CELL BIOCHEMISTRY AND FUNCTION
Cell Biochem Funct 2009; **27**: 402–405.
Published online 17 July 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/cbf.1582

SHORT COMMUNICATION

HERC1 polymorphisms: population-specific variations in haplotype composition

Isao Yuasa[1]*, Kazuo Umetsu[2], Hiroaki Nishimukai[3], Yasuo Fukumori[4], Shinji Harihara[5], Naruya Saitou[6], Feng Jin[7], Prasanta K. Chattopadhyay[8], Lotte Henke[9] and Jürgen Henke[9]

[1]*Division of Legal Medicine, Faculty of Medicine, Tottori University, Yonago, Japan*
[2]*Department of Forensic Medicine, Faculty of Medicine, Yamagata University, Yamagata, Japan*
[3]*Division of Legal Medicine, Ehime University School of Medicine, Ehime, Japan*
[4]*Osaka Red Cross Blood Center, Osaka, Japan*
[5]*Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo, Japan*
[6]*Division of Population Genetics, National Institute of Genetics, Mishima, Japan*
[7]*Institute of Psychology, Chinese Academy of Sciences, Beijing, China*
[8]*Amity Institute of Forensic Sciences, Defence Colony, New Delhi, India*
[9]*Institut für Blutgruppenforschung, Cologne, Germany*

Human HERC1 is one of six HERC proteins and may play an important role in intracellular membrane trafficking. The human *HERC1* gene is suggested to have been affected by local positive selection. To assess the global frequency distributions of coding and non-coding single nucleotide polymorphisms (SNPs) in the *HERC1* gene, we developed a new simultaneous genotyping method for four SNPs, and applied this method to investigate 1213 individuals from 12 global populations. The results confirmed remarked differences in the allele and haplotype frequencies between East Asian and non-East Asian populations. One of the three common haplotypes observed was found to be characteristic of East Asians, who showed a relatively uniform distribution of haplotypes. Information on haplotypes would be useful for testing the function of polymorphisms in the *HERC1* gene. This is the first study to investigate the distribution of HERC1 polymorphisms in various populations.
Copyright © 2009 John Wiley & Sons, Ltd.

KEY WORDS — genotyping; haplotype; HERC1; single nucleotide polymorphism; population study

INTRODUCTION

The human HERC protein family consists of six proteins, which are characterized by the presence of a HECT (homologous to E6-associated protein (UBE3A) carboxyl terminus) domain and one or more RCC1 (regulator of chromosome condensation)-like domains. The HECT domain is endowed with the ability to act as ubiquitin ligases, and the RCC1-like domain stimulates guanine nucleotide exchange on ADP-ribosylation factor 1 (ARF1) and on certain members of the RAS oncogene (Rab) family of proteins, including Rab3A and Rab5. HERC1 is an unusually large protein of 532 kDa containing 4861 amino acids. HERC1 is widely expressed in many tissues and is located in the cytosol and the Golgi apparatus. HERC1 may play an important role in intracellular membrane trafficking both in the cytoplasm and Golgi apparatus.[1] HERC1 also interacts with tuberous sclerosis 2 (TSC2, tuberin), which suppresses cell growth, and results in the destabilization of TSC2.[2] The biological function of HERC1, however, has not been well defined.

The human *HERC1* gene (GenBank GeneID: 8925) spans 225 kb on chromosome 15q22 and consists of 78 exons. The NCBI single nucleotide polymorphism (SNP) database lists 19 missense, 30 synonymous, and 7 frame shift mutations in its exons. Very little is known about the associations between the polymorphisms and functions of the HERC1 protein. A few HERC1 polymorphisms in coding exons show remarked differences in allele frequencies between East Asian populations (Chinese and Japanese) and other populations (Africans and Caucasians). Recent studies using publicly available datasets have suggested that these differences have been affected by positive selection.[3–6] However, no populations other than these four have been investigated with regard to the distribution of the polymorphisms in the *HERC1* gene. As an initial study toward a fuller understanding of HERC1 variation, we developed a novel

* Correspondence to: I. Yuasa, Division of Legal Medicine, Faculty of Medicine, Tottori University, Yonago 683-8503, Japan. Tel: +81-859-38-6123. Fax: +81-859-38-6120. E-mail: yuasai@med.tottori-u.ac.jp

Received 15 April 2009
Revised 18 May 2009
Accepted 19 May 2009

Copyright © 2009 John Wiley & Sons, Ltd.

simultaneous genotyping method for four SNPs in the 3′-half genomic region of the *HERC1* gene, based on the quadruplex amplified product length polymorphism (APLP) technique and investigated the distribution of genotypes in 12 global populations to obtain allele and haplotype frequencies.

