

## ORIGINAL ARTICLE

Yuka Yamane-Tanaka · Keiko Kogawa  
Toshihiro Tanaka Yusuke Nakamura · Minoru Isomura

## Heterozygosities and allelic frequencies of 358 dinucleotide-repeat marker loci in the Japanese population

Received: March 11, 1998 / Accepted: April 11, 1998

**Abstract** We examined 64 normal Japanese chromosomes to determine the heterozygosities and allelic frequencies of 358 dinucleotide-repeat marker loci spanning the whole human genome. Comparisons of the data for each marker in the Japanese population sample with data for the same markers among Caucasian samples in the Genome Database (GDB) revealed a slightly lower average of heterozygosity in Japanese (71% vs 79%). Although the majority of the markers were as informative as in Caucasians, some in our sample were uninformative due to low heterozygosity; 38 loci revealed heterozygosities lower than 50% and 11 of these were less than 30%. Furthermore, allelic distributions at many of the marker loci were quite different in the two racial groups. Since such differences will influence statistical analyses between markers and disease loci, our data will be essential for linkage analyses, sib-ship pair analyses, and association studies involving the Japanese population. Therefore we have archived this database on a home page on the Internet (<http://www.ims.u-tokyo.ac.jp/nakamura/Yamane.html>).

**Key words** Microsatellite · Heterozygosity · Japanese · population

### Introduction

Because polymorphic DNA markers can distinguish the paternal or maternal origin of homologous chromosomes, they are exploited widely for forensic medicine and studies

of genes associated with diseases. Among the various types of polymorphic DNA markers, microsatellites (i.e., dinucleotide-repeat loci) are the most useful because their heterozygosities in a given population tend to be much higher than 50% and because they can be analyzed in a very small amount of DNA by amplification with a polymerase chain reaction and subsequent electrophoresis. Such markers have contributed substantially to efforts to map disease loci by means of linkage analyses, sib-ship pair analyses, or association studies.

Dib et al. (1996) documented 5264 dinucleotide repeat markers in a linkage map that covers all human chromosomes. Genotypic information regarding each of these markers, such as allelic frequency and heterozygosity, is available through various databases. However, most of these data were derived from DNA analyses using Caucasian samples. Since allelic information sometimes significantly influences a statistical analysis, it is not certain that we can apply such information to studies of disease genes in other racial or ethnic groups. We previously examined the allelic frequencies of 12 dinucleotide markers on chromosome 13 in a Japanese population sample, and found significant differences between Japanese and Caucasians in the allelic distributions of two of these markers (Yamane et al. 1997).

Here we report results of an extended investigation of allelic distributions and heterozygosities at 358 dinucleotide-repeat marker loci lying an average of 10cM apart across the whole human genome, among normal Japanese subjects. The results represent an essential database for disease studies in Japanese families.

### Materials and methods

Genomic DNAs were extracted from 32 normal Japanese volunteers, according to a method described previously (Sato et al. 1990). The Perkin-Elmer Linkage Mapping Set (Perkin-Elmer, Norwalk, CT, USA) provided pairs of fluorescently labeled primers for genotyping, and PCR amplifications were carried out according to the

Y. Yamane-Tanaka · T. Tanaka · Y. Nakamura (✉) · M. Isomura  
Laboratory of Molecular Medicine, Human Genome Center,  
Institute of Medical Science, The University of Tokyo, 4-6-1  
Shirokanedai, Minato-ku, Tokyo 108-8639, Japan  
Tel. +81-3-5449-5372; Fax +81-3-5449-5433  
e-mail: yusuke@ims.u-tokyo.ac.jp

Y. Yamane-Tanaka · K. Kogawa · Y. Nakamura · M. Isomura  
Department of Human Genome Analysis, The Cancer  
Chemotherapy Center, Foundation for Cancer Research, Tokyo,  
Japan

