Polymorphisms in human heat shock factor-1 and analysis of potential biological consequences

Tiffany M. Bridges · Rachel G. Scheraga · Mohan E. Tulapurkar · Dante Suffredini · Stephen B. Liggett · Aparna Ramarathnam · Ratnakar Potla · Ishwar S. Singh · Jeffrey D. Hasday

Received: 24 November 2013 / Revised: 24 June 2014 / Accepted: 30 June 2014 / Published online: 16 July 2014
© Cell Stress Society International 2014

Abstract The stress-activated transcription factor, heat shock factor-1 (HSF1), regulates many genes including cytoprotective heat shock proteins (HSPs). We hypothesized that polymorphisms in *HSF1* may alter the level or function of HSF1 protein accounting for interindividual viability in disease susceptibility or prognosis. We searched for exomic variants in *HSF1* by querying human genome databases and directly sequencing DNA from 80 anonymous genomic DNA samples. Overall, *HSF1* sequence was highly conserved, with no common variations. We found 31 validated deviations from a reference sequence in the dbSNP database and an additional 5 novel variants by sequencing, with allele frequencies that were 0.06 or less. Of these 36, 2 were in 5′-untranslated region (5′UTR), 10 in 3′UTR, and 24 in the coding region. The potential effects of 5′UTR on secondary structure, protein structure/function, and 3′UTR targets of microRNAs were analyzed using RNAFold, PolyPhen-2, SIFT, and MicroSNiper. One of the 5′UTR variants was predicted to strengthen secondary structure. Eight of 3′UTR variants were predicted to modify microRNA target sequences. Eight of the coding region variants were predicted to modify HSF1 structure/function. Reducing HSF1 levels in A549 cells using short hairpin RNA (shRNA) increased sensitivity to heat-

induced killing demonstrating the impact that genetic variants that reduce HSF1 levels might have. Using the pmirGLO expression system, we found that the wild-type HSF1 3′UTR suppressed translation of a firefly luciferase reporter plasmid by 65 %. Introducing two of four 3′UTR single nucleotide polymorphisms (SNPs) increased HSF1 3′UTR translational suppression by 27–44 % compared with the wild-type HSF1 3′UTR sequence while a third SNP reduced suppression by 25 %. *HSF1* variants may alter HSF1 protein levels or function with potential effects on cell functions, including sensitivity to stress.

Keywords Heat shock · Heat shock transcription factor-1 · Single nucleotide variation · MicroRNA

Introduction

The heat shock response is a highly conserved ancient biological process that is essential for survival against a myriad of environmental stresses. This response is accompanied by reprogramming of the cellular transcriptional and translational machinery to preferentially express a set of stress-inducible proteins, including seven families of extensively conserved heat shock proteins (HSPs) (Feder and Hofmann 1999). While prokaryotic and eukaryotic HSP genes exhibit striking cross-domain homology, they use different mechanisms of transcriptional regulation. In eukaryotes, HSP expression is regulated at the transcriptional level by a heat-activated transcription factor called heat shock factor (HSF).

Mammals express three HSFs, HSF1, 2, and 4 (Nakai et al. 1997; Rabindran et al. 1991; Schuetz et al. 1991), of which HSF1 is the heat/stress-responsive orthologue of the single heat-inducible HSF expressed by nonvertebrate Metazoans (Rabindran et al. 1991). A recently discovered mammalian orthologue of avian HSF3 was discovered in mouse but it was

T. M. Bridges · R. G. Scheraga · M. E. Tulapurkar · D. Suffredini · A. Ramarathnam · R. Potla · I. S. Singh · J. D. Hasday
Pulmonary and Critical Care Medicine Division, Department of Medicine, University of Maryland School of Medicine, 110 S. Paca St. 2nd floor, Baltimore, Maryland, USA

I. S. Singh · J. D. Hasday (✉)
Medicine and Research services of the Baltimore VA Medical Center, Baltimore, MD 21201, USA
e-mail: jhasday@umaryland.edu

S. B. Liggett
Departments of Medicine and Molecular Physiology and Pharmacology, University of South Florida, Tampa, FL 22612, USA

only found to regulate atypical HSPs and did not affect the classical heat shock response or survival during heat stress (Fujimoto et al. 2010).