MATERIALS AND METHODS

Biological samples

DNA samples were obtained from 1213 unrelated individuals living in various areas of Eurasia and South America and were used for a population study. Sub-Saharan African (Nigerian and Ghanaian) and Turk samples were collected from people residing in Germany (Table 1). Most of these samples were from a set used in previous studies.[7,8] DNA samples from three chimpanzees were also investigated to determine the ancestral haplotype. This study was approved by the Ethical Committee at the Faculty of Medicine, Tottori University.

Genotyping procedure

In the 19 missense mutations in the database, only three SNPs are shown to have high heterozygosity values. Two SNPs are rs2228510 and rs2228511 investigated in this study. The other is rs36089909, but this was not polymorphic in our preliminary study. Although several synonymous mutations with frequency data are described in the 5′ genomic region, they are less polymorphic. Finally, four SNPs, rs2228510, rs2228511, rs2229749, and rs4523879 showing relatively high heterozygosity values in the SNP database, were selected as shown in Table 1. Three of them occurred in the coding sequence, and one was located in an intronic sequence. These SNPs were simultaneously genotyped by quadruplex PCR based on the APLP method.[7–9] The nucleotide sequence and final concentration of each primer are shown in Table 2. The PCR cocktail consisted of 50 μl of Multiple PCR Master Mix from a Multiplex PCR Kit (Qiagen, Hilden, Germany), 6.4 μl of 12 primers with a concentration of 100 pmol/μl, and 43.6 μl of water. PCR was performed in a volume of 8 μl containing 7.5 μl of the PCR cocktail and 0.5 μl of a solution containing 10–20 ng of genomic DNA. The cycle conditions were 95°C for 15 min; then 30 cycles of 94°C for 10 s, 56°C for 10 s, 72°C for 10 s, and a final extension step of 15 min at 72°C. The products were separated using a polyacrylamide gel (9%T, 5%C) together with positive and negative controls and then visualized by ethidium bromide staining.

Statistical analysis

The estimation of allele and haplotype frequencies, testing of the Hardy–Weinberg equilibrium, and analysis of molecular variance (AMOVA) were carried out using the Arlequin program ver. 3.11.[10]

RESULTS AND DISCUSSION

The APLP method requires three primers for the amplification of DNA fragments at a locus: two allele-specific primers differing in length and one common primer for the opposite

Table 1. Four SNPs and allele frequencies in 12 global populations with data from the HapMap database

			rs2228510	rs2228511	rs2229749	rs4523879
			Exon 37	Exon 45	Exon 57	Intron 73
			g.155692	g.172119	g.188939	g.210769
			c.6806A>G	c.9241A>G	c.11166C>G	INV73-289T>C
			p.I2220V	p.P3031P	p.D3722E	
No.	Population	*n*	A allele	G allele	G allele	C allele
1	African	68	0.5368	0.0956	0.0956	0.0956
2	German	92	0.4783	0.1141	0.1141	0.1141
3	Turk	80	0.6500	0.2750	0.2750	0.2750
4	Indian	107	0.6495	0.3411	0.3411	0.3411
5	Buryat	98	0.9031	0.8469	0.8469	0.8469
6	Mongolian	102	0.9559	0.8824	0.8775	0.8824
7	Korean (Kwangju)	120	0.9500	0.9208	0.9125	0.9208
8	Japanese (Tottori)	104	0.9183	0.9087	0.9087	0.9087
9	Han (Wuxi)	119	0.9370	0.8950	0.8950	0.8950
10	Han (Huizhou)	111	0.9054	0.8964	0.8919	0.9009
11	Thai	114	0.9211	0.8772	0.8684	0.8816
12	Columbian	98	0.5459	0.4745	0.4745	0.4745
	African (YRI)	59–60	0.608	0.169	0.183	0.178
	Caucasian (CEU)	58–60	0.342	0.043	0.050	0.043
	Japanese (JPT)	43–45	0.911	0.907	0.900	0.898
	Chinese (CHB)	45	0.956	0.933	0.911	0.933

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/cbf

Table 2. PCR primers for simultaneous typing of four SNPs in *HERC1* gene and their final concentrations

SNP	Primer name	Sequence (5′ → 3′)*	Final concentration (μM)
rs2228510	HERC1-E37F	TCCAGGAGACCCTATTTGTAGTC	0.2
	HERC1-E37RA	tataGTTCGGTGAAGGATtCGGAT	0.2
	HERC1-E37RG	GTTCGGTGAAGGATAaGGAC	0.2
rs2228511	HERC1-E45FA	tatAATGTGGGAGTGGtAATCCA	0.4
	HERC1-E45FG	ATGTGGGAGTGGAtATCCG	0.4
	HERC1-E45R	TCATGGCTAAGTACTTCTCCCTG	0.4
rs2229749	HERC1-E57F	CAGACCAATGTGACTAGTGCAG	0.2
	HERC1-E57RG	aatCCTGGCAATTTGATTCaTGC	0.2
	HERC1-E57RC	CCTGGCAATTTGATTaCTGG	0.2
rs4523879	HERC1-i73F	TTTCCAGCATTACCTAGTATCAGG	1.2
	HERC1-i73RT	tattaGTAGTGATGTTCAAATCGtATCAA	1.6
	HERC1-i73RC	GTAGTGATGTTCAAATCGAtTCAG	1.2

*Lowercase letters indicate non-complementary nucleotide.

