J Hum Genet (2002) 47:387–394

ORIGINAL ARTICLE

**Tomohiko Okuda · Yoshihiko Fujioka · Kei Kamide
Yuhei Kawano · Yoichi Goto · Yasunao Yoshimasa
Hitonobu Tomoike · Naoharu Iwai · Sotaro Hanai
Toshiyuki Miyata**

Verification of 525 coding SNPs in 179 hypertension candidate genes in the Japanese population: identification of 159 SNPs in 93 genes

Received: March 14, 2002 / Accepted: April 15, 2002

Abstract Single-nucleotide polymorphisms (SNPs) located in coding regions (coding SNPs; cSNPs) with amino acid substitution can potentially alter protein function. Therefore, identification of the nonsynonymous cSNPs of the genes of common diseases is valuable in tests of association with phenotypes. In this study, we validated 525 candidate cSNPs from 179 hypertension candidate genes deposited in the publicly available database dbSNP by DNA sequencing of samples from 32 Japanese individuals. We identified a total of 143 SNPs (27%) in 93 hypertension candidate genes. We also identified 16 new SNPs, for a total of 159 SNPs. Of the 159 SNPs thus identified, 104 were nonsynonymous. We estimate that approximately 20% of the SNPs deposited in dbSNP database showed a minor allele frequency of over 5%. The candidate SNPs for hypertension identified in this study would be valuable for association studies with hypertension to accelerate the identification of hypertension genes.

Key words Single-nucleotide polymorphism (SNP) · Hypertension · Japanese population · Allele frequency · Amino acid substitution · Nonsynonymous substitution · Validation of dbSNP

T. Okuda · K. Kamide · Y. Kawano · Y. Goto · Y. Yoshimasa · H. Tomoike · N. Iwai · S. Hanai · T. Miyata (✉)
National Cardiovascular Center, 5-7-1 Fujishirodai, Suita, Osaka 565-8565, Japan
Tel. +81-6-6833-5012 ext. 2512; Fax +81-6-6835-1176
e-mail: miyata@ri.ncvc.go.jp

Y. Fujioka[1]
Department of Geriatric Medicine, Ehime University School of Medicine, Ehime, Japan

Present address:
[1] Fukushima Hospital, Osaka, Japan

Introduction

Hypertension is a common disease that causes cardiovascular morbidity and mortality. In Japan, there are about 30 million patients with hypertension (Saito et al. 2000). Although it is inferred that the genetic background of each individual plays a specific role in the pathogenesis of essential hypertension, concrete responsible genetic alterations are still unclear owing to their complex nature in polygenic, heterogeneous, and multifactorial disease. Therefore, we must clarify the genetic effects that cause hypertension to improve effective diagnosis for personalized medicine (Risch and Merikanges 1996; Lander 1996; Collins et al. 1997).

Recently, attention has been focused on single-nucleotide polymorphisms (SNPs). SNPs are useful not only as markers for human genetic studies because of their high density, in every several hundred bases on average, throughout the human genome (Kruglyak 1999), but also because they make it easy to perform large-scale genotyping by high-throughput methods (Ohnishi et al. 2001). Furthermore, SNPs present in the coding region (coding SNPs; cSNPs) and promoter region (regulatory SNPs; rSNPs) of genes can potentially alter protein function and gene expression level, respectively (Collins et al. 1998; Cargill et al. 1999). In particular, nonsynonymous substitution in the translated protein probably has an impact on protein function. Thus, identification of the nonsynonymous SNPs or rSNPs of genes responsible for common diseases will facilitate progress in methods of diagnosis.

Many biallelic and multiallelic polymorphisms in genes have been identified, and various association studies with hypertension have been performed (Benetos et al. 1996; Siffert et al. 1998; Glenn et al. 2000). For instance, the Met235Thr variant of the human angiotensinogen gene has been found to be associated with an increased risk of hypertension (Jeunemaitre et al. 1992; Sato et al. 1997; Staessen et al. 1999). Therefore, the presence of functional SNPs within the gene is likely. However, the number of well-characterized genes for hypertension is still limited, and

additional genes responsible for hypertension await identification. Furthermore, certain polymorphisms reflect ethnic diversity; therefore, the identification of SNPs in potential hypertension genes must be performed by taking into account the ethnicity of the sample population.

A large number of SNPs are deposited in the public database, dbSNP (Sherry et al. 1999, 2001), at the U.S. National Center for Biotechnology Information (NCBI). The SNPs stored in dbSNP are mostly "candidate" SNPs found by computer data-mining procedures and have not been characterized. Therefore, the SNPs in dbSNP must be validated before use. We considered that to identify potential hypertension genes, we needed to target mainly nonsynonymous cSNPs.

We report here a total of 159 SNPs in 93 hypertension candidate genes, detected among 64 chromosomes in Japanese population samples. To achieve this, we selected candidate genes for hypertension, then searched for candidate cSNPs in those genes by using the publicly available database dbSNP, and finally sequenced the genes to validate the SNPs. This strategy also revealed the limited availability of candidate SNPs in the publicly available database relevant to Japanese population samples. About 73% of the SNPs in the publicly available database were monomorphic in Japanese samples, and about 5% showed low frequency (less than 5%).

