Approaches to Detecting Gene-Environment Interactions in Environmental Adaptability Using Genetic Engineering, Remote Sensing and Geographic Information Systems

Sumiko Anno[1], Kazuhiko Ohshima[2], Takashi Abe[3], Takeo Tadono[4], Aya Yamamoto[5], Tamotsu Igarashi[5]
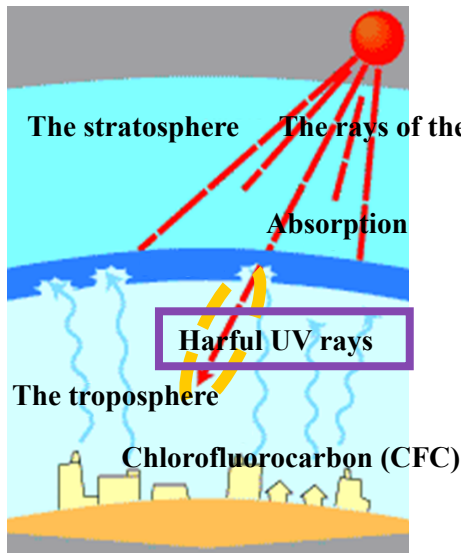
1: Shibaura Institute of Technology
2: Nagahama Institute of Bio-Science and Technology
3: Niigata University
4: Japan Aerospace Exploration Agency
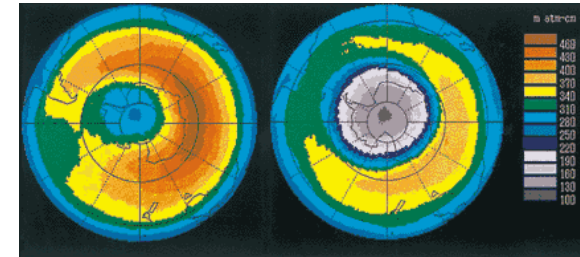**5: Remote Sensing Technology Center of Japan**

Introduction

The stratosphere   The rays of the sun

Absorption

Harful UV rays

The troposphere

Chlorofluorocarbon (CFC)

Industrial technology

Monthly average of ozone values for an entire month in the South pole, Oct., 1979 and Oct., 1999

Ozone depletion

An increased exposure to ultraviolet (UV) rays

conditions to which humans have not had to adapt

Affects the health of human beings

Skin cancer, a cataract, a viral disease such as AIDS due to a decline in immunity

Human survival requires to understand the essential of an adaptation to UV rays

The formation of environmental adaptability

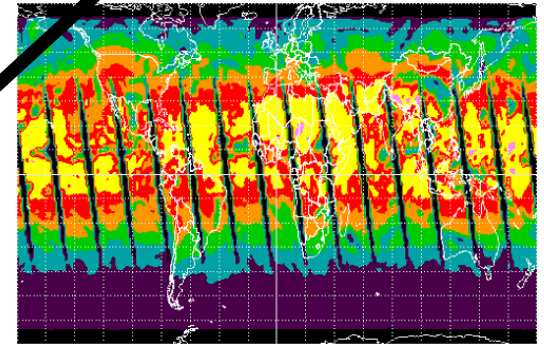Could be considered that **a stress from the outside influence a genetic mechanism in mutation form**

Gene ← **Mutation** — **Environment**

Gene → **Limbic System** ← **Environment**

Limbic System → **Hypothalamus**

Hypothalamus → **Endocrine System**

Hypothalamus → **Nervous System**

Endocrine System → **Effector Organs** ← **Nervous System**

Effector Organs → **Homeostasis**

Physiological functions of each population group can be observed or measured in terms of **phenotype emerged from genotype which is modified by culture and environment**

Skin pigmentation phenotype

One of the results of man's environmental adaptabilities
Results from localized adaptation to UVR conditions

Despite the importance of gene-environment interactions for the complex phenotype, there has been little progress in developing methods that can detect and clarify the interactions involved in skin color variation.