DNA chain. Non-complementary nucleotides were introduced into the primers to produce a difference in length between two PCR products, to enhance the specificity of the primers, and to optimize the annealing temperature of the primers.[7–9] Figure 1 shows the band patterns of the products obtained by quadruplex PCR. The nucleotide substitutions were clearly and unambiguously detected as bands of different sizes. The sizes ranged from 55 bp for *rs2229749*C* to 100 bp for *rs4523879*T*. Simultaneous genotyping by the present quadruplex PCR method is advantageous for investigating haplotypes because the amplicons at every locus are obtained by only one amplification without sampling errors. In addition, this method is technically simple and inexpensive and does not require specialized instruments.

We genotyped four SNPs in 1213 individuals from 12 global populations. The frequencies of an allele at each locus of the *HERC1* gene are shown in Table 1. None of the Hardy–Weinberg tests were significant. The allele frequencies in our African, German, Chinese, and Japanese samples were similar to those in the SNP database. The SNPs at three loci, rs2228511, rs2229749, and rs4523879, showed remarked differences in allele frequencies between the seven East Asian and five non-East Asian populations. The frequencies for the derived alleles at these three SNP loci were higher in the central part of East Asia than in the northern and southern parts.

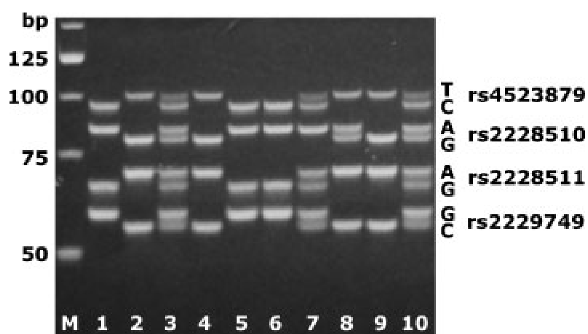

Figure 1. Simultaneous genotyping of the four SNPs by quadruplex PCR based on the APLP method. Lanes M: 25-bp ladder; 1, 5, 6: haplotypes AGGC; 2, 4, 9: GACT; 3, 10: GACT/AGGC; 7: AACT/AGGC; 8: AACT/GACT. The haplotypes estimated from the four SNPs are described in the text and Table 3

Haplotypes were constructed on the basis of the genotype data using the EM algorithm with phase-unknown samples, and three main haplotypes (AACT, GACT, and AGGC in 5′ to 3′ order of the four SNPs) were observed in all populations investigated here, indicating that the HERC1 protein was classified into three main types, Ile-Asp, Val-Asp, and Ile-Glu at the amino acid level. In addition, four minor haplotypes (AGCC, GACC, GGCC, and GGGC) were also observed at low or rare frequencies in some populations. Thus, most of the polymorphisms were in strong linkage disequilibrium, as evidenced by the low number of common haplotypes. Three chimpanzees showed the same homozygous genotype, consisting of the AACT haplotype, which was inferred to be an ancestral haplotype. Haplotypes GACT and AGGC arose by one and three mutation events, respectively, from the ancestral haplotype AACT.

As shown in Table 3, the distribution of the three main haplotypes was very diverse among various populations. Haplotype AGGC prevailed mainly in the East Asian populations, who showed relatively uniform frequencies, as the frequencies ranged from 0.91 in Koreans to 0.85 in Buryats. The Africans and Germans showed the lowest frequencies whereas the Turks, Indians, and Columbians had intermediate values. In contrast, the Africans and Germans were characterized by fairly high frequencies of haplotypes AACT and GACT. In the Columbians, the frequency of haplotype AACT was as low as its frequencies in the East Asian populations, but the frequency of haplotype GACT was as high as its frequencies in Africans and Germans. Haplotype heterozygosities were also calculated for the *HERC1* locus. East Asian populations had a mean haplotype heterozygosity of 0.2085. Haplotype heterozygosities were much higher outside of East Asia, with a mean value of 0.6162.