Materials and methods

Selection of hypertension candidate genes and SNPs

For the selection of hypertension candidate genes, we searched the OMIM database of NCBI by using the following 19 keywords; "hypertension," "essential hypertension," "blood pressure," "insulin resistance," "renin," "aldosterone," "angiotensin," "atherosclerosis," "obesity," "hypotension," "aging," "stroke," "renal failure," "vasoconstriction," "NIDDM," "cerebral infarction," "adrenaline," "cortisol," and "coronary artery disease." In addition, 75 hypertension candidate genes previously described by Halushka et al. (1999) were added. Thus, 505 hypertension candidate genes were obtained after removing duplicates. Next, we retrieved cSNPs in hypertension candidate genes from the publicly available database by using SNPper, a CHIP Bioinformatics Tool (Riva and Kohane 2001; http://bio.chip.org:8080/bio, as of November of 2001). At this point, we had retrieved 560 candidate nonsynonymous cSNPs in 201 hypertension candidate genes for further study. We also obtained information on candidate SNPs from the JSNP (Hirakawa et al. 2002) and the SNP Consortium (TSC; http://snp.cshl.org/) databases.

Direct sequencing for verification of the cSNPs

We obtained peripheral blood samples from 32 volunteer Japanese with written informed consent. Genomic DNA was extracted with an NA-3000 nucleic acid isolation system (KURABO, Osaka, Japan). We designed polymerase chain reaction (PCR) primers using the Primer3 program (Rozen and Skaletsky 2000; http://www-genome.wi.mit.edu/genome_software/other/primer3.html). Primer pairs were designed to produce about 200 base pairs by PCR, and their Tms ranged between 60° and 63°C as far as possible. We sequenced the 32 Japanese samples, thereby permitting allele frequencies to be estimated among 64 chromosomes. The PCR reaction was performed with 20 ng of genomic DNA as the template in a 20-μl reaction mixture by using a HotStar Taq Master Mix kit (QIAGEN, Hilden, Germany) as follows: activation of *Taq* polymerase at 95°C for 15 min, followed by 35 cycles of denaturing at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s. The PCR products were then treated with shrimp alkaline phosphatase and exonuclease I (PCR Product Pre-Sequencing Kit, USB Corporation, Cleveland, OH, USA), and used as templates for direct single-pass sequencing with a BigDye Terminator v3.0 Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA, USA). The reaction products were purified with a DyeEX 96 kit (QIAGEN) and analyzed on an ABI PRISM 3700 DNA analyzer (Applied Biosystems). The obtained sequences were examined for the presence of a polymorphism by using Sequencher software (Gene Codes Corporation, Ann Arbor, MI, USA), followed by visual inspection.

Results

We selected 505 genes as hypertension candidate genes. Of these 505 genes, 367 were registered on SNPper. We focused on the nonsynonymous SNPs. Therefore, we retrieved using SNPper 560 nonsynonymous SNPs in 201 hypertension candidate genes for the sequence study. There were 115 synonymous SNPs in their vicinity that were also evaluated. We verified each SNP and its allele frequency among the 32 Japanese samples by single-pass DNA sequencing. We found that 150 SNPs failed PCR and/or sequencing, and 525 SNPs containing DNA fragments produced sequences of sufficient quality to evaluate the candidate SNP reliably (Table 1). From this analysis, we found that 382 out of 525 SNPs (73%) were monomorphic; that is, only one of two predicted alleles was found in the

Table 1. Number of analyzed genes and SNPs

	Genes	SNPs
Candidate genes	505	—
Registered in SNPper database	367	—
Retrieved candidate SNPs	201	675
Candidate SNPs with amino acid substitution		560
Candidate SNPs with no amino acid substitution		115
Candidate SNPs successfully analyzed by sequencing	179	525
SNPs identified	93	143
New SNPs		16
SNPs with amino acid substitution	65	100
New SNPs with amino acid substitution		4