Investigated the evolutionary and adaptation mechanisms involved in skin color variation by detecting gene-environment interactions
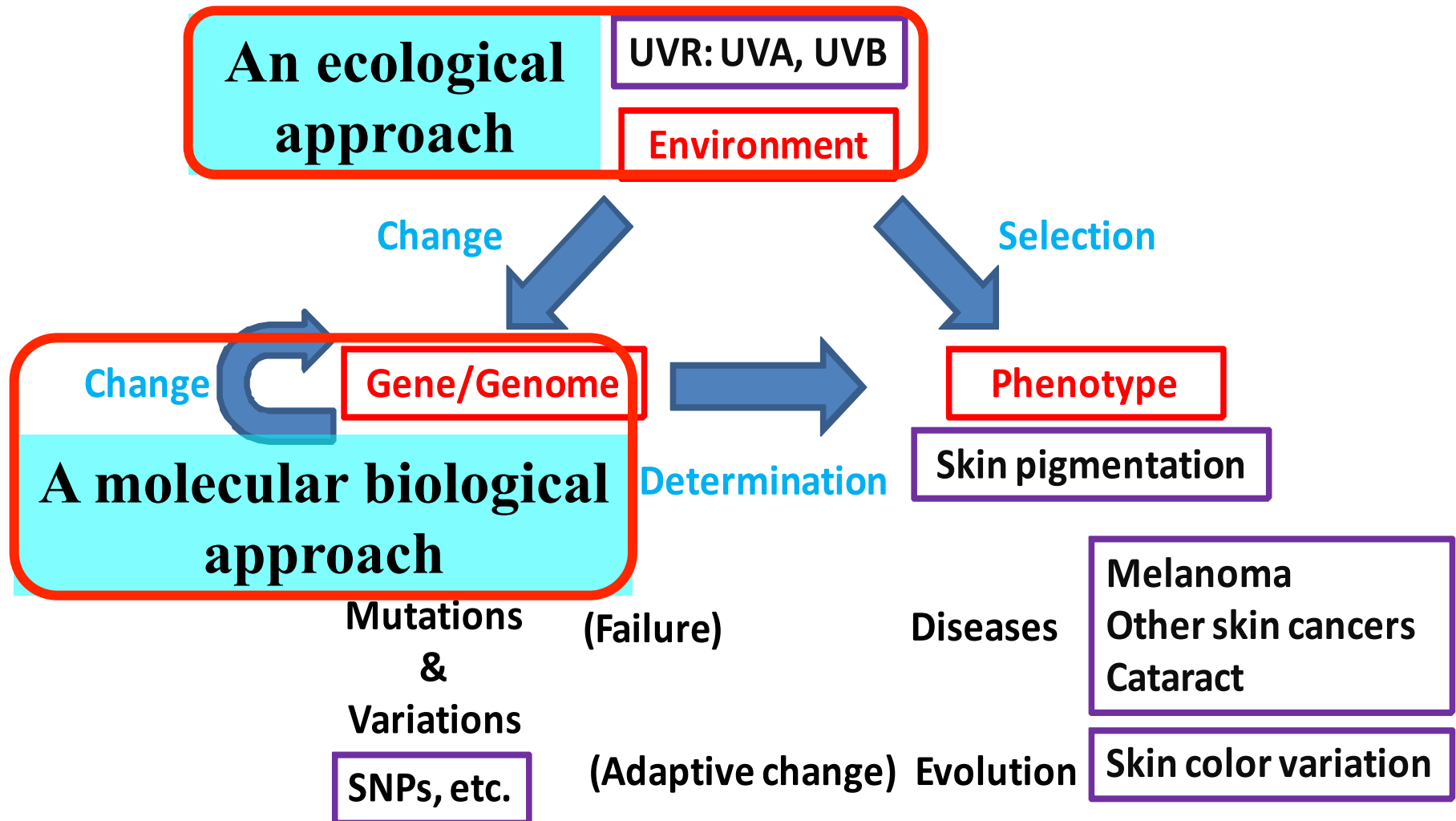
Material and Methods

An ecological approach

UVR: UVA, UVB

Environment

Change → → Selection

Change ↻ Gene/Genome → Phenotype

A molecular biological approach

Determination

Skin pigmentation

Mutations & Variations

(Failure)

Diseases

Melanoma
Other skin cancers
Cataract

SNPs, etc.

(Adaptive change) Evolution

Skin color variation

Figure The conceptual framework of new approaches to analyzing the gene-environment interactions

Two approaches to understanding the adaptation

A molecular biological approach
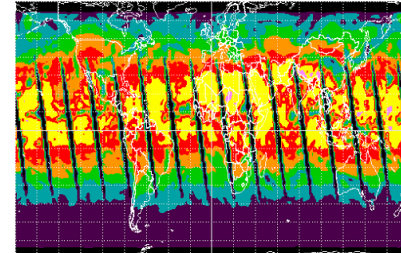
Genome DNA -- SNPs
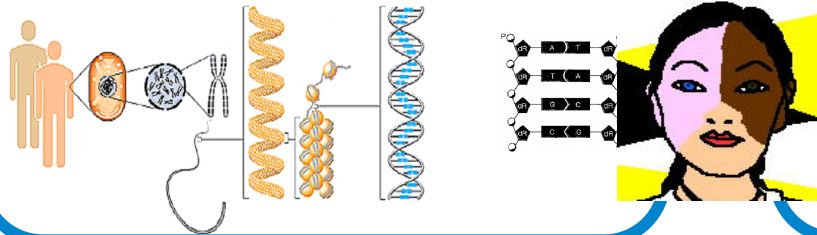
Genetic engineering was used as **a molecular biological approach**

To understand the dynamics of the human genome, which encodes the complex human phenotype
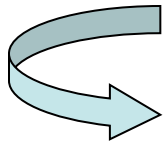
An ecological approach
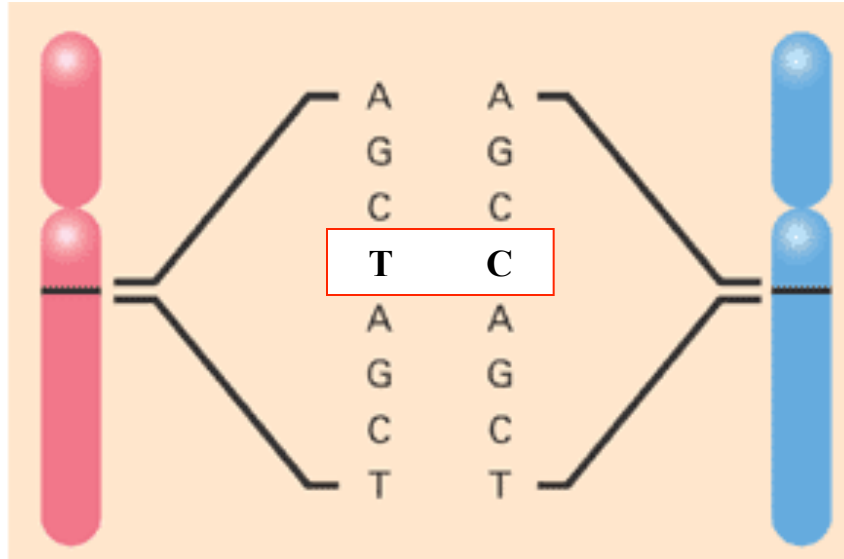
Atmospheric environments -- RS

RS/GIS was used as **an ecological approach**

To understand how the environment exerts pressures and effects on the genome such that it helps determine human traits

Genome DNA

SNPs (single nucleotide polymorphisms)

One chromosome   Another chromosome

Variations in DNA sequence

SNPs
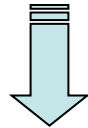
Human diversity

Definition

SNPs are **DNA sequence variations that occur** when a single nucleotide: adenine (A), thymine (T), cytosine (C) or guanine (G) - in the genome sequence is altered.