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/cbf

Table 3. Haplotype frequencies and heterozygosities in 12 global populations

No.	Population	*n*	AACT	GACT	AGGC	Others	Heterozygosity
1	African	68	0.4412	0.4632	0.0956		0.5859
2	German	92	0.3641	0.5217	0.1141		0.5854
3	Turk	80	0.3750	0.3500	0.2750		0.6654
4	Indian	107	0.3208	0.3380	0.3287	0.0124	0.6778
5	Buryat	98	0.0561	0.0969	0.8469		0.2715
6	Mongolian	102	0.0735	0.0441	0.8775	0.0049	0.2238
7	Korean (Kwangju)	120	0.0292	0.0500	0.9125	0.0083	0.1646
8	Japanese (Tottori)	104	0.0096	0.0817	0.9087		0.1684
9	Han (Wuxi)	119	0.0420	0.0630	0.8950		0.1941
10	Han (Huizhou)	111	0.0135	0.0856	0.8919	0.0090	0.1979
11	Thai	114	0.0439	0.0746	0.8684	0.0132	0.2393
12	Columbian	98	0.0714	0.4541	0.4745		0.5665

A high fixation index F_{ST} suggests positive selection. AMOVA showed that the estimated F_{ST} value was as high as 0.401 among the 12 populations. Population pairwise F_{ST} showed significantly differences between East Asian and non-East Asian populations. No significant differences were observed between the Africans and Germans, between the Turks and Indians, or between the East Asian populations except for between the Buryats and Koreans.

In this study, four SNPs of the *HERC1* gene in 12 global populations were simultaneously genotyped by the quadruplex PCR method. This is the first study to report on the distribution of allele and haplotype frequencies in various populations. This study confirmed that there are marked differences in allele and haplotype frequencies between East Asians and non-East Asians and showed that East Asians are deficient in genetic diversity. The frequency of haplotype AGGC has rapidly increased in East Asian populations. Recently, some East Asian-specific alleles have been demonstrated to be under local positive selection: *ABCC11* is responsible for ear wax type[11] and *EDAR* is associated with Asian hair thickness.[12] The HERC1 protein is classified into at least three main types of Ile-Asp, Val-Asp, and Ile-Glu. The frequencies of these types suggest the difference between Asp and Glu at amino acid position 3722 to be important, but nothing is currently known regarding the effect of the differences in amino acids on the function of the HERC1 proteins. Information on haplotypes would be useful in the design of biochemical experiments aimed at testing the function of polymorphisms in the *HERC1* gene. These data will also be informative and facilitate genetic association studies of HERC1-related diseases. Together with the clarification of the significance of the mutations in the HERC1 protein, the accumulation of data on the worldwide distribution of HERC1 polymorphisms using our genotyping method may provide additional insights into the role of this protein and the remarked differences in the distribution of its alleles and haplotypes.

ACKNOWLEDGEMENTS

This study was supported in part by a Grant-in-Aid for Scientific Research (20590678 to IY) from the Japan Society for the Promotion of Science.

REFERENCES

1. Garcia-Gonzalo FR, Rosa JL. The HERC proteins: functional and evolutionary insights. *Cell Mol Life Sci* 2005; **62**: 1826–1838.
2. Chong-Kopera H, Inoki K, Li Y, *et al*. TSC1 stabilizes TSC2 by inhibiting the interaction between TSC2 and the HERC1 ubiquitin ligase. *J Biol Chem* 2006; **281**: 8313–8316.
3. Carlson CS, Thomas DJ, Eberle MA, *et al*. Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Res* 2005; **15**: 1553–1565.
4. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. *PLoS Biol* 2006; **4**: e72.
5. Sabeti PC, Varilly P, Fry B, *et al*. Genome-wide detection and characterization of positive selection in human populations. *Nature* 2007; **449**: 913–918.
6. Johansson Å, Gyllensten U. Identification of local selective sweeps in human populations since the exodus from Africa. *Hereditas* 2008; **145**: 126–137.
7. Yuasa I, Umetsu K, Harihara S, *et al*. OCA2*481Thr, a hypofunctional allele in pigmentation, is characteristic of northeastern Asian populations. *J Hum Genet* 2007; **52**: 690–693.
8. Yuasa I, Nakagawa M, Umetsu K, *et al*. Molecular basis of complement factor I (CFI) polymorphism: one of two polymorphic suballeles responsible for CFI A is Japanese-specific. *J Hum Genet* 2008; **53**: 1016–1021.
9. Watanabe G, Umetsu K, Yuasa I, Suzuki T. Amplified product length polymorphism (APLP): a novel strategy for genotyping the ABO blood group. *Hum Genet* 1997; **99**: 34–37.
10. Excoffier L, Laval G, Schneider S. Arlequin ver 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005; **1**: 47–50.
11. Yoshiura K, Kinoshita A, Ishida T, *et al*. A SNP in the ABCC11 gene is the determinant of human earwax type. *Nat Genet* 2006; **38**: 324–330.
12. Fujimoto A, Kimura R, Ohashi J, *et al*. A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. *Hum Mol Genet* 2008; **17**: 835–843.

Copyright © 2009 John Wiley & Sons, Ltd.