SNP, single-nucleotide polymorphism

Table 2. Summary of 159 SNPs in 93 genes and their allele frequency data

Gene symbol	Reference SNP (dbSNP)	Allele 1/ allele 2	Amino acid change	Allele 1		Allele 2	Allele frequency		Hetero-zygosity	Additional information (SNP ID or flanking sequence of new SNPs)
				Homo	Hetero	Homo	Allele 1	Allele 2		
ABCC8	New SNP	C/T	NC(Asn234)	31	1	0	0.984	0.016	0.031	CTGGTGGATGAA(C/T)GCCTTCATCAAG
	rs757110	G/T	Ala1369Ser	2	12	17	0.258	0.742	0.383	TSC0101585
ACE	rs4309	C/T	NC(Pro302)	8	16	8	0.500	0.500	0.500	
	rs4331[a]	A/G	NC(Ala628)	5	15	12	0.391	0.609	0.476	IMS-JST003565
	rs4343[a]	G/A	NC(Thr673)	5	15	12	0.391	0.609	0.476	IMS-JST003570
ADD1	rs4961[b]	G/T	Gly460Trp	7	20	5	0.531	0.469	0.498	IMS-JST010969
	rs4963[b]	C/G	Ser586Cys	7	20	5	0.531	0.469	0.498	IMS-JST010968
ADRB2	rs1042713	G/A	Gly16Ala	9	16	7	0.531	0.469	0.498	
	rs1042714	G/C	Glu27Gln	0	6	24	0.100	0.900	0.180	
	rs1042718	C/A	NC(Arg175)	10	16	6	0.563	0.438	0.492	TSC1077946
AGT	rs4762	C/T	Thr207Met	23	7	1	0.855	0.145	0.248	
	rs699	C/T	Thr268Met	20	11	1	0.797	0.203	0.324	IMS-JST050962
AGTRL1	rs948847	C/A	NC(Gly45)	5	11	16	0.328	0.672	0.441	TSC0374468
AP3B1	rs42360	T/C	NC(Ala672)	25	6	1	0.875	0.125	0.219	
APOA4	rs5104	A/G	Asn147Ser	17	11	4	0.703	0.297	0.417	IMS-JST032726
APOB	rs1367117	C/T	Thr98Ile	29	2	0	0.968	0.032	0.062	TSC0514953, IMS-JST005907
	rs679899	C/T	Ala618Val	1	3	28	0.078	0.922	0.144	TSC0888425
APOC4	rs1132899	T/C	Leu36Pro	4	11	17	0.297	0.703	0.417	
	rs5167	G/T	Leu96Arg	9	12	11	0.469	0.531	0.498	
AQP2	rs426496	T/C	NC(Ser167)	4	13	15	0.328	0.672	0.441	IMS-JST074258
BCAR1	rs1035539	C/T	5′ UTR	6	11	15	0.359	0.641	0.460	TSC0339966, IMS-JST013273
C3	rs2230201	G/A	NC(Arg304)	10	13	8	0.532	0.468	0.498	
CACNA1A	rs16027	G/A	Gly889Ser	28	4	0	0.938	0.063	0.117	
	rs16051	T/C	3′ UTR	1	14	17	0.250	0.750	0.375	IMS-JST060983
	rs2304094	C/T	3′ UTR	31	1	0	0.984	0.016	0.031	IMS-JST060984
CALCA	rs5241	C/A	Ser76Arg	27	4	0	0.935	0.065	0.121	
CAST	rs754615	G/C	Val408Leu	24	8	0	0.875	0.125	0.219	TSC0145296
CCR2	rs1799864	G/A	Val64Ile	16	14	2	0.719	0.281	0.404	IMS-JST006604
CD14	New SNP	G/C	NC(Pro339)	31	1	0	0.984	0.016	0.031	AACTGCCCTCCC(C/G)CACGAGGGCTCA
CDKN1A	rs1801270	C/A	Ser31Arg	10	16	6	0.563	0.438	0.492	
	New SNP	T/C	NC(Asp35)	31	1	0	0.984	0.016	0.031	ACGCGACTGTGA(T/C)GCGCTAATGGCG
CFTR	rs213950	G/A	Val470Met	12	17	3	0.641	0.359	0.460	
CLCNKB	rs2015352	G/T	Arg27Leu	3	12	17	0.281	0.719	0.404	TSC0368577, IMS-JST052374
COL1A2	rs412777	A/C	NC(Pro482)	8	21	3	0.578	0.422	0.488	
	rs42524	C/G	Ala549Pro	29	3	0	0.953	0.047	0.089	TSC1239491
COMT	New SNP	C/G	NC(Leu136)	25	5	2	0.859	0.141	0.242	AGGGGCGAGGCT(C/G)ATCACCATCGAG
CPT2	rs1799821	G/A	Val368Ile	2	7	23	0.172	0.828	0.285	
	rs1799822	A/G	Met647Val	30	2	0	0.969	0.031	0.061	
CSF1	rs1058885	T/C	Leu408Pro	6	13	13	0.391	0.609	0.476	
CTLA4	rs231775	G/A	Ala17Thr	12	19	1	0.672	0.328	0.441	IMS-JST023322
CTNS	rs161400	T/C	Ile260Thr	27	5	0	0.922	0.078	0.144	
CYP17	rs6162	C/T	NC(His46)	8	16	8	0.500	0.500	0.500	TSC0116169, IMS-JST006036
CYP21A2	rs6474	G/A	Arg103Lys	12	17	3	0.641	0.359	0.460	IMS-JST008761
CYP27A1	New SNP	G/T	NC(Gly145)	31	1	0	0.984	0.016	0.031	CCTGACCTATGG(G/T)CCGTTCACCACG