locus symbol	Japanese	Caucasian	locus symbol	Japanese	Caucasian	locus symbol	Japanese	Caucasian	locus symbol	Japanese	Caucasian
D1S196	66	79	D2S364	75	93	D5S406	78	75	D7S531	78	70
D1S197	88	74	D2S367	91	86	D5S407	88	93	D7S550	78	89
D1S199	66	89	D2S368	91	89	D5S408	69	75	D7S636	88	93
D1S206	78	67	D2S391	84	71	D5S416	72	68	D7S640	88	85
D1S207	81	79	D2S396	78	79	D5S418	81	79	D7S657	69	78
D1S209	66	78	D3S1298	84	89	D5S419	78	82	D7S669	88	79
D1S213	78	89	D3S1266	59	79	D5S421	41	78	D7S684	78	82
D1S214	63	71	D3S1267	66	86	D5S422	66	71	locus symbol	Japanese	Caucasian
D1S216	72	85	D3S1289	72	79	D5S424	66	82	D8S258	38	86
D1S218	69	93	D3S1278	72	88	D5S426	81	89	D8S260	84	81
D1S220	78	93	D3S1580	78	93	D5S428	78	75	D8S270	56	75
D1S228	59	68	D3S1293	13	74	D5S429	88	82	D8S272	69	74
D1S234	81	82	D3S1566	94	93	D5S433	72	79	D8S277	59	71
D1S235	53	86	D3S1279	75	89	D5S436	78	68	D8S279	84	81
D1S238	75	82	D3S1262	63	78	D5S471	72	75	D8S283	69	82
D1S249	75	86	D3S1262	63	78	D5S630	97	86	D8S284	84	84
D1S252	88	89	D3S1300	69	71	D5S644	78	88	D8S285	78	81
D1S255	25	74	D3S1304	78	86	D5S647	81	86	D8S504	75	79
D1S413	56	78	D3S1569	75	71	D5S673	78	93	D8S514	66	79
D1S423	6	48	D3S1292	84	86	locus symbol	Japanese	Caucasian	D8S550	81	86
D1S424	19	75	D3S1297	75	82	D6S462	38	68	D8S556	31	80
D1S425	56	73	D3S1311	69	89	D6S289	91	64	locus symbol	Japanese	Caucasian
D1S468	63	68	D3S1565	78	64	D6S470	75	86	D9S164	78	82
D1S484	72	64	D3S1614	47	89	D6S271	0	85	D9S273	44	74
D1S498	81	78	D3S1263	97	93	D6S441	81	86	D9S157	91	84
D1S502	81	85	D3S1285	69	57	D6S308	66	62	D9S158	100	70
locus symbol	Japanese	Caucasian	D3S1271	56	75	D6S434	53	85	D9S161	56	86
D2S113	88	92	locus symbol	Japanese	Caucasian	D6S276	94	79	D9S167	81	82
D2S117	94	81	D4S1534	81	71	D6S264	41	74	D9S171	28	89
D2S125	59	86	D4S1535	81	81	D6S257	88	89	D9S175	72	86
D2S126	75	89	D4S1572	81	93	D6S292	91	82	D9S279	44	82
D2S139	84	82	D4S1575	31	48	D6S344	72	68	D9S283	72	89
D2S142	66	75	D4S1597	38	82	D6S287	63	85	D9S286	66	89
D2S151	72	86	D4S391	81	89	D6S309	81	82	D9S287	72	71
D2S160	72	78	D4S392	91	75	D6S262	78	81	D9S288	88	79
D2S162	91	75	D4S398	81	79	D6S305	78	79	D9S290	75	77
D2S164	59	93	D4S402	63	100	D6S281	78	78	locus symbol	Japanese	Caucasian
D2S165	84	96	D4S403	50	61	D6S426	81	89	D10S189	63	73
D2S168	59	75	D4S405	75	81	D6S422	59	89	D10S190	94	89
D2S206	69	86	D4S406	78	89	locus symbol	Japanese	Caucasian	D10S191	72	71
D2S281	47	81	D4S412	72	71	D7S484	81	78	D10S192	88	82
D2S285	75	89	D4S413	59	93	D7S486	63	80	D10S197	75	89
D2S286	81	67	D4S414	78	82	D7S493	78	100	D10S201	84	86
D2S305	86	68	D4S415	66	86	D7S507	81	93	D10S208	81	89
D2S319	69	75	D4S419	69	71	D7S510	69	86	D10S212	38	61
D2S325	75	79	D4S424	75	82	D7S513	81	84	D10S217	91	86
D2S326	84	82	D4S426	72	79	D7S515	78	83	D10S220	78	81
D2S337	81	100	D4S428	75	82	D7S516	78	79	D10S249	88	63
D2S338	81	78	locus symbol	Japanese	Caucasian	D7S517	94	93	D10S537	63	79
D2S347	59	79	D5S393	88	82	D7S519	63	86	D10S547	41	78
			D5S400	91	82	D7S530	69	79	D10S548	38	64

**Fig. 1** Heterozygosities among Japanese and Caucasian test populations for 358 dinucleotide-repeat marker loci. Respective heterozygosities are given in the columns to the right of each locus symbol. Shaded background indicates loci where heterozygosities deviated by more than 20% from one group to the other

manufacturer's instructions. Electrophoresis was performed with an ABI 377 DNA sequencer; the data were extracted by GeneScan Analysis software and analyzed by the Genotyper program (Perkin-Elmer). Allelic sizes for each of 358 microsatellite markers among the 64 chromosomes were determined as described previously (Yamane et al. 1997).

