep.* **39,** 5985–93 (2012).

Albalat, R. & Cañestro, C. Evolution by gene loss. *Nat. Rev. Genet.* **17,** 379–391 (2016).

Al-Dayel, F. *et al.* Polymorphisms of drug-metabolizing enzymes CYP1A1, GSTT and GSTP contribute to the development of diffuse large B-cell lymphoma risk in the Saudi Arabian population. *Leuk. Lymphoma* **49,** 122–9 (2008).

Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–402 (1997).

Amtha, R. *et al.* GSTM1, GSTT1 and CYP1A1 polymorphisms and risk of oral cancer: a case-control study in Jakarta, Indonesia. *Asian Pac. J. Cancer Prev.* **10,** 21–6 (2009).

Arakawa, S. *et al.* Evaluation of Hepatic Glutathione Transferase Mu 1 and Theta 1 Activities in Humans and Mice Using Genotype Information □ ABSTRACT : **40,** 497–503 (2012).

Arbiza, L., Zhong, E. & Keinan, A. NRE: a tool for exploring neutral loci in the human genome. *BMC Bioinformatics* **13,** 301 (2012).

Baclig, M. O. *et al.* Association of glutathione S-transferase T1 and M1 genotypes with chronic liver diseases among Filipinos. *Int. J. Mol. Epidemiol. Genet.* **3,** 153–9 (2012).

Bailey, J. a & Eichler, E. E. Primate segmental duplications: crucibles of evolution, diversity and disease. *Nat. Rev. Genet.* **7,** 552–64 (2006).

Bailey, J. a *et al.* Recent segmental duplications in the human genome. *Science* **297,** 1003–1007 (2002).

Beall, C. M. *et al.* Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders. *Proc Natl Acad Sci U S A* **107,** 11459–11464 (2010).

Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74,** 1111–20 (2004).

Buchard, A., Sanchez, J. J., Dalhoff, K. & Morling, N. Multiplex PCR detection of GSTM1, GSTT1, and GSTP1 gene variants: simultaneously detecting GSTM1 and GSTT1 gene copy number and the allelic status of the GSTP1 Ile105Val genetic variant. *J. Mol. Diagn.* **9,** 612–7 (2007).

Carbone, L. *et al.* Gibbon genome and the fast karyotype evolution of small apes. *Nature* **513,** 195–201 (2014).

Cartegni, L., Chew, S. L. & Krainer, A. R. Listening To Silence and Understanding Nonsense: Exonic Mutations That Affect Splicing. *Nat. Rev. Genet.* **3,** 285–298 (2002).

Chen, B. *et al.* Glutathione S-transferases T1 null genotype is associated with susceptibility to aristolochic acid nephropathy. *Int. Urol. Nephrol.* **44,** 301–7 (2012).

Conrad, D. F. *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* **464,** 704–12 (2010).

Dandara, C. *et al.* Genetic polymorphism of cytochrome P450 1A1 (Cyp1A1) and glutathione transferases (M1, T1 and P1) among Africans. *Clin. Chem. Lab. Med.* **40,** 952–7 (2002).

Devi, S. S. *et al.* Distribution of detoxifying genes polymorphism in Maharastrian population of central India. *Chemosphere* **70,** 1835–9 (2008).

Dialyna, I., Arvanitis, D. & Spandidos, D. Genetic polymorphisms and transcriptional pattern analysis of CYP1A1, AhR, GSTM1, GSTP1 and GSTT1 genes in breast cancer. *Int. J. Mol. Med.* **8,** 79–87 (2001).

Eaaswarkhanth, M., Pavlidis, P. & Gokcumen, O. Geographic distribution and adaptive significance of genomic structural variants: an anthropological genetics perspective. *Hum. Biol.* **86,** 260–75 (2014).

Ebeshi, B. U., Bolaji, O. O. & Masimirembwa, C. M. Glutathione-S-transferase (M1 and T1) polymorphisms in Nigerian populations. *J. Med. Genet.* **3,** 56–60 (2011).

Ebrahimi, M., Podolskaya, S. V. & Grigorieva, N. Genetic polymorphisms of glutathione S-transferase susceptibility in Ukrainian population. *Iran. J. Biotechnol.* **2,** 230–235 (2004).

