ORIGINAL PAPER

A protocol comparison for the analysis of heat shock protein A1B +A1538G SNP

**Carmen L. Contreras-Sesvold ·
Nyamkhishig Sambuughin · Andrei Blokhin ·
Patricia A. Deuster**

Abstract Heat shock proteins act as molecular chaperones, assist in peptide maturation, and transport nascent peptides across membranes. One commonly studied single nucleotide polymorphism (SNP) for one of the proteins is HSPA1B (+A1538G). However, several studies of this polymorphism have failed to achieve Hardy–Weinberg equilibrium (HWE) for their sample. We compared various published procedures for analyzing the HSPA1B +A1538G SNP and report reasons for HWE discrepancies. Samples from 141 apparently healthy, physically active, volunteers (99 men and 42 women) were analyzed. The first protocol, initially described by Schröder et al., resulted in a genotypic distribution of 22 GG (15.6%), 119 AG (84.4%), and 0 AA; results were confirmed by reanalysis and sequencing. Two other published protocols, one described by Klausz et al. and another by Fekete et al., were used to confirm these results: both resulted in 22 GG (15.6%), 46 AA (32.6%), and 73 AG (51.7%). Additionally, the results were within HWE and confirmed by sequence analysis. Of the original 119 subjects genotyped as AG by the Schröder protocol, 46 of those were confirmed as AA with the Klausz and Fekete methods. Mixing primers from the Schröder and Klausz protocol resulted in 100% concordance with the data generated by the Klausz and Fekete protocols. Some published data on HSP genotyping deviate from HWE; thus, primers used for analyzing these highly homologous genes must be carefully considered. Our results highlight the importance of reinvestigating data when HWE is not achieved for the HSPA1B, or another, polymorphism.

Keywords HSPA1B · HSP70-2 · Heat shock protein · Hardy–Weinberg equilibrium · Protocol comparison · Genetic validation

C. L. Contreras-Sesvold (✉) · P. A. Deuster
Department of Military and Emergency Medicine,
Uniformed Services University of the Health Sciences,
4301 Jones Bridge Road,
Bethesda, MD 20184, USA
e-mail: csesvold@usuhs.mil

N. Sambuughin · A. Blokhin
Department of Anesthesiology,
Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

Introduction

The human heat shock protein A1B (HSPA1B), a member of the HSP70 multigene family, encodes a highly conserved 70-kD protein with diverse functions. Genetic variations in this gene have been investigated, and research suggests that alterations may affect immune regulation and/or functional responses to stress (Maytin 1995; Favatier et al. 1997; Jalbout et al. 2003). Genetic variations in HSP70 have also been associated with individual susceptibility to several diseases through alterations in protein expression and/or function. In particular, the A to G single nucleotide polymorphism (SNP) at position +1538 in HSPA1B gene (GenBank accession # NM_005346; rs1061581) has been associated with pancreatitis, renal failure, Crohn's disease, and various carcinomas (Milner and Campbell 1990; Schroeder et al. 1999; Jalbout et al. 2003; Balog et al. 2005; Klausz et al. 2005).

Table 1 Forward and reverse primers for the Schröder, Klausz, Fekete, and mixed primer protocols

Protocol	Forward primer	Reverse primer
Schröder	5′ GTGCTCCGACCTGTTCCGAAGC 3′	5′ CGGAGTAGGTGGTGAAGATCTG 3′
Klausz	5′ TCCGAAGGACTGAGCTCTTG 3′	5′ CAGCAAAGTCCTTGAGTCCC 3′
Fekete	5′ ACCCTGGAGCCCGTGGAGAA 3′	5′ CTGCACCGCCGCCCCGTAGG 3′
MP1	5′ TCCGAAGGACTGAGCTCTTG 3′	5′ CGGAGTAGGTGGTGAAGATCTG 3′
MP2	5′ GTGCTCCGACCTGTTCCGAAGC 3′	5′ CAGCAAAGTCCTTGAGTCCC 3′

The validity of selected HSP70 genotyping data and their association with disease has been called into question because the genetic distribution was not in Hardy–Weinberg equilibrium (HWE) (Nemeth et al. 2004). Our initial evaluation of a HSPA1B SNP yielded such results suggesting a methodological error (Bogardus et al. 1999; Xu et al. 2002). Thus, the analyses were repeated, with sequencing and verification with several primer sets.