All Metazoan HSFs, including mammalian HSF1, have a conserved architecture with an N-terminal DNA-binding domain that binds to a conserved nGAAn pentanucleotide (Perisic et al. 1989) and two adjacent hydrophobic regions, HR-A and B in the N-terminal half of the molecule (Fig. 1). The C-terminal half of HSF1 contains a third HR, HR-C, a transactivation domain (TAD), and a regulatory domain that confers inducibility to TAD function (Green et al. 1995; Peteranderl and Nelson 1992). In response to stress, conversion from HSF1 monomers that lack stable DNA-binding activity to intranuclear DNA-binding HSF1 homotrimers has been consistently found to be critical for initiating the heat shock response (Sarge et al. 1993). HSF1 monomers are likely maintained by intramolecular binding of HR-A/B to HR-C. HSF1 trimerization is achieved by transitioning to an intermolecular triple-stranded coiled-coil structure comprising the HR-A/B domains on oligomerizing HSF1 molecules (Zuo et al. 1994). The HRs of HSF1 also mediate its association with several other proteins with potential functional consequences for both HSF1 and its binding partners (Singh et al. 2009a) (Reviewed).

Although mammalian HSF1 was originally identified as a thermal stress-activated transcriptional activator of HSP genes, subsequent studies have identified a much broader range of biological functions for HSF1 than previously thought. The concept that HSFs might have additional functions was initially suggested by Westwood et al. (Westwood et al. 1991) who used in situ hybridization analysis to show that exposure to heat stimulated the recruitment of HSF to 150 distinct chromosomal loci in *Drosophila* salivary gland polytene chromosomes, far more than could be accounted for by the known *Drosophila* HSP genes. These observations were subsequently complemented by Trinklein et al. (Trinklein et al. 2004) who used a combination of chromatin immunoprecipitation and human promoter microarray analyses to show recruitment of human HSF1 to multiple non-HSP genes in K562 cells.

Studies with the HSF1 knock-out mouse not only confirmed that HSF1 is the major regulator of the heat shock response (McMillan et al. 1998) but also demonstrated that

HSF1 participates in the regulation of extra-embryonic development, growth, and endotoxemia-induced systemic inflammation (Xiao et al. 1999), female (Xiao et al. 1999) and male (Izu et al. 2004) reproductive potential, the ubiquitin proteolytic pathway (Pirkkala et al. 2000), post-natal brain development (Santos and Saraiva 2004), maintenance of olfactory epithelium (Takaki et al. 2006), ciliary beating in the respiratory epithelium, ependymal cells, oviduct, and the trachea (Takaki et al. 2007), and tumorigenesis (Dai et al. 2007). Gene-specific studies have shown that HSF1 can modify the expression of various cytokines, chemokines and acute response genes, including pro-interleukin-1 β , c-fms, c-fos, TNF α , IL-6, IL-8, CXCL5, the pro-apoptotic factor, XIAP-associated factor 1 (XAF1), iNOS, cyclooxygenase-2 and the zinc finger AN1-type domain-2a gene, and arsenite-inducible RNA-associated protein (AIRAP) (Goldring et al. 2000; Inouye et al. 2007; Rossi et al. 2012; Rossi et al. 2010; Singh et al. 2000; Singh et al. 2008; Wang et al. 2006; Xie et al. 2002). Additional studies, using complementary DNA (cDNA) microarrays to analyze the gene expression pattern activated by heat exposure confirmed that such exposure also modifies expression of several non-HSP genes, including those involved in regulation of transcription, growth, DNA repair, apoptosis, signaling, and cytoskeletal function (Murray et al. 2004). More recently, Mendillo et al. (2012) showed that HSF1 was activated under basal conditions in cancers with high tumorigenic and metastatic potential but not in other cancers and, using high-throughput chromatin immunoprecipitation-sequencing (ChIPseq), that HSF1 was recruited to about 500 genes, many of which are distinct from those induced by heat exposure and some of which are down-regulated by HSF1.

In addition to its transcriptional activating activities, HSF1 may also exert additional biological effects by binding to and modifying function of proteins involved in diverse cellular processes, including HSPs (Singh et al. 2009a), the nuclear pore-forming TPR protein through which HSP72 is secreted (Skaggs et al. 2007), other transcription factors (Singh et al. 2009b; Xie et al. 2002), components of the TFIIB transcription complex (Yuan and Gurley 2000), the cell division cycle protein, Cdc20 (Lee et al. 2008), the apoptosis modulator DAXX (Charette et al. 2000), and the multidrug exporter, Ral-binding protein (Ralbp)-1 (Singhal et al. 2008).