Feuk, L., Marshall, C. R., Wintle, R. F. & Scherer, S. W. Structural variants: changing the landscape of chromosomes and design of disease studies. *Hum. Mol. Genet.* **15 Spec No,** 57–66 (2006).

Fujihara, J. *et al.* Cytochrome P450 1A1, glutathione S-transferases M1 and T1 polymorphisms in Ovambos and Mongolians. *Leg. Med. (Tokyo).* **11 Suppl 1,** S408-10 (2009).

Gamazon, E. R., Huang, R. S., Dolan, M. E. & Cox, N. J. Copy number polymorphisms and anticancer pharmacogenomics. *Genome Biol.* **12,** R46 (2011).

Garte, S. *et al.* Metabolic Gene Polymorphism Frequencies in Control Populations Metabolic Gene Polymorphism Frequencies in Control Populations. *Cancer Epidemiol Biomarkers Prev.* **10,** 1239–1248 (2001).

Gaspar, P. *et al.* Polymorphisms of CYP1a1, CYP2e1, GSTM1, GSTT1, and TP53 genes in Amerindians. *Am. J. Phys. Anthropol.* **119,** 249–56 (2002).

Gasper, A. V. *et al.* Glutathione S-transferase M1 polymorphism and metabolism of sulforaphane from standard and high-glucosinolate broccoli. *Am. J. Clin. Nutr.* **82,** 1283–1291 (2005).

Gibbs, R. *et al.* Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316,** 222–34 (2007).

Gokcumen, O. *et al.* Primate genome architecture influences structural variation mechanisms and functional consequences. (2013). *Proc Natl Acad Sci U S A.* **110,** 15764–15769 (2013).

Gudmundsdottir, K., Tryggvadottir, L. & Eyfjord, J. E. GSTM1 , GSTT1 , and GSTP1 Genotypes in Relation to Breast Cancer Risk and Frequency of Mutations in the p53 Gene GSTM1 , GSTT1 , and GSTP1 Genotypes in Relation to Breast Cancer Risk and Frequency of Mutations in the p53 Gene 1. *Cancer Epidemiol. Biomarkers Prev.* **10,** 1169–1173 (2001).

Gundacker, C. *et al.* Genetic background of lead and mercury metabolism in a group of medical students in Austria. *Environ. Res.* **109,** 786–96 (2009).

Hahn, M. W., Demuth, J. P. & Han, S. G. Accelerated rate of gene gain and loss in primates. *Genetics* **177,** 1941–1949 (2007).

Hamdy, S. I. *et al.* Genotype and allele frequencies of TPMT, NAT2, GST, SULT1A1 and MDR-1 in the Egyptian population. *Br. J. Clin. Pharmacol.* **55,** 560–9 (2003).

Handsaker, R. E., Korn, J. M., Nemesh, J. & McCarroll, S. a. Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. *Nat. Genet.* **43,** 269–276 (2011).

Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **45,** 51–88 (2005).

Hayes, J. D. & Pulford, D. J. The Glutathione S-Transferase Supergene Family : Regulation of GST * and the Contribution of the lsoenzymes to Cancer Chemoprotection and Drug Resistance. *Crit. Rev. Biochem. Mol. Biol.* **30,** 445–600 (1995).

Holla, L. I., Stejskalova, a & Vasku, a. Polymorphisms of the GSTM1 and GSTT1 genes in patients with allergic diseases in the Czech population. *Allergy* **61,** 265–7 (2006).

Hong, E. P. & Park, J. W. Sample size and statistical power calculation in genetic association studies. *Genomics Inform.* **10,** 117–22 (2012).

Hottes, A. K. *et al.* Bacterial Adaptation through Loss of Function. *PLoS Genet.* **9,** (2013).

Huson, D. H. *et al.* Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8,** 460 (2007).

Jablonski, N. G. & Chaplin, G. Colloquium paper: human skin pigmentation as an adaptation to UV radiation. *Proc. Natl. Acad. Sci. U. S. A.* **107 Suppl,** 8962–8 (2010).

Jiang, Y. *et al.* Polymorphisms in GSTM1, GSTTI and GSTP1 and nasopharyngeal cancer in the east of China: a case-control study. *Asian Pac. J. Cancer Prev.* **12,** 3097–100 (2011).