Table 2. Continued

Gene symbol	Reference SNP (dbSNP)	Allele 1/ allele 2	Amino acid change	Allele 1		Allele 2	Allele frequency		Hetero-zygosity	Additional information (SNP ID or flanking sequence of new SNPs)
				Homo	Hetero	Homo	Allele 1	Allele 2		
DBH	rs5320	G/A	5′ UTR	23	8	1	0.844	0.156	0.264	
	rs5322	C/T	5′ UTR	31	1	0	0.984	0.016	0.031	
	rs77905	A/G	NC(Thr207)	25	6	0	0.903	0.097	0.175	TSC0032361
DHCR7	rs909217	C/T	NC(Gly424)	17	10	5	0.688	0.313	0.430	TSC0190179
EBP	rs3048	G/T	NC(Ala5)	19		13	0.594	0.406	0.482	
EDN1	rs5370	G/T	Lys198Asn	16	10	6	0.656	0.344	0.451	IMS-JST007740
EDNRA	rs5333[c]	T/C	NC(His323)	16	15	1	0.734	0.266	0.390	IMS-JST045660
	rs5334[c]	G/A	NC(Glu335)	16	15	1	0.734	0.266	0.390	IMS-JST045659
F5	rs6020	G/A	Arg513Lys	5	15	12	0.391	0.609	0.476	
	rs6018	A/C	Asn817Thr	28	4	0	0.938	0.063	0.117	TSC0275221
	rs4524[d]	A/G	Lys858Arg	20	11	1	0.797	0.203	0.324	IMS-JST011184
	rs4525[d]	A/G	His865Arg	20	11	1	0.797	0.203	0.324	IMS-JST011183
	rs6030	A/G	Met1764Val	17	13	2	0.734	0.266	0.390	TSC1238664
	rs6027	C/T	3′ UTR	28	4	0	0.938	0.063	0.117	TSC1275087
F7	rs6046	G/A	Arg413Gln	27	5	0	0.922	0.078	0.144	
FBP1	rs1042144	T/C	NC(Ala217)	10	14	4	0.607	0.393	0.477	IMS-JST051477
FGF2	rs1048201	G/A	3′ UTR	11	14	7	0.563	0.438	0.492	
FRDA	New SNP	G/A	Asp178Asn	31	1	0	0.984	0.016	0.031	GTGTACTCCCAC(G/A)ACGGCGTGTCCC
G6PC	rs161620	A/C	3′ UTR	11	12	9	0.531	0.469	0.498	
GC	rs222037	T/G	Asp432Glu	19	13	0	0.797	0.203	0.324	
	rs1047220	A/C	Lys436Thr	2	15	15	0.297	0.703	0.417	
	New SNP	C/T	Arg445Cys	30	2	0	0.969	0.031	0.061	CTGGTTAACAAG(C/T)GCTCAGACTTTGCC
GCGR	rs5384	C/T	NC(Phe365)	21	9	2	0.797	0.203	0.324	
GHR	rs6182	G/T	Cys253Phe	30	2	0	0.969	0.031	0.061	
	rs6176	C/T	NC(Ser304)	29	2	0	0.968	0.032	0.062	
	rs6183	C/A	Pro308Thr	30	1	0	0.984	0.016	0.032	
	rs6180	C/A	Leu357Ile	5	19	8	0.453	0.547	0.496	TSC0472170
	rs6184	C/A	Pro392Thr	30	2	0	0.969	0.031	0.061	
GHRHR	rs740336	C/T	NC(His188)	31	1	0	0.984	0.016	0.031	TSC0103729
GIPR	rs1800437	G/C	Glu354Gln	17	14	1	0.750	0.250	0.375	
GNB3	rs5443	C/T	NC(Ser275)	7	12	12	0.419	0.581	0.487	IMS-JST057355
GYS1	rs5447	A/G	Met416Val	23	9	0	0.859	0.141	0.242	
HCFC1	rs1051152	T/C	Ser1164Pro	12		20	0.375	0.625	0.469	IMS-JST006441
HD	rs363125	C/A	Thr1719Asn	31	1	0	0.984	0.016	0.031	
	rs362331	T/C	Tyr2308His	10	18	4	0.594	0.406	0.482	TSC0211510
	rs362272	G/A	Val2785Ile	13	17	2	0.672	0.328	0.441	IMS-JST010951
HF1	rs800292	G/A	Val62Ile	8	18	6	0.531	0.469	0.498	
	rs1061170	C/T	His402Tyr	0	5	27	0.078	0.922	0.144	
	New SNP	A/G	NC(Gly879)	31	1	0	0.984	0.016	0.031	ATAGAACACGG(A/G)ACCATTAATTCA
	rs1065489	G/T	Glu452Asp	11	14	7	0.563	0.438	0.492	
HFE	rs1799945	C/G	His63Asp	31	1	0	0.984	0.016	0.031	IMS-JST006702
HP	rs587660	T/C	Ser243Pro	29	2	0	0.968	0.032	0.062	
	rs470428	A/G	Thr372Ala	29	3	0	0.953	0.047	0.089	
HPS	New SNP	C/T	NC(Thr99)	25	7	0	0.891	0.109	0.195	TGGTGACCACAC(C/T)GAGAGCGAGGGG
IAPP	rs1800203	A/G	Ser53Gly	31	1	0	0.984	0.016	0.031	