A variation must occur in at least 1% of the population to be considered an SNP.
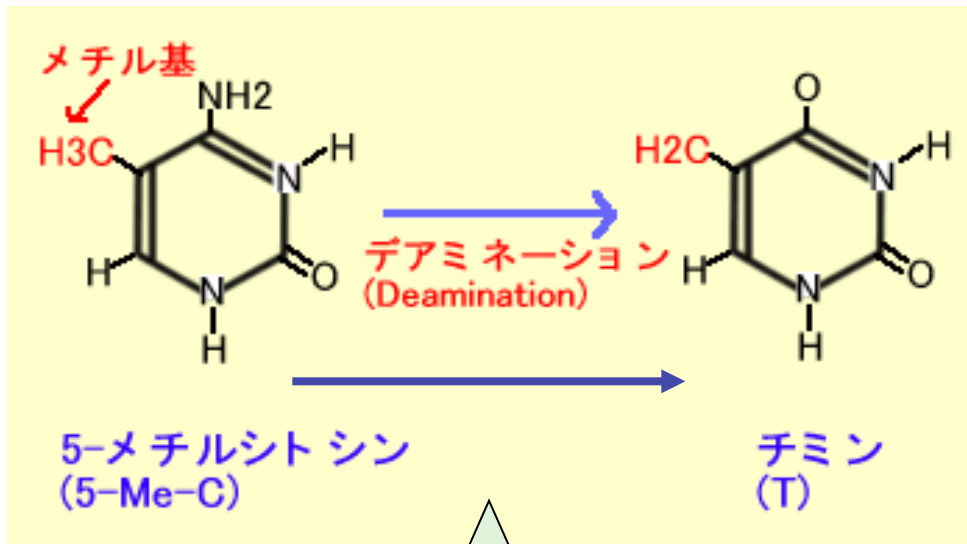
Skin color   **Eye color**

Physical constitution: Drug response, Metabolism, etc.

The mechanism of the formation of SNPs

The most case of SNPs that happen is transition; Cytosine (C) is substitute for Thymine (T) when a chemical change called deamination occurs.

メチル基
NH2
H3C
H
N
H
O
N
H
デアミ ネーション
(Deamination)
H2C
H
O
N
H
O
N
H
5-メチルシト シン
(5-Me-C)
チミン
(T)

C⇒T (Pyrimidine⇒Pyrimidine: Transition)
A,G⇔C,T（Purine⇔Pyrimidine: Transversion)

UV

260～320nm

Transition

C is substitute for T when a chemical change called deamination occurs.

The changes could be influenced by ultraviolet radiation (UVR).

The UV range absorbed by nucleotide is 260～320nm.
It is within the UV range.

8

SNP genotyping

Based on low-molecular-weight compounds known as tags

Masscode™ system for SNP genotyping
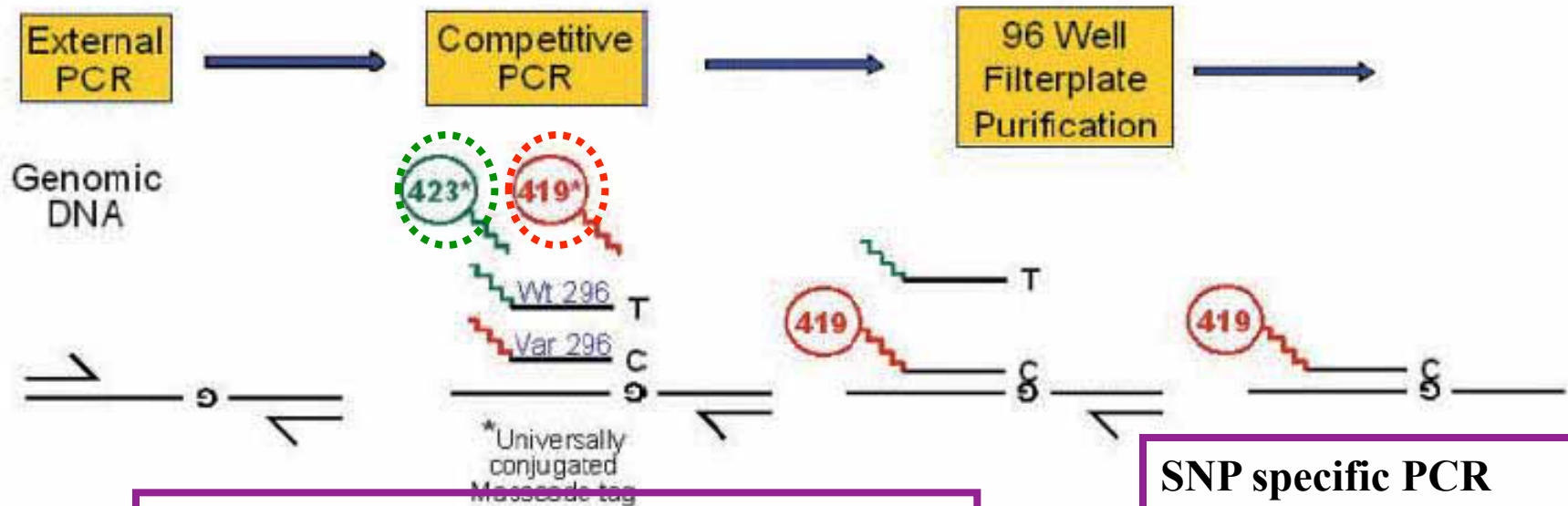
Tagged oligonucleotides are used as primers in an SNP-discrimination assay