In this report, analyses of the HSPA1B A1538G polymorphism among a group of apparently healthy, physically active controls are presented. The objective is to describe results derived from comparing three published procedures for analyzing the HSPA1B +A1538G polymorphism and emphasize the importance of using HWE for distribution analysis and verification.

Materials and methods

Human subjects

Apparently healthy, physically active volunteers ($n=141$) participated in a study previously described (Heled et al. 2007). The study was approved by the Institutional Review Board of the Uniformed Services University of the Health Sciences. Informed consent was obtained from all participants prior to the beginning of the study. Blood samples for DNA were collected for analysis of selected SNPs as part of a larger study.

Genetic analysis

Blood was collected in EDTA-containing tubes and then centrifuged at low speed to separate the cells and plasma. DNA was extracted from the buffy coat by using the QIAamp DNA mini kit 250; (Qiagen, Valencia, CA). The initial polymerase chain reaction (PCR) conditions for the HSPA1B +1538 A/G SNP were as described by Schröder et al. (Schröder et al. 2003). Table 1 describes the primers for this procedure. Table 2 describes the PCR conditions and amplicon length for this protocol. Amplicons were subsequently digested with the restriction endonuclease Pst1 as follows: each 25-µl digestion contained 5 µl of PCR product, 1 µl (20 U) Pst I, 2.5 µl NEB buffer #3, 0.25 µl BSA (New England Biolabs, Ipswich, MA, USA), and 16.25 µl PCR Grade Water (Fisher BioReagents, Fairlawn, NJ, USA). Allelic polymorphisms were identified by the presence or absence of the Pst I recognition site. Amplicons with the Pst I recognition site had the G allele, while those that lacked it were identified as the A allele. Imaging of digestion products shows the A allele band at 383 bp and G allele bands at 244 and 139 bp. All reactions resulting with the AG genotype were repeated and subsequently sequenced for validation by using forward primer with the ABI PRISM® BigDye® Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, Inc. Foster City, CA, USA) as per instructions. Amplicons and their digestion products were detected by 2% agarose gel electrophoresis. Subsequent documentation was performed with the Molecular Imager Gel Doc XR System (Bio-Rad, Hercules, CA, USA).

Based on the results from Schröder et al. (2003), two other methods were selected for genotyping our samples: the methods of Klausz et al. (Klausz et al. 2005) and Fekete et al. (Fekete et al. 2003). These two methods were selected because (1) they have been used by others (Chouchane et al. 1997); (2) the primers for the Klausz protocol, but initially described by Chouchane et al. (1997), were located outside the Schröder primer location; and (3) the primers of

Table 2 Polymerase chain reaction conditions for Schröder, Klausz, Fekete, and mixed primer protocols

| Protocol	Denature	Annealing	Extension	Cycles	Amplicon length (bp)
Schröder	96°C 1 min	57°C 1 min	72°C 1 min	30	383
Klausz	96°C 20 s	60°C 20 s	72°C 90 s	30	2,075
Fekete	94°C 30 s	64°C 30 s	72°C 30 s	40	189
MP1	96°C 20 s	60°C 20 s	72°C 90 s	30	1,385
MP2	96°C 20 s	60°C 20 s	72°C 90 s	30	1,075

Fekete were located within the amplicon generated by the Schröder primers. Figure 1 presents the orientation of the three primer sets, and Table 1 presents the primer sequences for all three protocols. The Klausz protocol, which initially generated a 2,075-bp amplicon, yielded a band at 2,075 bp for the A allele and bands at 1,140 and 935 bp for the G allele following digestion (Klausz et al. 2005). All reactions were sequenced by using the forward Klausz primer for genotype validation.

The Fekete protocol (Fekete et al. 2003) resulted in a 189-bp amplicon. A typographical error was noted for their reverse primer, which was corrected as indicated in Table 1. Furthermore, the annealing temperature was raised to 64°C to increase primer specificity. The initial amplicon produced, which was 189 bp long, yielded two additional products following digestion: one band at 189 bp for the A allele and two bands at 116 and 73 bp for the G allele. All samples identified as heterozygotes by the Schröder protocol were reanalyzed by these two methods.