Table 2. Continued

Gene symbol	Reference SNP (dbSNP)	Allele 1/ allele 2	Amino acid change	Allele 1		Allele 2	Allele frequency		Hetero-zygosity	Additional information (SNP ID or flanking sequence of new SNPs)
				Homo	Hetero	Homo	Allele 1	Allele 2		
ICAM1	rs5491[e]	A/T	Lys56Met	27	5	0	0.922	0.078	0.144	
	New SNP[e]	C/T	NC(Asn365)	27	5	0	0.922	0.078	0.144	CCCAGAGGACAA(C/T)GGGCGCAGCTTC
	rs5498	A/G	Lys469Glu	10	15	7	0.547	0.453	0.496	
IKBKAP	rs1538660	C/T	Pro1158Leu	17	9	6	0.672	0.328	0.441	TSC0414869, IMS-JST071690
KCNJ11	rs5219	A/G	Lys23Glu	2	12	18	0.250	0.750	0.375	
	rs5218	C/T	NC(Ala190)	6	16	10	0.438	0.563	0.492	IMS-JST001084
	rs5215	G/A	Val337Ile	2	12	18	0.250	0.750	0.375	
KLK1	rs5516[f]	G/C	Glu145Gln	1	15	15	0.274	0.726	0.398	IMS-JST060444
	rs1054713[f]	C/T	NC(Asp135)	1	15	15	0.274	0.726	0.398	TSC1656146, IMS-JST060445
	rs5517	A/G	Lys186Glu	9	16	7	0.531	0.469	0.498	
LDLR	rs5930	A/G	NC(Arg471)	10	14	7	0.548	0.452	0.495	TSC0080580, IMS-JST040310
LEPR	rs1137101	A/G	Gln223Arg	1	9	22	0.172	0.828	0.285	
	rs1805096	G/A	NC(Pro1019)	0	8	24	0.125	0.875	0.219	TSC1006465
LIPA	rs1051338	A/C	Thr16Pro	15	14	1	0.733	0.267	0.391	
	rs1051339	G/A	Gly23Arg	28	1	1	0.950	0.050	0.095	
LIPC	rs6078	G/A	Val95Met	19	11	2	0.766	0.234	0.359	
	rs6083	A/G	Asn215Ser	0	4	28	0.063	0.938	0.117	
	rs6084	C/G	NC(Thr224)	29	3	0	0.953	0.047	0.089	
LPL	rs328	C/G	Ser474Stop	25	7	0	0.891	0.109	0.195	
LYPLA1	rs1935795	A/G	5′ UTR	4	10	18	0.281	0.719	0.404	TSC1002141
MC4R	New SNP	A/T	Thr162Ser	31	1	0	0.984	0.016	0.031	CATAACATTATG(A/T)CAGTTAAGCGGG
NDUFV2	rs906807	T/C	5′ UTR	7	13	12	0.422	0.578	0.488	TSC0186343
NOTCH3	rs1044009	C/T	Ala2223Val	6	14	12	0.406	0.594	0.482	IMS-JST003420
	New SNP	G/T	NC(Ala2223)	31	1	0	0.984	0.016	0.031	GGAGTACCCGGY(G/T)GCTGGGGCACAC
NR2E3	rs1805020	A/G	Glu140Gly	16	13	3	0.703	0.297	0.417	
	rs1805021	T/C	Met163Thr	15	14	3	0.688	0.313	0.430	
	rs1805025	G/A	Val302Ile	31	1	0	0.984	0.016	0.031	
P2RY2	rs1626154	T/C	Cys334Arg	0	7	25	0.109	0.891	0.195	
PCNA	rs1050525	C/A	Ser39Arg	30	1	0	0.984	0.016	0.032	
PCSK1	rs1799904	G/A	Arg80Gln	31	1	0	0.984	0.016	0.031	
	rs6234[g]	C/G	Gln665Glu	22	8	2	0.813	0.188	0.305	
	rs6235[g]	G/C	Ser690Thr	22	8	2	0.813	0.188	0.305	
PLA2G1B	rs5634	T/C	NC(Tyr45)	24	8	0	0.875	0.125	0.219	
	rs5637	G/A	NC(Ser69)	28	4	0	0.938	0.063	0.117	
PLA2G7	rs1805017	G/A	Arg92His	23	6	1	0.867	0.133	0.231	
	rs1805018	T/C	Ile198Thr	19	10	3	0.750	0.250	0.375	IMS-JST038438
	rs1051931	T/C	Val379Ala	0	8	23	0.129	0.871	0.225	TSC0623317, IMS-JST059363
PON1	rs854560	T/A	Leu55Met	28	4	0	0.938	0.063	0.117	
	rs662	A/G	Gln192Arg	2	13	15	0.283	0.717	0.406	
PON2	rs1058082	C/G	Ala148Gly	19	13	0	0.797	0.203	0.324	
RET	rs1800858	A/G	NC(Ala45)	3	18	11	0.375	0.625	0.469	
	rs1799939[h]	G/A	Gly691Ser	31	1	0	0.984	0.016	0.031	
	New SNP[h]	C/T	NC(Pro679)	31	1	0	0.984	0.016	0.031	CCTTCCGGAGGCC(C/T)GCCCAGGCCTTC
SALL1	rs1965024	C/T	NC(Leu858)	1	14	17	0.250	0.750	0.375	TSC0375750
SCNN1B	New SNP	G/A	NC(Thr594)	30	2	0	0.969	0.031	0.061	CCAGCCTGACAC(G/A)GCCCCCCGCAGC
SELE	rs5368	C/T	His468Tyr	17	14	1	0.750	0.250	0.375	IMS-JST006027
	rs5355	C/T	Leu575Phe	28	3	0	0.952	0.048	0.092	

Table 2. Continued

Gene symbol	Reference SNP (dbSNP)	Allele 1/ allele 2	Amino acid change	Allele 1		Allele 2	Allele frequency		Hetero-zygosity	Additional information (SNP ID or flanking sequence of new SNPs)
				Homo	Hetero	Homo	Allele 1	Allele 2		
SLC18A1	rs1390938	T/C	Ile136Thr	3	12	15	0.300	0.700	0.420	TSC0561500, IMS-JST065226
SLC1A2	rs752949	G/A	NC(Pro201)	14	12	3	0.690	0.310	0.428	IMS-JST069329
SLC2A2	rs1800572	G/A	Val101Ile	30	2	0	0.969	0.031	0.061	
	rs5400	C/T	Thr110Ile	31	1	0	0.984	0.016	0.031	
	rs5398	C/T	NC(Phe479)	19	13	0	0.797	0.203	0.324	
SLC4A1	rs5035	A/C	Asp38Ala	30	2	0	0.969	0.031	0.061	
	rs5036	A/G	Lys56Glu	21	10	1	0.813	0.188	0.305	
	New SNP	G/A	NC(Ser438)	28	4	0	0.938	0.063	0.117	GATGGGAGTGTC(G/A)GAGCTGCTGATC
	rs2285644	C/T	Pro854Leu	27	5	0	0.922	0.078	0.144	IMS-JST035906
SOD2	rs1799725	T/C	Val16Ala	25	6	0	0.903	0.097	0.175	
TBXA2R	rs5745	C/T	NC(Thr81)	31	1	0	0.984	0.016	0.031	
	New SNP	G/T	Arg60Leu	31	1	0	0.984	0.016	0.031	GTTCGCACACGC(G/T)CTCCTCCTTCCT
TBXAS1	rs2286199	G/A	Ala428Thr	31	1	0	0.984	0.016	0.031	IMS-JST036645
TGFB1	rs1982073	C/T	Pro10Leu	6	15	8	0.466	0.534	0.498	TSC0929271
TNFRSF1B	rs1061622	T/G	Met196Arg	24	7	0	0.887	0.113	0.200	
TP53	rs1800370	G/A	NC(Pro36)	28	1	0	0.983	0.017	0.034	
	rs1042522	G/C	Arg72Pro	11	16	5	0.594	0.406	0.482	
TRH	rs5658	G/C	Val8Leu	3	16	13	0.344	0.656	0.451	
VWF	rs1800377	G/A	Val471Ile	18	13	0	0.790	0.210	0.331	
	rs1800378	A/G	His484Arg	6	10	16	0.344	0.656	0.451	IMS-JST010192
	rs1063856	A/G	Thr789Ala	26	6	0	0.906	0.094	0.170	TSC0445949, IMS-JST037640
	rs216321	A/G	Gln852Arg	1	13	17	0.242	0.758	0.367	IMS-JST010205
WRN	rs1346044	T/C	Cys1297Arg	24	6	0	0.900	0.100	0.180	TSC0472626, IMS-JST039413