Fig. 1 Schematic of HSF1 protein based on Swiss-prot accession no. Q00613. DNA-binding, regulatory, and transactivation domains, and hydrophobic regions (HR) A and B and HR-C indicated. Amino acid numbers are shown above

Considering the central participation of HSF1 in so many important biological functions, it is surprising that so little is known about genetic variations in elements of the human heat shock response, especially HSF1, and the potential impact on human health and disease. While specific studies of human HSF1 genetic variation have not yet been performed, genome-wide association studies (GWAS) have linked 8q24.3, the location of the HSF1 gene, to schizophrenia, bipolar disorder (Wang et al. 2010), and attention deficit hyperactivity disorder (Lasky-Su et al. 2008), suggesting HSF1 genetic variation may contribute to neuropsychiatric disease. Recently, Li et al. (2011) reported two novel single nucleotide polymorphisms (SNPs) in the bovine *HSF1* gene that are disproportionately associated with thermal tolerance in Chinese Holstein cattle, including one SNP in the 3'-untranslated region (3' UTR) that disrupts a potential microRNA (miRNA)-binding sequence in *HSF1* (Li et al. 2011) and is disproportionately associated with heat tolerance in Chinese Holstein cows. Huang et al. (2012) found a SNP in a mature microRNA, miR-608, that is predicted to weaken its interaction with its target sequence in human *HSF1* 3'UTR and which is associated with increased human epithelial growth factor receptor-2 (HER2)-positivity and larger size of human breast cancers. While a reduction in HSF1 protein expression that might be caused by reduced miRNA targeting is consistent with HER2-positivity and larger tumors, the effect of this SNP on HSF1 protein level has not been confirmed experimentally.

To begin to understand the potential biological and clinical importance of genetic variants in *HSF1*, we analyzed the human *HSF1* gene for variation compared to a reference sequence by analyzing existing databases and performing exomic sequencing from anonymous genomic DNA samples from 80 healthy Caucasian, African American, and Asian subjects. The identified variants were analyzed in silico to predict potential functional ramifications and the effect of 3' UTR SNPs on HSF1 protein levels, and the impact of varying HSF1 protein levels on stress resistance was tested.

Methods

DNA sequencing DNA was obtained from Human Variation Panel of the Coriell Institute (<http://ccr.coriell.org/nigms/cells/humdiv.html>) representing 80 apparently healthy individuals of whom 26 described themselves as African American, 24 as Asian, and 30 as Caucasian as described (Small et al. 2006). The 13 exons were amplified using PCR using Taq™ high fidelity (Invitrogen, Grand Island, NY) with primer pairs that contained the forward and reverse M13 sequences to facilitate sequencing. Nine primer pairs were designed to generate amplification products between 302 and 781 nt containing 1 to 3 exons and intervening intron sequence and 60 to 200 nt 3'- and 5'-flanking intronic sequence. All reactions were

performed using 35 cycles comprising 30 s of denaturation at 94 °C, 30 s of annealing, and 90 s of elongation at 68 °C with a 7-min final extension at 68 °C. Some reactions required the addition of 7.5–10 % dimethyl sulfoxide (DMSO) to facilitate amplification. The primer sequences, annealing temperature used, and requirement for DMSO for each reaction are summarized in Table 1. PCR products were purified (Qiagen™) and sequenced using dye-terminator-based automated DNA sequencing and M13 primers on an Applied Biosystems model 3730XL 96-capillary high-throughput sequencer in the University of Maryland Sequencing Core. Sequences were read automatically by the Applied Biosystems software and confirmed by visual inspection of the chromatograms using Finch TV™ (Geospiza; Seattle, WA) and by comparing the results of bidirectional sequencing of each sequence. Variations as defined by a reference sequence (GenBank accession no. NT_037704) were identified using BLAST (National Center for Biotechnology Information).

Variant discovery from an existing database We searched the dbSNP, 1000 Genomes, HapMap, and NHLBI Exome Sequencing Project (ESP) databases for HSF1 exomic variants. Specifically, we searched the dbSNP (build 138 phase I) (Sherry et al. 2001) database for variants in the HSF1 messenger RNA (mRNA) sequence (NM005526.2) within contig NT037704.5 (Nusbaum et al. 2006). We used HSF1 as the search term to query the 1000 Genomes Project database (Abecasis et al. 2012) using the 1000 Genomes browser, the merged HapMap phase I–III database (The International HapMap Project 2003) using the HapMap Genome Browser Release #28, and the NHLBI ESP Exome Variant Server (version 0.0.22) (Exome Variant Server NGSPE).

In silico analysis of potential variant effects The potential effect of 5'UTR variations on secondary structure of 5'UTR was analyzed using the RNAFOLD™ program version 2.0 using both the minimal free energy (Zuker and Stiegler 1981) and centroids (Gruber et al. 2008) models. The potential effect of 3'UTR on microRNA target sequences was analyzed using the MicroSNiPer™ program (Barenboim et al. 2010). The potential effect of nonsynonymous variants on protein function was estimated using the PolyPhen-2 (Adzhubei et al. 2010) and SIFT (Kumar et al. 2009) prediction programs.