Jing, C. *et al.* Glulathione-S-transferases gene polymorphism in prediction of gastric cancer risk by smoking and Helicobacter pylori infection status. *Asian Pac. J. Cancer Prev.* **13,** 3325–8 (2012).

Karban, A. *et al.* Non-Jewish Israeli IBD patients have significantly higher glutathione S-transferase GSTT1-null frequency. *Dig. Dis. Sci.* **56,** 2081–7 (2011).

Kargas, C., Krupa, R. & Walter, Z. Combined Genotype Analysis of GSTM1 and GSTT1 Polymorphisms in a Polish Population. *Hum. Biol.* **75,** 301–307 (2003).

Katoh, T. *et al.* Glutathione S-transferase M1 (GSTM1) and T1 (GSTT1) genetic polymorphism and susceptibility to gastric and colorectal adenocarcinoma. *Carcinogenesis* **17,** 1855–9 (1996).

Khan, M. I. *et al.* The association of glutathione S-transferase GSTT1 and GSTM1 gene polymorphism with pseudoexfoliative glaucoma in a Pakistani population. *Mol. Vis.* **16,** 2146–52 (2010).

Khrunin, A. V. *et al.* GSTM1 copy number variation in the context of single nucleotide polymorphisms in the human GSTM cluster. *Mol. Cytogenet.* **9,** 30 (2016).

Kietthubthew, S., Sriplung, H. & Au, W. W. Genetic and environmental interactions on oral cancer in Southern Thailand. *Environ. Mol. Mutagen.* **37,** 111–6 (2001).

Kirk, G. D. *et al.* Hepatocellular carcinoma and polymorphisms in carcinogen-metabolizing and DNA repair enzymes in a population with aflatoxin exposure and hepatitis B virus endemicity. *Cancer Epidemiol. Biomarkers Prev.* **14,** 373–9 (2005).

Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*. **33**, msw054 (2016).

Lakhdar, R. *et al.* Association of GSTM1 and GSTT1 polymorphisms with chronic obstructive pulmonary disease in a Tunisian population. *Biochem. Genet.* **48,** 647–57 (2010).

Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).

Li, C.-G., Zhao, Z.-M., Hu, M.-G. & Liu, R. Predictive role of glutathione-S-transferase gene polymorphisms in risk and prognosis of hepatocellular carcinoma. *Asian Pac. J. Cancer Prev.* **13,** 3247–52 (2012).

Lin, Y. L., Pavlidis, P., Karakoc, E., Ajay, J. & Gokcumen, O. The evolution and functional impact of human deletion variants shared with archaic hominin genomes. *Mol. Biol. Evol.* **32,** 1008–1019 (2015).

Little, J. *et al.* Colorectal cancer and genetic polymorphisms of CYP1A1, GSTM1 and GSTT1: a case-control study in the Grampian region of Scotland. *Int. J. Cancer* **119,** 2155–64 (2006).

Liu, L. *et al.* Genetic polymorphisms of glutathione S-transferase and risk of vitiligo in the Chinese population. *J. Invest. Dermatol.* **129,** 2646–52 (2009).

Locke, D. P. *et al.* Comparative and demographic analysis of orang-utan genomes. *Nature* **469,** 529–33 (2011).

Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45,** 580–585 (2013).

Low, W. Y. *et al.* Molecular evolution of glutathione S-transferases in the genus Drosophila. *Genetics* **177,** 1363–75 (2007).

Machiela, M. J. & Chanock, S. J. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31,** 3555–3557 (2014).

Malakar, M. *et al.* Genetic Polymorphism of Glutathione S-transferases M1 and T1, Tobacco Habits and Risk of Stomach Cancer in Mizoram, India. *Asian Pac. J. Cancer Prev.* **13,** 4725–32 (2012).

Malik, M. A., Upadhyay, R., Mittal, R. D., Zargar, S. A. & Mittal, B. Association of xenobiotic metabolizing enzymes genetic polymorphisms with esophageal cancer in Kashmir Valley and influence of environmental factors. *Nutr. Cancer* **62,** 734–42 (2010).

Matakova, T. *et al.* G ENE P OLYMORPHISMS OF B IOTRANSFORMING E NZYMES (GST S) AND THEIR A SSOCIATION WITH L UNG C ANCER IN. **14,** 275–279 (2009).