Genotypic data derived from Caucasians for each marker were obtained from the GDB (<http://gdbwww.gdb.org>). Data derived from a control individual (Centre d'Etudes du Polymorphisme Humain CEPH 884-15) were obtained from the CEPH database (<http://www.cephb.fr>)

## Results and discussion

Figure 1 summarizes heterozygosities calculated at each of the 358 marker loci for both the Japanese and Caucasian populations (the data for Caucasians were obtained from

the GDB). Heterozygosities in our test population varied from 0% to 100%, with a mean value of 71% (vs 79% for Caucasians; Table 1). On the basis of heterozygosities and allelic distributions at each marker, we constructed a Japanese database that is available through our www site on the Internet (<http://www.ims.u-tokyo.ac.jp/nakamura/Yamane.html>). Since the primers for PCR amplification of some dinucleotide repeat loci had been redesigned by the supplier (Perkin-Elmer), we adjusted allelic sizes in the GDB to our system on the basis of data from a known individual (CEPH 884-15) in the CEPH database. The adjusted data are reflected in Table 2 and in Tables on our web site.

Among the 358 microsatellite markers we found some that showed distinctly different heterozygosities between Japanese and Caucasian samples. For example, 89% heterozygosity was reported in Caucasians at D9S171, but heterozygosity at this locus was only 28% in Japanese. None of the 32 Japanese individuals examined was heterozygous at the locus defined by D6S271 (only a 129-bp allele was

D10S561	88	89
D10S583	81	86
D10S587	78	89
D10S591	38	75
D10S597	44	77
locus symbol	Japanese	Caucasian
D11S1313	59	89
D11S1314	78	78
D11S1320	44	56
D11S1338	63	79
D11S1358	59	68
D11S898	25	71
D11S901	75	88
D11S902	72	82
D11S904	72	86
D11S905	63	71
D11S908	50	78
D11S922	84	86
D11S934	44	84
D11S935	66	77
D11S937	84	93
D11S968	38	82
D11S987	75	79
D11S1345	56	74
locus symbol	Japanese	Caucasian
D12S310	59	69
D12S324	63	69
D12S326	69	80
D12S336	66	82
D12S345	84	87
D12S346	66	57
D12S351	63	84
D12S352	81	73
D12S364	81	87
D12S367	75	76
D12S368	59	81
D12S78	75	92
D12S79	69	96
D12S83	72	82
D12S85	84	68
D12S86	59	89
D12S99	81	83
locus symbol	Japanese	Caucasian
D13S153	100	81
D13S156	78	80
D13S158	75	81
D13S159	84	89
D13S170	88	89
D13S171	59	72
D13S173	88	82
D13S175	56	76
D13S217	78	67
D13S263	88	84
D13S265	72	70
D13S285	81	81
locus symbol	Japanese	Caucasian
D14S258	66	71
D14S261	81	68
D14S276	63	68
D14S280	66	64
D14S283	69	86
D14S288	84	86
D14S292	13	70
D14S63	69	79
D14S65	75	89
D14S68	63	82
D14S70	75	71
D14S74	91	78
D14S78	38	57
D14S80	47	82
locus symbol	Japanese	Caucasian
D15S128	75	81
D15S153	84	93
D15S117	66	85
D15S120	69	79
D15S126	66	81
D15S127	78	71
D15S130	75	71
D15S131	75	82
D15S165	34	82
D15S205	84	85
locus symbol	Japanese	Caucasian
D16S401	63	70
D16S405	72	73
D16S407	91	92
D16S411	66	79
D16S415	66	64
D16S420	34	79
D16S423	81	71
D16S503	63	75
D16S511	91	88
D16S515	72	89
D16S516	72	73
D16S520	81	86
locus symbol	Japanese	Caucasian
D17S784	66	78
D17S787	78	78
D17S791	88	88
D17S798	53	71
D17S799	69	61
D17S802	81	79
D17S808	78	79
D17S849	91	71
D17S925	72	88
D17S928	84	82
D17S938	66	89
D17S945	88	82
D17S949	78	85
locus symbol	Japanese	Caucasian
D18S452	81	82
D18S462	59	75
D18S464	56	70
D18S469	47	57
D18S474	75	82
D18S478	59	57
D18S52	69	59
D18S53	78	75
D18S57	78	89
D18S59	75	61
D18S61	88	78
D18S64	81	88
D18S68	75	89
D18S70	84	86
locus symbol	Japanese	Caucasian
D19S216	53	81
D19S209	84	89
D19S210	63	89
D19S220	81	93
D19S221	72	89
D19S226	78	75
D19S414	50	89
D19S418	63	56
D19S420	78	74
locus symbol	Japanese	Caucasian
D20S100	69	64
D20S107	78	81
D20S115	56	57
D20S117	78	88
D20S118	47	82
D20S119	69	96
D20S171	63	79
D20S173	81	75
D20S178	56	85
D20S186	91	87
D20S189	53	75
D20S195	78	82
D20S196	84	81
D20S95	66	84
locus symbol	Japanese	Caucasian
D21S1252	78	79
D21S1253	75	82
D21S1256	69	75
D21S263	72	75
D21S266	88	68
locus symbol	Japanese	Caucasian
D22S274	88	82
D22S280	84	89
D22S283	75	85
D22S315	81	96
D22S420	63	70
D22S423	75	79
locus symbol	Japanese	Caucasian
DXS1001	75	83
DXS1047	75	89
DXS1055	25	67
DXS1060	84	81
DXS1068	69	59
DXS1106	22	67
DXS1202	66	89
DXS1214	78	68
DXS1226	72	78
DXS1227	72	65
DXS986	75	73
DXS987	78	92
DXS990	84	69
DXS991	69	95
DXS993	81	86