Finally, to confirm our results, additional mixed primer experiments (MP) were performed using the forward primer of one method (Schröder or Klausz) with the reverse primer of either Klausz or Schröder (Table 1). Table 2 describes the PCR conditions, and digestions were performed as described above (Klausz et al. 2005). Using the Klausz forward (Kl-f) and Schröder reverse primers (Sc-r) or the MP1 protocol generated a 1,385-bp amplicon, with subsequent digestion generating a 1,385-bp product for the A allele and bands of 1,139 and 246 bp for the G allele. In contrast, using the Schröder forward (Sc-f) and Klausz reverse primers (Kl-r), or the MP2 protocol generated a 1,075-bp amplicon, whereupon following digestion yielded a 1,075-bp for the A allele and 138 and 937 bp for the G allele.

Results

Initial results with the Schröder protocol indicated that of the 141 subjects tested, 22 or 15.6% were homozygous for the G allele and 119 or 84.4% were heterozygous; none
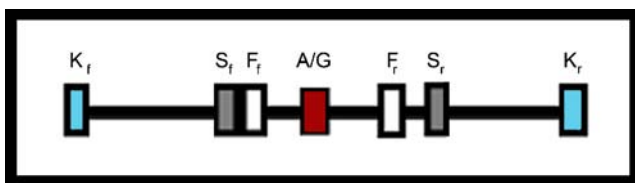
Fig. 1 Depicts the orientation of the Klausz forward (K_f), Klausz reverse (K_r), Schröder forward (S_f), Schröder reverse (S_r), Fekete forward (F_f), Fekete reverse (F_r), primer sets with respect to the heat shock protein A1B +A1538G SNP (A/G)

of the samples were homozygous for the A allele, which resulted in a failure to meet HWE (Chi square=75.15; p=0.000). Sequencing with the forward primer confirmed these genotypes and resulted in reanalysis with other previously published primers to resolve the discrepancy.

The Klausz protocol resulted in a different distribution. Of the 141 subjects tested, 22 or 15.6% were homozygous for the G allele, 46 or 32.6% were homozygous for the A allele, and 73 or 51.7% were heterozygous. This distribution was in HWE (Chi square=0.62; p=0.43). The genetic distribution for the control population described by Klausz was 13.1% GG, 39% AA, and 48% AG: it was similar and in HWE (Chi square=0.071; p=0.79; Klausz et al. 2005).

The Fekete, MP1, and MP2 genotype results mirrored those generated by the Klausz protocol, and the observed distributions were all in HWE (Chi square=0.0840; p=0.77). All samples that converted from AG by the Schröder protocol to AA by the Klausz and Fekete protocols were sequenced as described above and confirmed to be AA. Figure 2 presents genotype results for five subjects derived from the five different primer sets. Table 3 shows the genotype distribution by Schröder and Klausz protocols.

A sequence homology search using the basic local alignment search tool (BLAST) was performed using the amplicon sequences generated Schröder, Klausz, and Fekete protocols. The BLAST results for Schröder and Fekete sequences revealed a 99% sequence match to another HSP family member (HSPA1A), whereas no significant match was found for DNA sequence fragments amplified using Klausz primers, which included in its amplicon the primer sequences for both Fekete and Schröder.