Generation of A549 clones with reduced HSF1 expression The A549 cell line was purchased from the American Type Cell Collection (ATCC; Manassas, VA) and maintained in RPMI 1640 supplemented with 50 U/ml penicillin, 50 µg/ml streptomycin, 2 mM L-glutamine, 1 mM sodium pyruvate, 10 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) buffer (Invitrogen), pH 7.3 (CRPMI), and containing 10 % defined fetal bovine serum

Table 1 PCR reaction conditions used for exomic sequencing

Reaction no.	5' primer sequence	3' primer sequence	Product size (nt)	Exon (s)	Annealing temperature (°C)	DMSO (%)
1	5'-TGTAACACGACGGCCAGT-CGCTTGGGGTGGAGGGA	5'-CAGGAACAGCTATGACC-CGCGTAAAGTGAAGTGGGGC	524	1	56	10
2	5'-TGTAACACGACGGCCAGT-TGCCTGGTTTCAGAGGGGC	5'-CAGGAACAGCTATGACC-GGACGCCCGTCTTCATCA	302	2	56	0
3	5'-TGTAACACGACGGCCAGT-CGCTCTTCAGGGGTTCTGGTCCC	5'-CAGGAACAGCTATGACC-ACCGGGGGCTCTTGGTGG	567	3, 4	56	0
4	5'-TGTAACACGACGGCCAGT-GCTCTGCATGGGGGACAGG	5'-CAGGAACAGCTATGACC-GCAGGAGGTCCATGCAAGTGC	608	5, 6, 7	56	0
5	5'-TGTAACACGACGGCCAGT-AGTCCGGAGACAGGGCACC	5'-CAGGAACAGCTATGACC-AGCCCCACCTCATCCCCAACTC	319	8	56	0
6	5'-TGTAACACGACGGCCAGT-CCTCAATGAGGGAGGCCCTTG	5'-CAGGAACAGCTATGACC-CCAGCTGCCCTGCTCAGTCC	473	9	56	0
7	5'-TGTAACACGACGGCCAGT-GCAGGGCAGAGCCCTCCAC	5'-CAGGAACAGCTATGACC-GGGAGGGAGTCAAGACAGCACC	486	10	56	10
8	5'-TGTAACACGACGGCCAGT-CCGGGTGCTGTCTGACTTC	5'-CAGGAACAGCTATGACC-GAGAAGTAGGAGCCCTCTCCAGCTC	428	11, 12	56	7.5
9	5'-TGTAACACGACGGCCAGT-CAGAGAACAGCAGCCGGATTACG	5'-CAGGAACAGCTATGACC-GCAGCCGGCCCTTCTGTG	781	13	58	7.5

(FBS; Hyclone, Logan, UT) at 37 °C in 5 % CO₂-enriched air. A549 cells were transfected with plasmids expressing either short hairpin (sh)RNA targeting human HSF1 or negative control shRNA plasmid (SABiosciences; Valencia, CA) using Lipofectamine (Invitrogen). Transfected cells were selected with puromycin (2 µg/ml) and cloned by limiting dilution.

Susceptibility to heat-induced cytotoxicity Wild-type A549 cells and clones expressing HSF1-targeting shRNA or a scrambled shRNA were grown to confluence in 96-well plates, then exposed to 45 °C in a forced air incubator for 0.5 to 2.0 h, and returned to a 37 °C incubator for the remainder of the 24 h. The monolayers were washed three times with PBS, stained with 0.2 % crystal violet (w/v in 2 % ethanol) for 15 min, washed five times with water, and solubilized in 100 µl 1 % (w/v) SDS for 45 min, and absorbance was measured at 570 nm using a Molecular Devices plate reader (Sunnyvale, CA). Absorbance of heat-exposed monolayers was normalized to the absorbance in the 37 °C monolayer for each clone.

Effect of HSF1 SNPs on translational efficiency The entire 391 nt 3'UTR of human HSF1 was amplified by PCR using Taq HiFi, 2 mM Mg(SO₄)₂, cDNA reverse transcribed from SAEC total RNA, and the primers 5'-GACAGACA-GAGC TC-cagtgcagggtggtctt and 5'-GACAGACA-TCTAGAGcatagacgtgtctgtttatagat. The gel-purified (Qiagen) product was double digested with SacI and XbaI and cloned into the multicloning site of pmirGLO (Promega). Individual SNPs were introduced into the wild-type pmirGLO-HSF1 plasmid (pmirGLO-HSF1-wt) by site-directed mutagenesis (Quikchange II™; Agilent; Santa Clara, CA) using the manufacturer's protocol. All plasmids were confirmed by bidirectional sequencing. HEK 293 cell monolayers were transfected with 10 ng of plasmid DNA and 36 h later were lysed using a commercial lysis buffer (Passive Lysis Buffer; Promega) and Renilla and firefly luciferase activities measured using commercial reagents (Promega) and a Victor III luminometer (Perkin-Elmer). The ratio of firefly:renilla activities was calculated for each sample.