Fig. 1 Continue

**Table 1** Comparison of heterozygosities at 358 test loci between Japanese and Caucasians

Observed heterozygosity (%)	No. of markers in each category among Japanese tested	No. of markers in each category in GDB (Caucasian)
0–20	5	0
20–30	6	0
30–40	11	0
40–50	16	2
50–60	32	10
60–70	70	42
71–80	108	114
81–90	87	161
91–100	23	29

GDB, Genome Database

observed) although heterozygosity at this locus is 85% among 30 Caucasians registered in the GDB. As Table 1 shows, in the Japanese sample, 38 marker loci revealed heterozygosity lower than 50% but only two markers did so in Caucasians; moreover, 11 of these markers in Japanese (vs none in Caucasians) showed heterozygosity lower than

30%. Although heterozygosity was higher overall in Caucasians, a few markers showed higher heterozygosity in Japanese; for example, at D9S158 we observed 100% heterozygosity in Japanese as opposed to the 70% reported in Caucasians. Our results should provide information for the selection of appropriate markers for disease analyses in Japanese families.

Many marker loci revealed quite different allelic distributions between the two populations; representative data are shown in Table 2. For instance, a 133-bp allele at the D16S423 locus accounted for 25% of Japanese alleles, but for only 3% of Caucasian alleles. On the other hand, the 137-bp and 149-bp alleles at this locus accounted for more than 70% of all chromosomes in Caucasians, and for just 8% in Japanese. At the locus defined by D14S292, an 80-bp allele was seen in 93% of Japanese chromosomes and in the remainder, allele sizes varied from 80 to 88bp. However, in Caucasians, the alleles varied from 86 to 94bp, with four alleles being most common. It is apparent that miscalculation at some loci can occur if allelic distribution information derived from Caucasians is applied to sib-ship pair analyses or association studies in the Japanese population. For

**Table 2** Allelic distributions of D16S423 and D14S292 among 64 Japanese and 60 Caucasian chromosomes

Locus: D16S423		
Allele (bp)	Japanese frequency	Caucasian frequency
131	0.0156	
133	0.25	0.036
135	0.0469	0.018
137	0.0938	0.357
139	0.0313	0.018
141	0.0469	0.018
143	0.3125	0.054
145	0.0781	0.036
147	0.0938	0.054
149	0.0313	0.357
151		0.054

  

Locus: D14S292		
Allele (bp)	Japanese frequency	Caucasian frequency
80	0.9375	
82	0.0313	
84	0.0156	
86		0.259
88	0.0156	0.278
90		0.315
92		0.013
94		0.014

Frequencies are indicated to the right of each allele size. Caucasian data were obtained from the GDB.

genetic analyses of families carrying genetic diseases, for homozygosity mapping, for sib ship-pair analyses, and for association studies, one should avoid uninformative markers and, for any statistical calculations, one should consider allelic distributions in the population under study. Hence, the results reported here provide essential data for investigating genes related to diseases in Japanese families. These results may also contribute to anthropologic research, such as investigations into the origin of the Japanese people.

**Acknowledgments** This work was supported by a Grant from the Ministry of Health and Welfare of Japan.

## References

- Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5264 microsatellites. *Nature* 380: 152–154
- Sato T, Tanigami A, Yamakawa K, Akiyama F, Kasumi F, Sakamoto G, Nakamura Y (1990) Allelotype of breast cancer: Cumulative allele losses promote tumor progression in primary breast cancer. *Cancer Res* 50: 7184–7189
- Yamane Y, Nakamura Y, Isomura M (1997) Allelic frequencies of twelve dinucleotide repeat marker loci on chromosome 13 in the normal Japanese population. *Jpn J Hum Genet* 42: 533–537