SNP specific PCR products are purified to remove unincorporated tagged primers
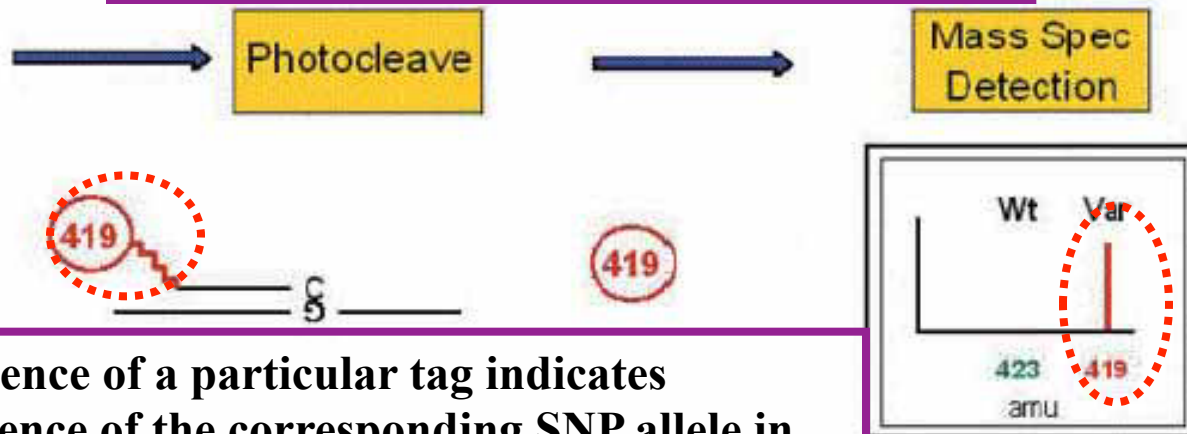
Presence of a particular tag indicates presence of the corresponding SNP allele in the genomic DNA used as the PCR template

Genotyping results

The SNP allele is classified into three types:
wild type homo, variant type homo and hetero type.
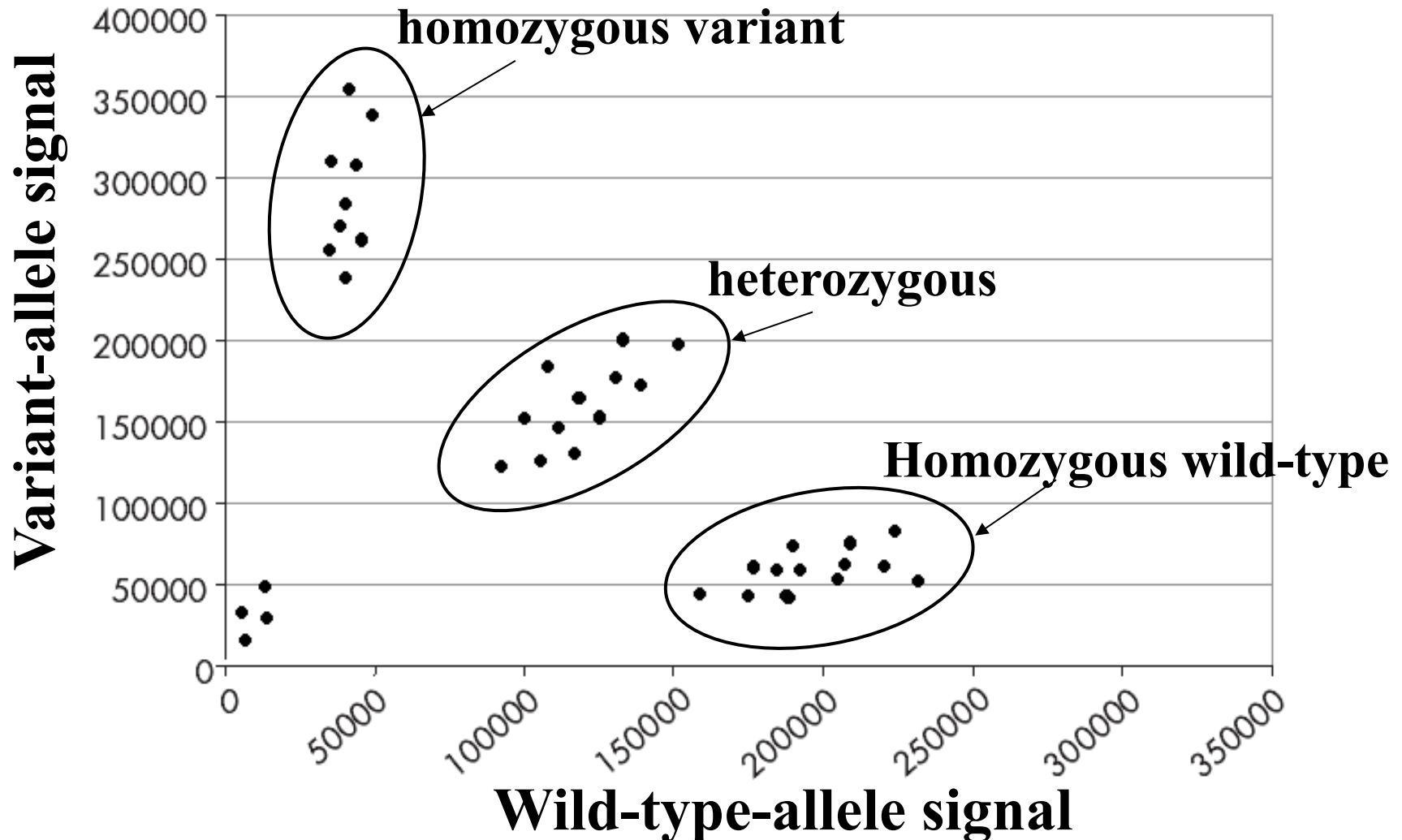
Alleles were reported using binary nomenclature, in which 1 represented wild-type alleles and 2 represented variant alleles