Western blotting Cell extracts were prepared in RIPA buffer containing protease inhibitors (Roche Diagnostics; Indianapolis, IN), clarified by centrifugation at 15,000×g, and protein was assayed using the Bradford method (Biorad; Hercules, CA) using a bovine serum albumin standard curve. A sample containing 20 µg of total protein was resolved by SDS-PAGE, electrostatically transferred to polyvinylidene difluoride membrane for immunoblotting. The membranes were blocked with 5 % nonfat dry milk in TTBS (25 mM Tris pH 7.4, 0.5 M NaCl, 0.05 % Tween-20) and probed for HSF1 and β-tubulin using antibodies from Santa Cruz Biotechnology (Santa Cruz, CA) and Millipore/Chemicon (Billerica, MA), respectively.

Bands were developed using HRP-conjugated secondary antibody (BioRad) and ECL detection system (Renaissance™; New England Nuclear; Boston, MA), and quantified using a gel documentation system (Fuji LAS-1000) as described earlier (Singh et al. 2008).

Statistics Data are displayed as mean±SEM. Differences among multiple groups over time was analyzed by MANOVA using the JMP 9 statistics package (SAS Institute; Cary, NC). The effects of 3'UTR SNPs on pmirGLO activity were compared using Mann-Whitney tests.

Results

DNA from a group of 80 healthy subjects comprising 30 Caucasians, 26 African Americans, and 24 Asians was analyzed for *HSF1* variants by bidirectional exomic sequencing. Regions of genomic DNA were amplified using high-fidelity polymerase, and all variants were confirmed by re-amplifying from the original DNA sample and resequencing. With this sample size, we had a 90 % probability of detecting a variant with an allele frequency of >0.014 at least once (Liggett 2003). We found eight variants in the *HSF1* exomic sequence in the 80 DNA samples (Table 2). Seven of these variants had mean allele frequencies (MAF) <1 %. Five of these variants were not found in any of the databases analyzed and three were found in multiple individuals. All three variants that were found in multiple individuals were found only in individuals of the same race. Of the eight variants identified, two were located in the 5'UTR, two in the 3'UTR, and four in the coding sequence of which three were nonsynonymous.

This unexpected lack of more common variants led to in silico mining of multiple databases to ascertain variability of *HSF1* reported by others. Our search of the HapMap and 1000

Genome databases using the web-based browsers available for each database did not identify any variants in *HSF1*. The dbSNP database contained 94 variants within the human *HSF1* exomic sequence of which only 1 (rs78202224) had MAF >1 % and this variant was also identified in our own sequencing analysis (Table 3). Twenty nine of these variants had been validated by having multiple independent submissions, detection within two chromosomes, or identification within the 1000 Genome sequence database, or were validated by our own sequencing analysis. Our search of the NHLBI ESP database found a total of 67 variants of which 55 were present in the dbSNP database, including 2 that were also found in our own sequencing results, and 12 unique variants that were not found in any of the other analyzed databases or in our sequencing results. Of the 12 *HSF1* variants unique to the ESP database, only 2 (rs375790976 and rs369366319) had MAF >1 %. Of the 31 validated variants from dbSNP and NHLBI ESP databases and the 5 novel variants identified by our own sequencing, 2 were located in the 5'UTR, 10 were located in the 3'UTR, and 24 were located in the coding region. Of the variants in the coding region, 16 were nonsynonymous.