McCarroll, S., Hadnott, T. & Perry, G. Common deletion polymorphisms in the human genome. *Nat. gen.* **38,** 86–92 (2005).

McIlwain, C. C., Townsend, D. M. & Tew, K. D. Glutathione S-transferase polymorphisms: cancer incidence and therapy. *Oncogene* **25,** 1639–48 (2006).

McLellan, R. A. *et al.* Characterization of a Human Glutathione S -Transferase Cluster Containing a Duplicated GSTM1 Gene that Causes Ultrarapid Enzyme Activity. **965,** 958–965 (1997).

McVean, G. A. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).

Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338,** 222–6 (2012).

Mitrunen, K. *et al.* Polymorphisms and Susceptibility to Breast Cancer. *Cancer Epidemiol. Biomarkers Prev.* **10,** 229–236 (2001).

Moasser, E., Kazemi-Nezhad, S. R., Saadat, M. & Azarpira, N. Study of the association between glutathione S-transferase (GSTM1, GSTT1, GSTP1) polymorphisms with type II diabetes mellitus in southern of Iran. *Mol. Biol. Rep.* **39,** 10187–92 (2012).

Moyer, A. M. *et al.* Glutathione S-transferase T1 and M1: gene sequence variation and functional genomics. *Clin. Cancer Res.* **13,** 7207–16 (2007).

Nair, U. J., Nair, J., Mathew, B. & Bartsch, H. Glutathione S-transferase M1 and T1 null genotypes as risk factors for oral leukoplakia in ethnic Indian betel quid/tobacco chewers. *Carcinogenesis* **20,** 743–8 (1999).

Nebert, D. W. Polymorphisms in drug-metabolizing enzymes: what is their clinical relevance and why do they exist? *Am. J. Hum. Genet.* **60,** 265–271 (1997).

Niwa, Y. *et al.* Association of the NAD(P)H: quinone oxidoreductase C609T polymorphism and the risk of cervical cancer in Japanese subjects. *Gynecol. Oncol.* **96,** 423–9 (2005).

Pakakasama, S. *et al.* Polymorphisms of drug-metabolizing enzymes and risk of childhood acute lymphoblastic leukemia. *Am. J. Hematol.* **79,** 202–5 (2005).

Palli, D. *et al.* GSTT1 and GSTM1 gene polymorphisms and gastric cancer in a high-risk italian population. *Int. J. Cancer* **115,** 284–9 (2005).

Pan, S. *et al.* Human GSTs polymorphisms in the Hakka population of south China and their associations with family history of several chronic diseases. *Biomed. Environ. Sci.* **24,** 491–8 (2011).

Perry, G. H. *et al.* Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39,** 1256–60 (2007).

Perry, G. H. *et al.* Hotspots for copy number variation in chimpanzees and humans. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8006–8011 (2006).

Piacentini, S. *et al.* GSTT1 and GSTM1 gene polymorphisms in European and African populations. *Mol. Biol. Rep.* **38,** 1225–30 (2011).

Piacentini, S. *et al.* GSTM1 null genotype as risk factor for late-onset Alzheimer's disease in Italian patients. *J. Neurol. Sci.* **317,** 137–40 (2012).

Prado-Martinez, J. *et al.* Great ape genetic diversity and population history. *Nature* **499,** 471–475 (2013).

Prüfer, K. *et al.* The bonobo genome compared with the chimpanzee and human genomes. *Nature* **486,** 527–31 (2012).

Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505,** 43–9 (2014).

Pybus, M. *et al.* 1000 Genomes Selection Browser 1.0: a genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res.* **42,** D903-9 (2014).

Ramalhinho, A. C., Fonseca-Moutinho, J. A. & Breitenfeld, L. Glutathione S-transferase M1, T1, and P1 genotypes and breast cancer risk: a study in a Portuguese population. *Mol. Cell. Biochem.* **355,** 265–71 (2011).

Robinson, J. T. & Helga Thorvaldsdóttir1, Wendy Winckler1, Mitchell Guttman1, 2, Eric S Lander1–3, G. G. & J. P. M. 1Broad. Integrative Genomics Viewer. *Broad Inst.* **29,** 24–26 (2011).

Roodi, N., Dupont, W. D., Moore, J. H. & Parl, F. F. Association of Homozygous Wild-Type Glutathione S-Transferase M1 Genotype with Increased Breast Cancer Risk. *Cancer Res.* **64,** 1233–1236 (2004).