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		SNP ID⇒	SBA0001	SBA1003	SBA0005	SBA1007	SBA0010	SBA0012	SBA0015	SBA0017	SBA0018	SBA0019	SBA0020
2			rs819136	rs1129414	rs2075508	rs10960756	rs3793976	rs2298458	rs3212363	rs1805008	rs3212371	rs2279727	rs4778182
3		Allele Type⇒	A/?	A/?	?/T	A/?	?/T	?/T	A/T	?/T	A/?	A/?	A/?
4	Position	Sample ID											
119	2-E06	115	22	11	22	22	12	11	12	11	12	11	22
120	2-E07	116	22	11	22	22	11	11	11	11	11	11	12
121	2-E08	117	12	11	22	22	11	11	11	11	11	11	12
122	2-E09	118	22	11	22	22	11	11	11	11	12	12	22
123	2-E10	119	22	11	22	22	11	11	12	12	11	11	12
124	2-E11	120	22	11	22	22	11	12	12	11	12	11	12
125	2-E12	121	22	11	12	22	11	11	11	11	11	11	11
126	2-F01	122	12	11	22	22	11	11	11	11	12	11	12
127													
128		N11	1	120	0	0	107	115	64	108	98	112	19
129		N12	22	0	22	1	14	7	54	14	24	10	65
130		N22	99	0	100	118	0	0	4	0	0	0	38
131			122	120	122	119	121	122	122	122	122	122	122
132		XXorGray	0	2	0	3	1	0	0	0	0	0	0
133													
134		Allele Freq.(%)											
135		1	9.8	100.0	9.0	0.4	94.2	97.1	74.6	94.3	90.2	95.9	42.2
136		2	90.2	0.0	91.0	99.6	5.8	2.9	25.4	5.7	9.8	4.1	57.8
137													
138		HWE11	1.2	122.0	1.0	0.0	108.3	115.1	67.9	108.4	99.2	112.2	21.7
139		HWE12	21.6	0.0	20.0	1.0	13.3	6.8	46.2	13.2	21.6	9.6	59.5
140		HWE22	99.2	0.0	101.0	121.0	0.4	0.1	7.9	0.4	1.2	0.2	40.7
141													
142		G Freq N11	0.008	1.000	0.000	0.000	0.884	0.943	0.525	0.885	0.803	0.918	0.156
143		G Freq N12	0.180	0.000	0.180	0.008	0.116	0.057	0.443	0.115	0.197	0.082	0.533
144		G Freq N22	0.811	0.000	0.820	0.992	0.000	0.000	0.033	0.000	0.000	0.000	0.311
145			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
146													
147		A Freq 1	0.098	1.000	0.090	0.004	0.942	0.971	0.746	0.943	0.902	0.959	0.422
148		A Freq 2	0.902	0.000	0.910	0.996	0.058	0.029	0.254	0.057	0.098	0.041	0.578
149			1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

A homozygous wild-type allele was designated as 1,1 and a heterozygous allele was designated as 1,2

A molecular biological approach

1. Collected samples of 122 European and 100 East Asian

2. Extract DNA from the collected samples

3. Amplify whole genome

4. PCR to amplify the regions containing the 20 SNPs in the 7 candidate genes, encoding ASIP, TYRP1, TYR, MC1R, OCA2, MITF, MYO5A

5. SNP genotyping: Masscode™ system

Results

from data analysis with the results of SNP genotyping

Detecting natural selection in the human population would have profound implications for the study of human adaptation/ evolutionary history and for medicine (Sabeti, 2002)

Conducted Tajima's D test and Fu and Li's F test to determine whether any of the SNPs were under natural selection pressures

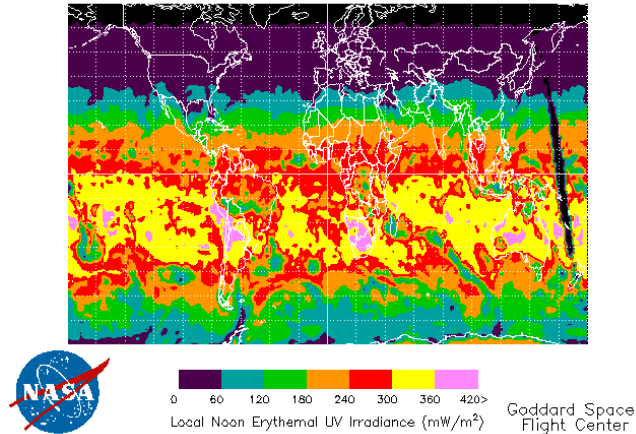**Detected natural selection in candidate pigmentation genes in haplotypes revealed by the SNP analyses**