For those variants identified in dbSNP and validated by detecting in the 1000 Genomes database, the estimated MAF varied between 0.0005 and 0.04 (C1263A). The variants identified in the NHLBI ESP database had an estimated MAF for combined European and African Americans of 0.00008 to 0.025. Our own sequencing identified the C1263A single nucleotide polymorphism (SNP) in 3 of 84 subjects but this SNP was exclusively found in the DNA samples from the 26 African American subjects. The novel nonsynonymous C1220A variant was detected in 3 individuals and exclusively in the 24 samples from Asian subjects. The mean allelic frequency of the C2066T variant was identified with a MAF=0.0055 in the dbSNP database, but we found it in the heterozygous state by our sequencing in 2 of 30

Table 2 Results of sequencing. DNA from 80 anonymous subjects was subjected to exomic sequencing of *HSF1* and analyzed for nucleotide variations compared with a reference sequence (GenBank accession no. NT_037704.5)

Genomic address (nt) ^a	mRNA location (nt) ^b	Nucleotide substitution	Number in Caucasians ^c (MAF)	Number in African Americans	Number in Asians
82728 ^d	47	C > T	0	1 (0.0125)	0
82,783 ^d	102	G > A	0	1 (0.0125)	0
82,923	242	C > A	0	1 (0.0125)	0
103,233 ^d	1,203	G > A	0	1 (0.0125)	0
103,250 ^d	1,220	C > A	0	0	3 (0.0375)
103,293	1,263	C > A	0	3 (0.0375)	0
105,525 ^d	1,880	G > A	1 (0.0125)	0	0
105,711	2,066	C > T	2 (0.025)	0	0

^aBased on GenBank file NT_037704.5

^bBased on GenBank file NM_005526.2

^cNumber of individuals with indicated variant in 30 Caucasians, 26 African Americans, and 24 Asians

^dUnique variants not found in any databases

Table 3 Position of variants identified by database analysis and resequencing

Cluster ID number	mRNA location (nt) ²	Nucleotide substitution	AA no. ^b	AA substitution	Validated ^c	MAF ^d	Database ^e	Protein domain ^f
Sequenced ^a	47	C > T	n/a	n/a			Seq	5'UTR
Sequenced	102	G > A	n/a	n/a			Seq	5'UTR
rs368505579	229	C > T	20	T > I		0.00008	dbSNP, ESP	DNA-binding
rs189283411	242	C > A	24	Synonymous	1, 3	0.0041	dbSNP, ESP Seq	
rs150835821	293	G > A	41	Synonymous	1, 2	0.0006	dbSNP, ESP	
rs138392604	371	C > G	67	Synonymous		0.00008	dbSNP, ESP	
rs70999330	405	C > T	79	R > W		0.00008	ESP	
rs143928888	422	C > T	84	Synonymous	2	0.00038	dbSNP, ESP	
rs199606239	437	G > A	89	Synonymous	3	0.0009	dbSNP	
rs147295049	460	C > T	97	T > M		0.00015	dbSNP, ESP	DNA-binding
rs202242509	477	T > A	103	C > S	1	0.00008	dbSNP, ESP	DNA-binding
rs140943256	479	C > G	103	T > W			dbSNP	DNA-binding
rs149839114	490	G > C	107	G > A		0.00008	dbSNP, ESP	DNA-binding
rs377742831	489	G > T	107	G > C		0.00008	ESP	DNA-binding
rs144843756	550	G > A	127	S > N		0.00008	dbSNP, ESP	
rs200810784	555	G > T	129	D > T	3	0.0005	dbSNP	
rs148595793	568	G > A	133	R > H		0.00015	dbSNP, ESP	
rs11538907	581	C > T	137	Synonymous			dbSNP	
rs142508322	595	C > T	142	T > M		0.00008	dbSNP, ESP	HR-A/B
rs146430632	599	C > T	143	Synonymous	1, 2, 3	0.0041	dbSNP, ESP	
rs371996718	616	G > A	149	G > E		0.00008	ESP	HR-A/B
rs140885087	730	A > G	187	N > S		0.0005	dbSNP, ESP	HR-A/B
rs376720142	746	C > T	192	Synonymous		0.00008	ESP	
rs369105325	759	G > A	197	V > M		0.00008	ESP	HR-A/B
rs372957533	771	C > A	201	Synonymous		0.00008	ESP	
rs375270900	772	G > A	201	R > Q		0.00008	ESP	HR-A/B
rs377537624	797	C > T	209	Synonymous		0.00008	ESP	
rs201430550	813	G > A	215	D > N	3	0.0005	dbSNP	
rs143289066	826	C > T	219	A > V	1, 2	0.00038	dbSNP, ESP	
rs146718106	872	C > G	234	Synonymous		0.00008	dbSNP, ESP	Regulatory
rs200339075	876	G > A	236	G > S			dbSNP	Regulatory
rs377435416	881	G > A	237	Synonymous		0.00031	dbSNP, ESP	
rs140301460	923	C > T	251	Synonymous		0.00015	dbSNP, ESP	
rs201143946	930	G > A	254	A > T	1	0.00008	dbSNP	Regulatory
rs150366176	931	C > T	254	A > V	1, 2	0.00038	dbSNP	Regulatory
rs371469057	948	A > G	260	S > G		0.00008	ESP	Regulatory
rs138083089	957	C > T	263	P > S		0.00008	dbSNP, ESP	Regulatory
rs370907821	999	A > G	277	M > V		0.00008	ESP	Regulatory
rs149124758	1,007	C > T	279	Synonymous		0.00023	dbSNP, ESP	Regulatory
rs375790976	1,021	T > C	284	I > T		0.00023	ESP	Regulatory
rs201488133	1,038	T > C	290	S > P	1, 3	0.0023	dbSNP, ESP	Regulatory
rs371903195	1,071	C > A	301	P > T		0.00008	dbSNP, ESP	Regulatory
rs147030605	1,095	C > T	309	R > W			dbSNP	Regulatory
rs376607063	1,105	A > C	312	E > A		0.00030	dbSNP, ESP	
rs369608182	1,120	G > A	317	R > H		0.00008	dbSNP, ESP	
rs138330396	1,124	A > G	318	Synonymous			dbSNP	
rs143898779	1,130	C > T	320	Synonymous		0.00023	dbSNP, ESP	
rs375748356	1,131	G > A	321	V > M		0.00008	dbSNP	
rs201422904	1,167	T > G	333	S > A	1	0.00031	dbSNP, ESP	
rs369045018	1,181	G > A	337	Synonymous		0.00008	dbSNP, ESP	
rs373370243	1,184	T > A	338	S > R		0.00008	dbSNP, ESP	
rs200842781	1,190	T > C	340	Synonymous	3	0.0005	dbSNP	