Discussion

In the present study, HSPA1B fragments containing the +A1538G polymorphism were initially amplified using primers and protocols described by Schröder et al. (1999, 2003); however, the results, although comparable, were not in HWE. Deviation from HWE required reanalysis of the samples with another method. We reanalyzed the data with the protocol described by Klausz et al. (Klausz et al. 2005), which had primer sequences outside the area of sequence homology with other known HSPs, Fekete et al. (2003) and two mixed primer protocols. All of the new results were in HWE, such that 46 samples of the original 119 subjects genotyped as AG by the Schröder protocol were genotyped as AA with the Klausz and Fekete methods as well as mixed primer protocols. These results were validated by sequencing,
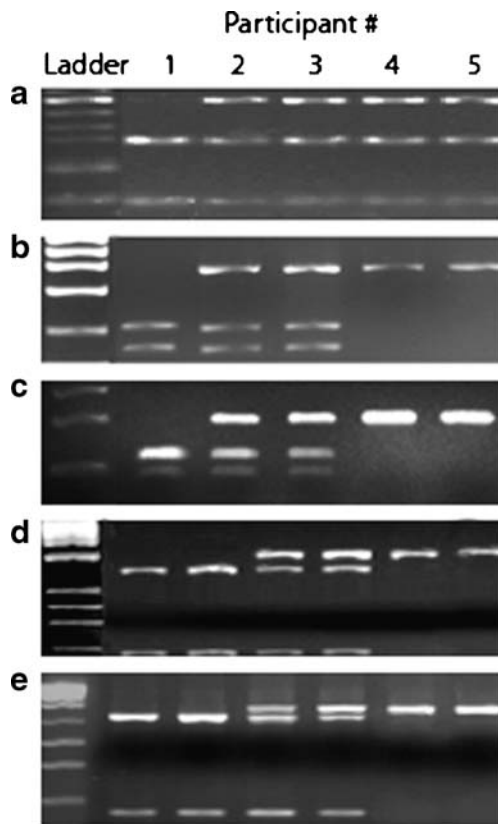
Fig. 2 The images represent Pst1 digestion products (2% agarose gels) of amplicons generated by the (*A*) Schröder, (*B*) Klausz, (*C*) Fekete (*D*), mixed primer 1 (MP1), and (*E*) mixed primer 2 (MP2) primer pairs. (*A*) Participant #*1* is a GG genotype. Participants #*2*, #*3*, #*4*, and #*5* are all AG genotypes. (*B*) Participants #1, #2, #3 genotypes remain GG, AG, and AG, respectively. Participants #4 and #5 have converted genotypes from AG to AA with the Klausz primers. (*C*) Participants #1, #2, #3 genotypes remain GG, AG, and AG, respectively. Participants #4 and #5 have converted from AG to AA genotypes, consistent with the Klausz protocol. MP data demonstrate Schröder primers are non-specific and may be generating amplicons for more than one gene. (*D*) Participant #1 remains a GG. Participants #2 and #3 are both AG (banding at 246 bp is not visible in this photograph). Participants #4 and #5 genotypes are AA. (*E*) Participants #1, #2, #3 genotypes remain GG, AG, and AG respectively. Participants #4 and #5 have converted from AG to AA genotypes with this primer set

such that we found 100% concordance with the data generated by Klausz et al. (data not provided).

BLAST search of the Klausz primer pairs revealed no match with other DNA sequences, whereas BLAST analysis of DNA sequence fragments of HSPA1B suggested that the Schröder and Fekete primers were nonspecific and could possibly amplify another sequence on a homologous gene, HSPA1A (GenBank accession # NM_005345.5; National Center for Biotechnology Information 2008). Unlike the Schröder assay, the Fekete primers became more specific when the annealing temperature was raised from 61°C to 64°C. Regardless of annealing temperatures below 62°C, nonspecific amplification by the Schröder assay, as demonstrated in our analyses, resulted in an excess of AG genotypes and deviation from HWE. Thus, we strongly suggest using the Klausz method because of its high primer specificity.

Our results highlight the importance of reinvestigating data when HWE is not achieved. Two different published protocols were used to reanalyze the discordant data, resulting in significantly different results. The primers used for analyzing HSPA1B polymorphism must be carefully considered as the sequence near this variant is highly homologous to other HSP genes (Bonnycastle et al. 1994; National Center for Biotechnology Information 2008). HSPs high homology and multiple isoforms contribute to an increased probability of nonspecificity. Thus, stringency must be emphasized when assessing any HSP polymorphism. Further, genetic databases contributed to the confusion by identifying similar sequences with the same name but has now been corrected (Kampinga et al. 2009). There is also confusion in the literature with regard to this polymorphism as two nucleotide locations (A1267G and A1538G) are reported for the same SNP (Milner and Campbell 1990; Fekete et al. 2003; Schröder et al. 2003; Klausz et al. 2005).