Our results indicate the possibility that the haplotype in the OCA2 gene in the East Asians population has been under selective pressure.
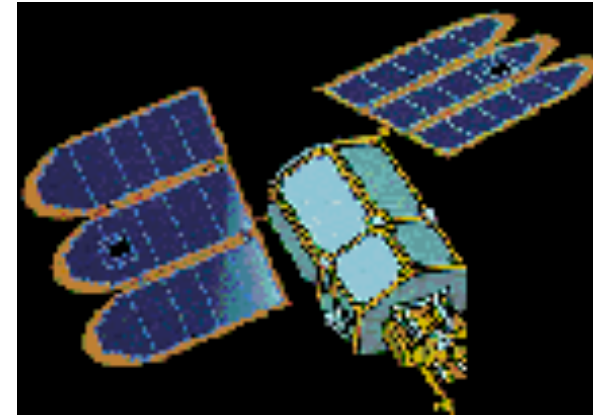
An ecological approach

Atmospheric Remote Sensing data   Earth Probe

Earth Probe TOMS Version 8 Local Noon Erythemal UV Irradiance
on January 01, 2002

NASA

| | 60 | 120 | 180 | 240 | 300 | 360 | 420> |
Local Noon Erythemal UV Irradiance (mW/m²)

Goddard Space
Flight Center

The UVR data were derived from readings taken from the NASA Total Ozone Mapping Spectrometer (TOMS), which was flown aboard the Nimbus-7/Earth Probe satellites.

The TOMS sampled single wavelengths representative of long-wave and medium-wave UVR: 324-nm and 380-nm wavelengths for UVA (range, 315-400 nm) and 305-nm and 310-nm wavelengths for UVB (range 280-315 nm).

Processing Atmospheric RS data

|	W	X	Y	Z
1	310-0101	310-0201	310-0301	ave
2	27.44444444	25.71428571	22.83333333	21.91194255
3	24.33333333	21.6	27	21.30913901
4	27	24.6	22.125	21.77637944
5	23.27272727	23.85714286	25	21.93607847
6	24.11111111	25.42857143	25.4	22.59792139
7	23.5	24.4	21.33333333	21.56429087
8	21.75	23.57142857	29	21.47188209
9	24.55555556	23.5	24.5	21.61377551
10	22.90909091	23.7	24.71428571	21.41095135
11	26.33333333	23.625	24	21.23501512

The original data set for 310-nm UVB (which induces deamination and causes a barely perceptible reddening of light skin) was read as integers into a single integer array of dimensions by using a sample C routine with longitude and latitude.

The data sets were calculated for the average reading for each month from 1979 to 2003.

Spatial Analysis in GIS
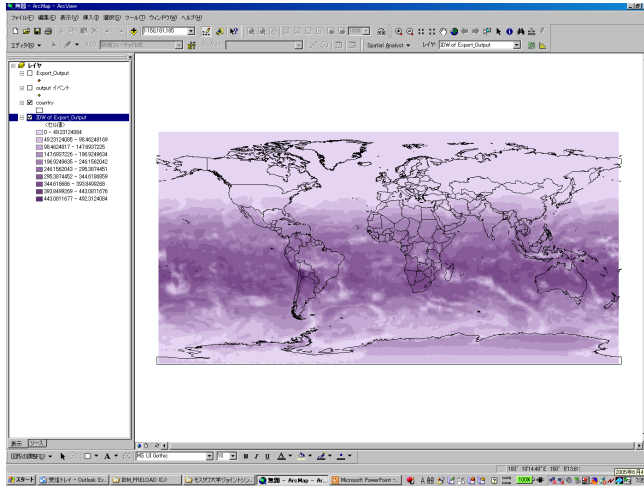
The processed data were integrated into a GIS and interpolated using inverse distance weighting (IDW) for the values of a raster.

Legend

Africans

East Asians
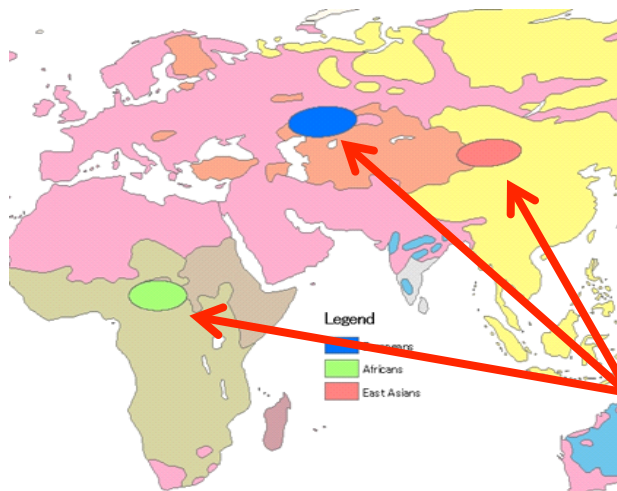
The raster layer was overlaid with the polygon layer, which contained three polygons representing the **birthplaces of the human race**.

Spatial Analysis in GIS

Each polygon contained and summarized the raster values within its area and reported the results as a table listed below.

Cell A1: 種別

	A	B	C	D	E	F	G	H	I	J
1	種別	ZONE_CODE	COUNT	AREA	MIN	MAX	RANGE	MEAN	STD	SUM
2	モンゴロイド_起源	1	73	73.0000	7.5762	18.7564	11.1802	13.2361	2.9605	966.2370
3	ネグロイド_起源	2	76	76.0000	98.8161	117.5170	18.7010	108.2350	4.3304	8225.8500
4	コーカソイド_起源	3	89	89.0000	1.4756	4.9928	3.5172	2.7215	0.8634	242.2170

Chaplin's study found that the evolution of skin reflectance could be almost fully modeled as **a linear effect of UVR in fall alone.**

The monthly mean values by the human racial population were **organized for the seasonal mean values for winter, spring, summer, and fall.**

Gene-Environment (G×E) Interaction Analysis

A molecular biological approach: Genetic engineering

UVA, UVB, etc.

Environment

An ecological approach : RS/GIS

Change

Selection

Change

Gene/Genome

Phenotype

Determination

Skin color etc.