SNPs, followed by the same letters (a–h) are in linkage disequilibrium

The following SNPs were monomorphic among 32 Japanese individuals: rs1048098, rs1048096, rs1048095, rs1048094, rs4971, rs4972, rs4962, rs11792, rs4986, rs4982, rs1800034, rs1800888, rs1042090, rs5039, rs5041, rs1805090, rs1064533, rs1801021, rs5191, rs1042860, rs1063469, rs5077, rs4882, rs5078, rs1053223, rs5080, rs5102, rs1042372, rs675, rs1864423, rs1041952, rs180743, rs40832, rs5122, rs5126, rs5164, rs5168, rs769452, rs1800557, rs1442730, rs1060788, rs5195, rs1051744, rs5196, rs5198, rs5200, rs1046248, rs1801256, rs5232, rs5234, rs1047286, rs16019, rs16026, rs16028, rs16029, rs16036, rs16052, rs5239, rs1643702, rs27672, rs13763, rs4914, rs5957, rs1803256, rs1802775, rs5881, rs5880, rs5887, rs1800072, rs1800073, rs1800074, rs1800076, rs1800075, rs1800079, rs1800080, rs1800083, rs1800084, rs1800085, rs1799834, rs1800086, rs1800087, rs1800089, rs1800090, rs1800092, rs1801178, rs1800093, rs1800094, rs1800097, rs1800098, rs766874, rs1800099, rs1800100, rs1800101, rs1800103, rs1800106, rs1800107, rs1800108, rs1800109, rs1800110, rs1800120, rs1800121, rs1800122, rs1800123, rs167037, rs5256, rs5259, rs1803289, rs5246, rs5247, rs1800211, rs1800214, rs1800215, rs1059454, rs1800220, rs1800227, rs1800228, rs1800229, rs1800230, rs368468, rs1800231, rs1800232, rs414408, rs408535, rs407903, rs1800240, rs392609, rs1800250, rs422361, rs384444, rs384472, rs1800253, rs384487, rs384497, rs418570, rs1516446, rs1800230, rs368468, rs1800231, rs1800232, rs414408, rs408535, rs407903, rs1800240, rs392609, rs1800250, rs422361, rs384444, rs384472, rs1800253, rs384487, rs384497, rs418570, rs1516446, rs712979, rs1871748, rs130015, rs1064527, rs333971, rs1065442, rs1800527, rs1804006, rs1804007, rs1049968, rs6161, rs762563, rs6445, rs1051087, rs692003, rs5319, rs5321, rs5325, rs4531, rs6271, rs4298, rs4299, rs4302, rs4303, rs4975, rs4976, rs1318506, rs760242, rs5327, rs5330, rs6272, rs155417, rs1049367, rs1049389, rs1076669, rs5798, rs5345, rs5346, rs1801710, rs5347, rs5350, rs6019, rs6025, rs6031, rs6032, rs6005, rs6011, rs6034, rs6026, rs6045, rs1769259, rs1052201, rs1052195, rs1044310, rs1062809, rs15208, rs161628, rs1050827, rs1050828, rs1050829, rs1064651, rs1042498, rs1799960, rs222038, rs5650, rs1801483, rs5385, rs1805085, rs1805086, rs1800436, rs5392, rs1802357, rs1802356, rs8986, rs5442, rs5444, rs1803692, rs1803691, rs1804593, rs5461, rs5465, rs5450, rs5451, rs1065746, rs1065747, rs363129, rs1061171, rs1060821, rs515299, rs534399, rs1800730, rs1800562, rs1802821, rs694328, rs1803751, rs1801285, rs1801286, rs5492, rs422429, rs1799969, rs5495, rs1801714, rs5497, rs5503, rs5504, rs1800563, rs1887415, rs838827, rs1802207, rs1799816, rs1801108, rs1801118, rs1801276, rs1801120, rs1801277, rs5918, rs5917, rs5921, rs646617, rs5217,rs1800467, rs5515, rs5518, rs15140, rs5931, rs5932, rs1137099, rs1805095, rs1800447, rs1800448, rs1056913, rs1801177, rs1800011, rs268, rs5934, rs5180, rs5181, rs5170, rs5171, rs1935794, rs1935793, rs1800973, rs1016862, rs630257, rs1800885, rs1800884, rs1063542, rs1050849, rs5227, rs5229, rs16139, rs5571, rs1805023, rs6232, rs1050622, rs1802258, rs241572, rs1804176, rs5632, rs5635, rs5636, rs1804181, rs1804161, rs1056929, rs5639, rs5640, rs5641, rs5642, rs5643, rs1051612, rs5670, rs20426, rs5272, rs5273, rs1212604, rs1062846, rs1044704, rs1800859, rs5713, rs1047229, rs1799980, rs1799979, rs5360, rs5361, rs5364, rs5366, rs1042109, rs30832, rs5811, rs5397, rs5434, rs5037, rs5018, rs5019, rs5022, rs5023, rs5025, rs5026, rs5557, rs1804450, rs1804451, rs1737957, rs1042017, rs1737959, rs2297507, rs1062463, rs1042854, rs5743, rs5744, rs5746, rs5747, rs5751, rs6138, rs5768, rs6137, rs5769, rs5770, rs5771, rs6140, rs4529, rs5760, rs5761, rs5762, rs4526, rs4527, rs5763, rs1800575, rs1805035, rs1800471, rs1800472, rs1042346, rs1126665, rs1804965, rs1804532, rs5659, rs1800458, rs1803083, rs1803084, rs1804117, rs1804118, rs1052853, rs1804886, rs1800387, rs1800386, rs566362, rs669884, rs1042036