Saadat, M. GSTM1 null genotype associated with age-standardized cancer mortality rate in 45 countries from five continents: an ecologic study. *Int. J. cancer Res.* **3,** 74–91 (2007).

Sabeti, P. C. *et al.* The case for selection at CCR5-delta32. *PLoS Biol.* **3,** 1963–1969 (2005).

Safarinejad, M. R., Shafiei, N. & Safarinejad, S. H. Glutathione S-transferase gene polymorphisms (GSTM1, GSTT1, GSTP1) and prostate cancer: a case-control study in Tehran, Iran. *Prostate Cancer Prostatic Dis.* **14,** 105–13 (2011).

Saify, K., Saadat, I. & Saadat, M. Genetic polymorphisms of glutathione S-transferase T1 (GSTT1) and M1 (GSTM1) in selected populations of Afghanistan. *Mol Biol Rep,* **39,** 7855–7859 (2012).

Salem, A. H. *et al.* Genetic polymorphism of the glutathione S-transferase M1 and T1 genes in three distinct Arab populations. *Dis. Markers* **31,** 311–6 (2011).

Scally, A. & Durbin, R. Revising the human mutation rate: implications for understanding human evolution. *Nat. Rev. Genet.* **13,** 745–53 (2012).

Scally, A. *et al.* Insights into hominid evolution from the gorilla genome sequence. *Nature* **483,** 169–75 (2012).

Sequencing, T. C. & Consortium, A. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437,** 69–87 (2005).

Setiawan, V. W., Zhang, Z., Yu, G., Li, Y. & Lu, M. GSTT1 and GSTM1 Null Genotypes and the Risk of Gastric Cancer : A Case-Control Study in a Chinese Population GSTT1 and GSTM1 Null Genotypes and the Risk of Gastric Cancer : A Case-Control Study in a Chinese Population. *Cancer Epidemiol Biomarkers Prev* **9,** 73–80 (2000).

Settheetham-Ishida, W., Yuenyao, P., Kularbkaew, C., Settheetham, D. & Ishida, T. Glutathione S-transferase (GSTM1 and GSTT1) polymorphisms in cervical cancer in Northeastern Thailand. *Asian Pac. J. Cancer Prev.* **10,** 365–8 (2009).

Shaikh, R. S. *et al.* Frequency distribution of GSTM1 and GSTT1 null allele in Pakistani population and risk of disease incidence. *Environ. Toxicol. Pharmacol.* **30,** 76–9 (2010).

She, X. *et al.* A preliminary comparative analysis of primate segmental duplications shows elevated substitution rates and a great-ape expansion of intrachromosomal duplications. 576–583 (2006).

Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* **16,** 1–16 (2001).

Singh, N., Sinha, N., Kumar, S., Pandey, C. M. & Agrawal, S. Glutathione S-transferase gene polymorphism as a susceptibility factor for acute myocardial infarction and smoking in the North Indian population. *Cardiology* **118,** 16–21 (2011).

Singh, S. *et al.* Genetic polymorphism of glutathione S-transferase M1 and T1 in Delhi population of Northern India. *Environ. Toxicol. Pharmacol.* **28,** 25–9 (2009).

Skjelbred, C. F. *et al.* Influence of GSTM1, GSTT1, GSTP1, NAT1, NAT2, EPHX1, MTR and MTHFR polymorphism on chromosomal aberration frequencies in human lymphocytes. *Carcinogenesis* **32,** 399–405 (2011).

Smith, G., Weidlich, S., Dawe, R. S. & Ibbotson, S. H. Glutathione S-transferase M1 (GSTM1) genotype but not GSTT1 or MC1R genotype influences erythemal sensitivity to narrow band (TL-01) UVB phototherapy. *Pharmacogenet. Genomics* **21,** 217–24 (2011).

Song, D.-K. *et al.* Association of NAT2, GSTM1, GSTT1, CYP2A6, and CYP2A13 gene polymorphisms with susceptibility and clinicopathologic characteristics of bladder cancer in Central China. *Cancer Detect. Prev.* **32,** 416–23 (2009).

South, A. rworldmap: A New R package for Mapping Global Data. *R J.* **3,** 35–43 (2011).