Mutations & Variations

(Failure)

Disease

Melanoma etc.

SNPs, etc.

(Adaptive change)

Evolution

Skin color variation etc.
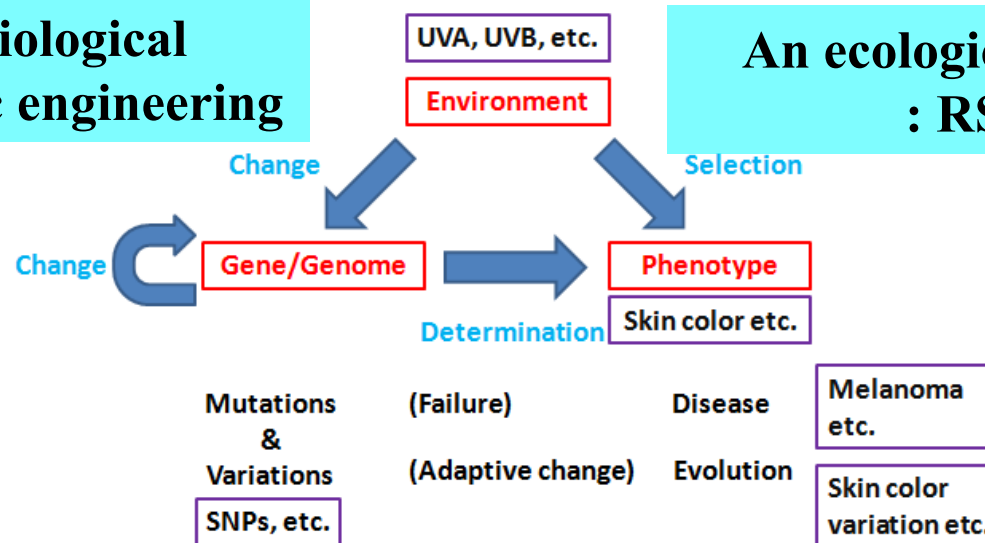
With the data obtained from these analyses based on genetic engineering/RS/GIS data, we conducted spatial statistical analysis to evaluate G×E interactions while assessing the relationship between the SNPs at candidate genes involved in human skin pigmentation and seasonal UVR exposure in the three regions denoted using polygons.

18

The data

To perform spatial statistical analysis, we used the following data:

● 20 SNPs in seven candidate genes for human skin pigmentation obtained in our previous study

● 553 SNPs in the OCA2 gene (a common gene in both European and East Asian populations in our previous study) from the HapMap database

● Seasonal UVR exposure in the three regions denoted using polygons

Spatial Statistical Analysis 1

Calculated the mean values for the heterozygous genotype frequency for every allele pattern in each population

Spatial Statistical Analysis 2

Conducted PCA using the correlation matrix to determine the relationships between **SNP frequencies for heterozygous subjects** and **the seasonal UVR data**. The following four kinds of PCA analyses were conducted:

(I)20 SNPs in seven candidate genes from our previous study + UVR data for the fall;

(II)20 SNPs in seven candidate genes from our previous study + UVR data for all four seasons;

(III) 553 SNPs in the OCA2 gene from the HapMap database + UVR data for the fall; and

(IV) 553 SNPs in the OCA2 gene from the HapMap database + UVR data for all four seasons

Results 1

Table: Mean SNP frequencies for the heterozygous alleles

Environment	UVR (310 nm, Autumn)	123.523	26.708	51.715
	20 SNP in 7 genes	Africans	Europeans	East Asians
Genome	SNP (C/T)	0.247	0.345	0.323
Genome	SNP (A/G)	0.206	0.208	0.399
Genome	SNP (G/T)	0.405	0.093	0.131
Genome	SNP (A/T)	0.545	0.481	0.413
Genome	SNP (A/C)	0.367	0.100	0.500

	553 SNP in OCA2 gene	Africans	Europeans	East Asians
Genome	SNP (A/C)	0.337	0.299	0.180
Genome	SNP (A/G)	0.260	0.234	0.187
Genome	SNP (A/T)	0.199	0.271	0.157
Genome	SNP (C/G)	0.292	0.246	0.153
Genome	SNP (C/T)	0.279	0.241	0.208
Genome	SNP (G/T)	0.251	0.155	0.137