NC, no amino acid change

Japanese population samples. All the SNPs identified were biallelic, and none were multiallelic. A total of 143 SNPs (27%) out of the 525 SNPs sequenced had specific allele frequency, showing more than one allele. We also identified 16 new SNPs that have not yet been registered on the dbSNP database. Out of the 159 SNPs, 104 (including 4 new SNPs) caused amino acid substitution (Table 2). In general, the new SNPs were present at low frequency. Ten new SNPs were observed in only one individual and two new SNPs were found in two individuals.

We observed perfect linkage disequilibrium between two SNPs in eight genes (*ACE*, *ADD1*, *EDNRA*, *F5*, *ICAM1*, *KLK1*, *PCSK1*, *RET*) (Table 2). The genotypes of these two SNPs were completely concordant and generated only two haplotypes. Forty-three (30%) and 30 (21%) out of the 143 SNPs were registered in the JSNP database (Release 8) or the TSC database (Release 9), respectively.

Allele frequencies of SNPs are an important characteristic in defining their utility for human genetic application. The allele frequency distribution of 159 SNPs by the frequency of the minor allele is shown in Table 3. Forty SNPs had a frequency between 1.0% and 5.0%, 56 had between 5.1% and 25.0%, and 63 had between 25.1% and 50.0%. We can estimate the prevalence of SNPs deposited in the dbSNP database by excluding the 16 new SNPs from the calculation. Thus, 10% and 12% of the SNPs deposited in the dbSNP database showed a minor allele frequency between 5.1% and 25.0% or between 25.1% and 50.0%, respectively. The allele frequencies obtained in this study were in good agreement with those in the JSNP database (Table 4).

Table 3. Distribution of minor allele frequency of verified SNPs

Minor allele frequency (%)	Number of SNPs (%)	Number of new SNPs	Total number of SNPs
1.0–5.0	27 (5)	13	40
5.1–25.0	53 (10)	3	56
25.1–50.0	63 (12)	0	63
Total	143 (27)	16	159

A total of 143 (27%) of 525 SNPs sequenced showed more than one allele. We identified 16 new SNPs that were not registered in the SNPper database. Therefore, we identified a total of 159 SNPs in this study

Table 4. Comparison of allele frequency of verified SNPs and JSNP database data

SNP rs number (dbSNP)	Polymorphism	Present data	JSNP data
rs4943	C/G	0.531/0.469	0.551/0.449
rs5104	A/G	0.703/0.297	0.661/0.339
rs1035539	C/T	0.359/0.641	0.380/0.620
rs6162	C/T	0.500/0.500	0.532/0.468
rs5370	G/T	0.656/0.344	0.704/0.296
rs4525	A/G	0.797/0.203	0.781/0.219
rs5516	G/C	0.274/0.726	0.226/0.774
rs1051931	T/C	0.129/0.871	0.092/0.908
rs5368	C/T	0.750/0.250	0.806/0.194
rs1390938	T/C	0.300/0.700	0.257/0.743
rs2285644	C/T	0.922/0.078	0.957/0.043
rs216321	A/G	0.242/0.758	0.191/0.809

Discussion

We verified the hypertension candidate SNPs deposited in the publicly available database dbSNP by single-pass DNA sequencing to see if they represented true SNPs among Japanese population samples. To achieve this, we sequenced 32 samples from Japanese individuals. A total of 143 (27%) out of 525 SNPs retrieved from the publicly available database were confirmed. Our study also provided allele frequency information on SNPs in hypertension candidate genes. Our estimates indicated that approximately 20% of the SNPs deposited in the dbSNP database showed a minor allele frequency of over 5%. Therefore, if a researcher uses the candidate SNPs in the dbSNP for a Japanese population study, there is only about a 20% chance that the SNPs have appreciable minor allele frequency. These results show the limited utility of the publicly available SNPs for the Japanese population.