In summary, genotyping errors can be generated anywhere along the analysis process, from DNA harvesting to data analysis; a reliable method to check for these errors is through analysis of the distribution for HWE. In some cases, deviation from HWE may imply genetic associations of clinical relevance, but in many cases, it may suggest problems with sample processing, bias in allele frequency estimates, population stratification, and/or failure in one or more HW assumptions (Salanti et al. 2005). In conclusion, results that deviate from HWE require verification with another method to establish data reliability and validity.

Table 3 Heat shock protein A1B (+1538) genotypic distribution by protocol and Hardy–Weinberg equilibrium

Protocol (*n*=141)	Genotype (%)			Chi square
	AA	AG	GG	
Schröder	0	119 (84.4)	22 (15.6)	75.15; *p*=0.000
Klausz	46 (32.6)	73 (51.7)	22 (15.6)	0.621; *p*=0.43
Fekete	46 (32.6)	73 (51.7)	22 (15.6)	0.621; *p*=0.43

References

Balog A, Gyulai Z et al (2005) Polymorphism of the TNF-alpha, HSP70-2, and CD14 genes increases susceptibility to severe acute pancreatitis. Pancreas 30(2):e46–e50

Bogardus ST Jr, Concato J et al (1999) Clinical epidemiological quality in molecular genetic research: the need for methodological standards. JAMA 281(20):1919–1926

Bonnycastle LL, Yu CE et al (1994) Cloning, sequencing, and mapping of the human chromosome 14 heat shock protein gene (HSPA2). Genomics 23(1):85–93

Chouchane L, Ahmed SB et al (1997) Polymorphism in the tumor necrosis factor-alpha promoter region and in the heat shock protein 70 genes associated with malignant tumors. Cancer 80(8):1489–1496

Favatier F, Bornman L et al (1997) Variation in hsp gene expression and Hsp polymorphism: do they contribute to differential disease susceptibility and stress tolerance? Cell Stress Chaperones 2 (3):141–155

Fekete A, Treszl A et al (2003) Association between heat shock protein 72 gene polymorphism and acute renal failure in premature neonates. Pediatr Res 54(4):452–455

Heled Y, Bloom MS et al (2007) CK-MM and ACE genotypes and physiological prediction of the creatine kinase response to exercise. J Appl Physiol 103(2):504–510

Jalbout M, Bouaouina N et al (2003) Polymorphism of the stress protein HSP70-2 gene is associated with the susceptibility to the nasopharyngeal carcinoma. Cancer Lett 193(1):75–81

Kampinga HH, Hageman J et al (2009) Guidelines for the nomenclature of the human heat shock proteins. Cell Stress Chaperones 14 (1):105–111

Klausz G, Molnár T et al (2005) Polymorphism of the heat-shock protein gene Hsp70-2, but not polymorphisms of the IL-10 and CD14 genes, is associated with the outcome of Crohn's disease. Scand J Gastroenterol 40(10):1197–1204

Maytin EV (1995) Heat shock proteins and molecular chaperones: implications for adaptive responses in the skin. J Invest Dermatol 104(4):448–455

Milner CM, Campbell RD (1990) Structure and expression of the three MHC-linked HSP70 genes. Immunogenetics 32(4):242–251

National Center for Biotechnology Information (2008) HSPA1A heat shock 70 kDa protein 1A (*Homo sapiens*) National Library of Medicine

Nemeth E, Vasarhelyi B et al (2004) Unreported deviations of genotype distributions from Hardy-Weinberg equilibrium in articles published in Critical Care Medicine between 1999 and 2003. Crit Care Med 32(6):1431–1433

Salanti G, Amountza G et al (2005) Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. Eur J Hum Genet 13(7):840–848

Schröder O, Schulte KM et al (2003) Heat shock protein 70 genotypes HSPA1B and HSPA1L influence cytokine concentrations and interfere with outcome after major injury. Crit Care Med 31 (1):73–79

Schroeder S, Reck M et al (1999) Analysis of two human leukocyte antigen-linked polymorphic heat shock protein 70 genes in patients with severe sepsis. Crit Care Med 27(7):1265–1270

Xu J, Turner A et al (2002) Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? Hum Genet 111(6):573–574