Table 3 (continued)

Cluster ID number	mRNA location (nt) ²	Nucleotide substitution	AA no. ^b	AA substitution	Validated ^c	MAF ^d	Database ^e	Protein domain ^f
rs377529771	1,192	C > T	341	A > V		0.00008	ESP	
rs369366319	1,202	C > T	344	Synonymous		0.00023	ESP	
Sequenced	1,203	G > A	345	V > I			Seq	
rs147273395	1,210	C > T	347	A > V		0.00008	dbSNP	
rs139202082	1,214	C > A	348	Synonymous			dbSNP, ESP	
rs141516077	1,217	G > A	349	Synonymous		0.00031	dbSNP, ESP	
rs374134476	1,220	C > T	350	Synonymous		0.00008	dbSNP, ESP	
sequenced	1,220	C > A	350	D > E			Seq	
rs377258216	1,221	G > A	351	A > T		0.00008	dbSNP, ESP	
rs370819503	1,241	C > A	357	Synonymous		0.00008	dbSNP, ESP	
rs373838093	1,242	G > A	358	E > K		0.00008	dbSNP, ESP	
rs150889947	1,249	G > A	360	R > Q	1, 2, 3	0.0036	dbSNP, ESP	
rs377031622	1,253	T > A	361	Synonymous		0.00008	dbSNP, ESP	
rs139397969	1,254	C > T	362	P > S		0.00008	dbSNP, ESP	
rs370103805	1,262	C > A	364	Synonymous		0.00023	dbSNP, ESP	
rs78202224	1,263	C > A	365	P > T	1, 3	0.04	dbSNP, Seq, ESP	
rs368025244	1,265	G > A	365	Synonymous		0.00008	ESP	
rs369616383	1,320	C > T	384	L > F		0.00015	dbSNP, ESP	HR-C, TAD
rs373501871	1,331	C > T	387	Synonymous		0.00008	dbSNP, ESP	
rs149648662	1,358	T > A	396	D > E		0.00008	dbSNP, ESP	HR-C, TAD
rs372022206	1,386	G > A	406	G > S		0.00008	dbSNP, ESP	HR-C, TAD
rs199954509	1,389	T > C	407	F > L	3	0.0005	dbSNP	HR-C, TAD
rs145929063	1,428	C > T	420	P > S	2	0.00023	dbSNP, ESP	TAD
rs369997109	1,461	G > C	431	D > H		0.00008	dbSNP, ESP	TAD
rs373385331	1,480	C > T	437	A > V		0.00008	dbSNP, ESP	TAD
rs375849065	1,496	C > T	442	Synonymous		0.00015	dbSNP, ESP	
rs370804984	1,503	C > A	445	P > T		0.00008	dbSNP, ESP	TAD
rs139825694	1,524	C > A	452	P > T		0.00008	dbSNP, ESP	TAD
rs201606037	1,536	A > G	456	N > D	1, 3	0.0005	dbSNP, ESP	TAD
rs375071738	1,538	C > G	456	N > K		0.00008	ESP	TAD
rs200404974	1,546	C > T	459	P > L	1, 3	0.0005	dbSNP	TAD
rs372568305	1,571	C > T	467	Synonymous		0.00008	dbSNP, ESP	
rs377250319	1,575	A > G	469	T > A		0.00008	dbSNP, ESP	TAD
rs367584997	1,576	C > T	469	T > I		0.00008	dbSNP, ESP	TAD
rs377499162	1,593	C > T	475	Synonymous		0.00034	dbSNP, ESP	
rs376075311	1,604	C > T	478	Synonymous		0.00017	dbSNP, ESP	
rs369415484	1,637	G > A	489	Synonymous		0.00017	dbSNP, ESP	
rs373631789	1,680	C > T	504	Synonymous		0.00008	ESP	
rs375062485	1,686	C > T	506	Synonymous		0.00008	ESP	
rs34454068	1,680	G >	504	D > T		0.00009	dbSNP	TAD
rs373631789	1,682	C > T	504	Frameshift				
rs375062485	1,688	C > T	506	Synonymous			dbSNP	
rs369564684	1,777	G > A	n/a	n/a			dbSNP	3'UTR
rs373326153	1,779	G > A	n/a	n/a			dbSNP	3'UTR
rs376478959	1,780	T > C	n/a	n/a			dbSNP	3'UTR
rs202020329	1,783	G > A	n/a	n/a	3		dbSNP	3'UTR
rs201939853	1,785	T > C	n/a	n/a	1, 3	0.0018	dbSNP	3'UTR
rs191177777	1,853	G > A	n/a	n/a	3	0.0005	dbSNP	3'UTR
rs373669643	1,860	T > C	n/a	n/a			dbSNP	3'UTR
rs201104313	1,869	T > C	n/a	n/a			dbSNP	3'UTR
sequenced	1,880	G > A	n/a	n/a			Seq	3'UTR
rs11538908	1,953	G > A	n/a	n/a			dbSNP	3'UTR