Sreeja, L. *et al.* Possible risk modification by CYP1A1, GSTM1 and GSTT1 gene polymorphisms in lung cancer susceptibility in a South Indian population. *J. Hum. Genet.* **50,** 618–27 (2005).

Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313 (2014).

Sudmant, P. H. *et al.* An integrated map of structural variation in 2,504 human genomes. *Nature* **526,** 75–81 (2015).

Sudmant, P. H. *et al.* Global diversity, population stratification, and selection of human copy number variation. *Science* **80,** 1–16 (2015).

Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123,** 585–595 (1989).

Tamaki, Y. *et al.* Association between Cancer Risk and Drug-metabolizing Enzyme Gene (CYP2A6, CYP2A13, CYP4B1, SULT1A1, GSTM1, and GSTT1) Polymorphisms in Cases of Lung Cancer in Japan. *Drug Metab. Pharmacokinet.* **26,** 516–522 (2011).

Tatewaki, N., Maekawa, K., Atori, N. K. & Urose, K. K. Genetic Variations and Haplotype Structures of the Glutathione S-transferase Genes, GSTT1 and GSTM1, in a Japanese Patient Population Naoko. **24,** 118–126 (2009).

Tiemersma, E. W., Omer, R. E. & Bunschoten, A. Role of Genetic Polymorphism of Glutathione- S -Transferase T1 and Microsomal Epoxide Hydrolase in Aflatoxin-associated Hepatocellular Carcinoma Role of Genetic Polymorphism of Glutathione-S-Transferase T1 and Microsomal Epoxide Hydrolase in Aflatoxin-ass. *Cancer Epidemiol. Biomarkers Prev.* **10**, 785–791 (2001).

Timofeeva, M. *et al.* Genetic polymorphisms of MPO, GSTT1, GSTM1, GSTP1, EPHX1 and NQO1 as risk factors of early-onset lung cancer. *Int. J. Cancer* **127,** 1547–61 (2010).

Tiosano, D. et al. Latitudinal Clines of the Human Vitamin D Receptor and Skin Color-Genes. *G3*. **6**, 1251–1266 (2016).

Tiwawech, D., Srivatanakul, P., Karalak, A. & Ishida, T. Glutathione S-transferase M1 gene polymorphism in Thai nasopharyngeal carcinoma. *Asian Pac. J. Cancer Prev.* **6,** 270–5 (2005).

To-Figueras, J. polymorphisms and lung cancer risk among Northwestern Mediterraneans. *Carcinogenesis* **18,** 1529–1533 (1997).

Uhm, Y. K. *et al.* Association of glutathione S-transferase gene polymorphisms (GSTM1 and GSTT1) of vitiligo in Korean population. *Life Sci.* **81,** 223–7 (2007).

Uno, Y., Murayama, N., Kunori, M. & Yamazaki, H. Systematic identification and characterization of glutathione S-transferases in cynomolgus macaque. *Biochem. Pharmacol.* **86,** 679–90 (2013).

Weir, BS. and Cockerham, CC., Estimating F-Statistics for the Analysis of Population Structure, *Evolution*, **38**, 1358-1370 (1984).

Wooding, S. *et al.* Independent evolution of bitter-taste sensitivity in humans and chimpanzees. *Nature* **440,** 930–934 (2006).

Worley, K. C. *et al.* The common marmoset genome provides insight into primate biology and evolution. *Nat. Genet.* **46,** 850–7 (2014).

Wu, X., Muzny, D. M., Chi Lee, C. & Thomas Caskey, C. Two independent mutational events in the loss of urate oxidase during hominoid evolution. *J. Mol. Evol.* **34,** 78–84 (1992).

Xue, Y., Sun, D., Daly, A., Yang, F. & Zhou, X. Adaptive Evolution of UGT2B17 Copy-Number Variation. *Am. J. Hum. Gen.* **83,** 337–346 (2008).

Yan, H., Sun, X., Den, W., Fan, X. & Liu, T. GSTP1, GSTM1, and GSTT1 polymorphism in Tibetan mountaineers. *Biol. Sport* **23,** (2006).

Zhang, A.-P. *et al.* Glutathione S-transferase gene polymorphisms and risk of gastric cancer in a Chinese population. *Asian Pac. J. Cancer Prev.* **12,** 3421–5 (2011).