The results showed more significant correlations in the mean G/T SNP frequencies for heterozygous alleles for the SNPs identified in our previous study than in those from the HapMap database.
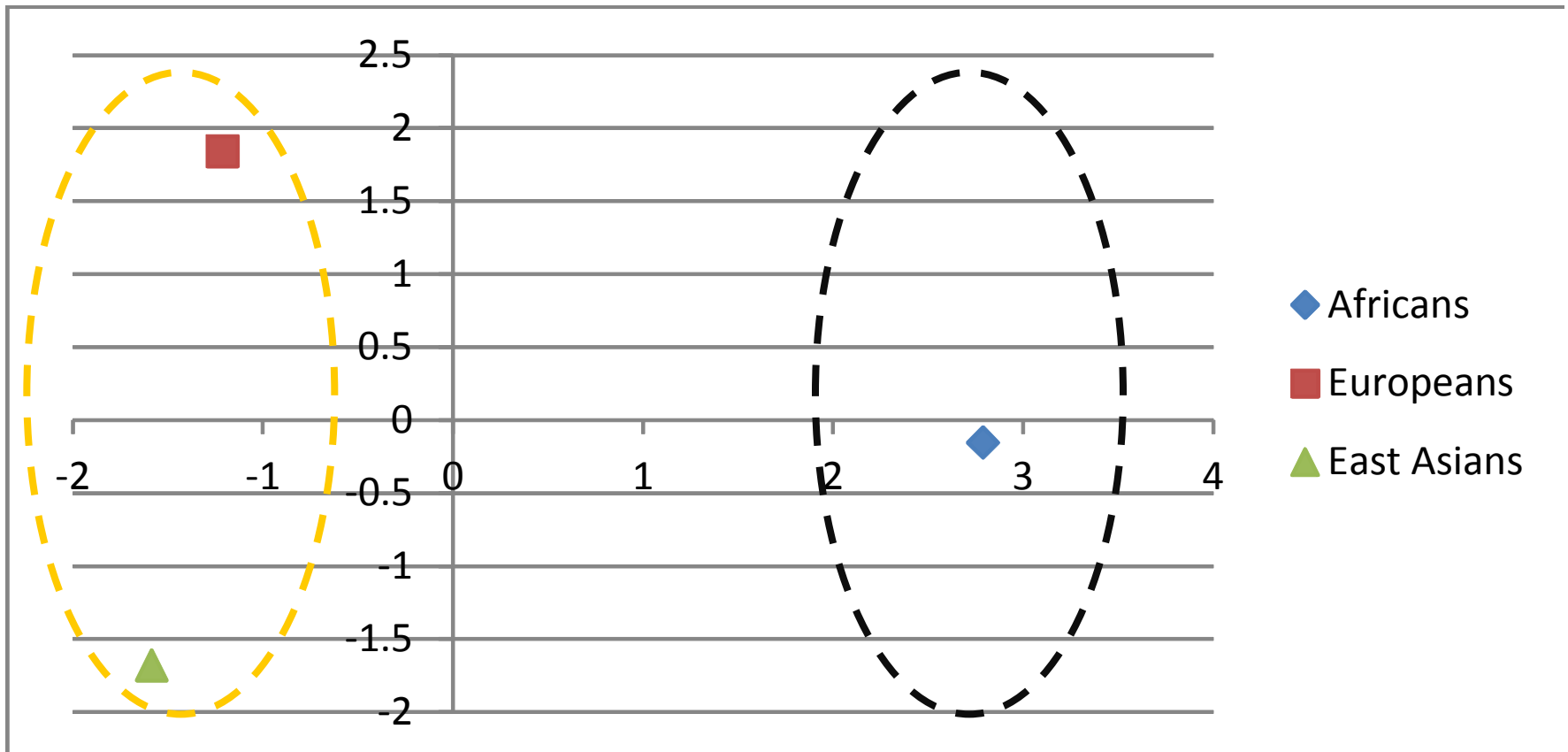
Results 2

Table: The correlation matrices for cases I and II

	UVR (Autumn)	UVR (Summer)	UVR (Winter)	UVR (Spring)
UVR (Autumn)	1.00000	0.90878	0.99078	0.99573
UVR (Summer)	0.90878	1.00000	0.84387	0.94341
UVR (Winter)	0.99078	0.84387	1.00000	0.97405
UVR (Spring)	0.99573	0.94341	0.97405	1.00000
SNP (C/T)	-0.99905	-0.88973	-0.99574	-0.99076
SNP (A/G)	-0.27792	0.14828	-0.40549	-0.18807
SNP (G/T)	0.99023	0.84172	0.99999	0.97314
SNP (A/T)	0.70147	0.34009	0.79156	0.63270
SNP (A/C)	0.42726	0.76556	0.30084	0.50888

The correlation matrix results showed **very high correlations** between the **G/T SNP** and **the mean seasonal 310-nm UVR levels**.

The mean 310-nm UVR levels in **fall** and **winter** were more highly correlated with the G/T SNP than were the mean seasonal 310-nm UVR levels in spring and summer.

Results 3

The points of the principal component scores of each distinct population were distributed separately, dividing the data into two groups: **one group of Africans** and **another group of Europeans and East Asians** for the cases I, II, and IV.

Results 4

The principal component loading results showed **a high first principal component loading of 0.98 for the G/T SNP**, which was **positively correlated to the G/T SNP variable** in case I.

The other principal component loadings were low for the other SNPs and did not correlate with other SNPs for cases II, III and IV.

This indicates **a strong relationship between the G/T SNP and fall.**

Discussion

Our results suggest the possibility that genes involved in skin pigmentation might be subject to UVR-induced mutations.

Available data also indicate that mutations possibly occurred due to modification of bases, i.e. 8-oxoG (8-oxoguanine).

8-oxoG is a potent premutagenic lesion because it can pair with adenine as well as cytosine during DNA replication and can thus cause a G:C→T:A transversion mutation.

It is thus possible that 8-oxoG formation, which is involved in active oxygen through UV exposure, may contribute to G:C→T:A transversion mutations.
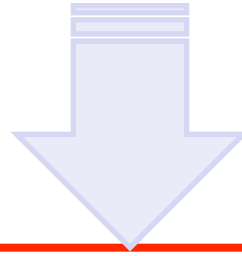
The mutations may then be passed on to the next generation, leading to skin pigmentation mutations and resulting in skin pigmentation variations.

This supports the theory that depigmented and tannable skin may have evolved numerous times in hominin evolution via independent genetic pathways under natural selection.

Conclusions

It is important to investigate human adaptability to UVR exposure to **enable better prediction of the health risks caused by extreme environmental conditions** and to **develop preventive interventions**.

Understanding human adaptability to UVR may allow us to **predict the impact of environmental changes on health risks**, particularly regarding high UV exposure secondary to ozone depletion.

Given the newly developed RS technologies that capture **more detailed**, **higher resolution** UV wavelength measurements and the emerging approaches to G×E interactions, it would clarify the mechanism underlying human adaptability to varying levels of UVR.