There are two independent efforts to experimentally identify SNPs. TSC is a nonprofit foundation organized for identifying SNPs distributed throughout the human genome. The SNPs identified by the TSC are deposited in the publicly available database by TSC number. In Release 9 of this database, 30 SNPs (21%) of the 143 SNPs we detected were registered. In Japan, JSNP is a repository of Japanese single-nucleotide polymorphism data (Hirakawa et al. 2002), and the SNPs identified by resequencing 34 samples from anonymous Japanese individuals have been stored there. In Release 8 of this database (117,427 entries), 43 SNPs (30%) of the 143 SNPs we detected were registered. Furthermore, in the JSNP database, allele frequency information was supplemented for some SNPs by the latest release, Release 9, in January 2002. We compared the allele frequency of 12 SNPs between the JSNP data and ours (Table 4). The results showed that the allele frequency from our data was quite consistent with that from the JSNP data.

dbSNP contains more than one million submissions from about 100 registered groups describing five species, including humans (Sherry et al. 1999; Sherry et al. 2001). Submissions are divided into four categories, including those from SNP mining of the human genome project sequences (65%) and those from private investigator/corporate experimental results (28%). Public and private initiatives to discover new SNPs in humans identified more than 306,000 variations in the period 1999–2000 (Brookes et al. 2000). We retrieved candidate SNPs from dbSNP without any preliminary screening. Had we retrieved only experimentally identified SNPs, as Marth et al. (2001) did, the monomorphic SNP sites would have been less frequent and confirmed SNPs would have been more prevalent. Marth et al. (2001) reported that more than 80% of TSC and Washington University (Taillon-Miller et al. 1998) candidate SNPs are polymorphic and that approximately 50% of the candidate SNPs from these two sources are common SNPs with a minor allele frequency of more than 20% (Marth et al. 2001).

In conclusion, we identified a total of 159 SNPs in 93 hypertension candidate genes among 64 chromosomes in samples from the Japanese population. These SNPs are

extremely valuable for further association studies with hypertension.

Acknowledgments We are grateful to Dr. Tetsuro Miki at Ehime University School of Medicine for his continuous supports. We also thank Yoko Tokunaga, Chiyako Imai, and Natsuko Yamashita for their technical assistance. This study was supported by the Program for Promotion of Fundamental Studies in Health Science of the Organization for Pharmaceutical Safety and Research of Japan.

References

Benetos A, Gautier S, Ricard S, Topouchian J, Asmar R, Poirier O, Larosa E, Guize L, Safar M, Soubrier F, Cambien F (1996) Influence of angiotensin-converting enzyme and angiotensin II type 1 receptor gene polymorphisms on aortic stiffness in normotensive and hypertensive patients. Circulation 94:698–703

Brookes AJ, Lehväslaiho H, Siegfried M, Boehm JG, Yuan YP, Sarkar CM, Bork P, Ortigao F (2000) HGBASE: a database of SNPs and other variations in and around human genes. Nucleic Acids Res 28:356–360

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ, Lander ES (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet 22:231–238

Collins FS, Guyer MS, Chakravarti A (1997) Variations on a theme: cataloging human DNA sequence variation. Science 278:1580–1581

Collins FS, Patrinos A, Jordan E, Chakravati A, Gesteland R, Walters L (1998) New goals for U.S. human genome project: 1998–2003. Science 282:682–689

Glenn CL, Wang WY, Benjafield AV, Morris BJ (2000) Linkage and association of tumor necrosis factor receptor 2 locus with hypertension, hypercholesterolemia and plasma shed receptor. Hum Mol Genet 9:1943–1949

Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A (1999) Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. Nat Genet 22:239–247

Hirakawa M, Tanaka T, Hashimoto Y, Kuroda M, Takagi T, Nakamura Y (2002) JSNP: a database of common gene variations in the Japanese population. Nucleic Acids Res 30:158–162

Jeunemaitre X, Soubrier F, Kotelevtsev YV, Lifton RP, Williams CS, Charru A, Hunt SC, Hopkins PN, Williams RR, Lalouel JM, Corvol P (1992) Molecular basis of human hypertension: role of angiotensinogen. Cell 71:169–180

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139–144

Lander ES (1996) The new genomics: global views of biology. Science 274:536–539

Marth G, Yeh L, Minton M, Donaldson R, Li Q, Duan S, Davenport R, Miller RD, Kwok PY (2001) Single-nucleotide polymorphisms in the public domain: how useful are they? Nat Genet 27:371–372

Ohnishi Y, Tanaka T, Ozaki K, Yamada R, Suzuki H, Nakamura Y (2001) A high-throughput SNP typing system for genome-wide association studies. J Hum Genet 46:471–477

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Riva A, Kohane IS (2001) A web-based tool to retrieve human genome polymorphisms from public databases. Proceedings of the American Medical Informatics Association Symposium, Washington, D.C., November 3–7, 2001, pp 558–562

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Saito S, Ohnishi H, Takagi S, Shimamoto K (2000) Epidemiology of hypertension in Japanese. Nippon Rinsho Suppl 1:593–596

Sato N, Katsuya T, Rakugi H, Takami S, Nakata Y, Miki T, Higaki J, Ogihara T (1997) Association of variants in critical core promoter element of angiotensinogen gene with increased risk of essential hypertension in Japanese. Hypertension 30:321–325

Sherry ST, Ward M, Sirotkin K (1999) dbSNP — database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome Res 9:677–679

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–311

Siffert W, Rosskopf D, Siffert G, Busch S, Moritz A, Erbel R, Sharma AM, Ritz E, Wichmann HE, Jakobs KH, Horsthemke B (1998) Association of human G-protein B3 subunit variant with hypertension. Nat Genet 18:45–48

Staessen JA, Kuznetsova T, Wang JG, Emelianov D, Vlietinck R, Fagard R (1999) M235T angiotensinogen gene polymorphism and cardiovascular renal risk. J Hypertens 17:9–17

Taillon-Miller P, Gu Z, Li Q, Hillier L, Kwok PY (1998) Overlapping genomic sequences: a treasure trove of single-nucleotide polymorphisms. Genome Res 8:784–754