Table 3 (continued)

Cluster ID number	mRNA location (nt) ²	Nucleotide substitution	AA no. ^b	AA substitution	Validated ^c	MAF ^d	Database ^e	Protein domain ^f
rs181811585	2,012	G > A	n/a	n/a	3	0.0009	dbSNP	3'UTR
rs111911280	2,066	C > T	n/a	n/a	1, 2, 3	0.0055	dbSNP, Seq	3'UTR
rs149080079	2,071	G > C	n/a	n/a	1, 3	0.0032	dbSNP	3'UTR
rs143150506	2,089	T > C	n/a	n/a	3	0.0005	dbSNP	3'UTR
rs1042568	2,096	C > A	n/a	n/a			dbSNP	3'UTR
rs186977343	2,099	G > C	n/a	n/a	3	0.0005	dbSNP	3'UTR
rs368034256	2,100	T > C	n/a	n/a			dbSNP	3'UTR
rs191891262	2,144	T > G	n/a	n/a	1, 3	0.0028	dbSNP	3'UTR

^a Unique SNPs not found in any databases^b Based on GenBank file NM_005526.2^c As per dbSNP database; 1 denotes multiple independent submissions; 2 denotes minor allele detection in at least two chromosomes; 3 denotes sequence found in 1000 Genomes Database^d Minor allele frequency from dbSNP database or NHLBI database whichever was greater^e Database in which SNP found. "Seq" indicates SNP identified by sequencing test population (Table 2)^f Protein domains from Swiss-Prot accession no. Q00613; 5' and 3'UTR from GenBank accession no. NM_005526.2

samples from Caucasian subjects (MAF=0.03 in Caucasians and 0.0125 overall).

The location of the nonsynonymous coding sequence variants in HSF1 protein was determined based on established structure-function analyses of HSF1 (Rabindran et al. 1993) (Fig. 1; Tables 3 and 4). One variant was located in the DNA-binding domain, three in the regulatory domain, two in the HR-C, and three in the TAD. We used two analysis programs, PolyPhen-2 and SIFT, to predict the biological consequences of the nonsynonymous coding sequence variants identified in

HSF1 (Table 4). The predictions of the two programs were concordant for 11 of the 16. Of the 16 nonsynonymous variants analyzed, 8 were identified as possibly damaging, probably damaging, or damaging by one or both of the prediction programs. Of these, 1 potentially damaging variation was in the DNA-binding domain, 1 in the regulatory domain, 1 in the HR-C, and 2 in the TAD (Table 4).

Analysis of 5'UTR secondary structure stability using the RNAFold program in the Vienna Suite (Gruber et al. 2008) predicted stronger 5'UTR secondary structure with the 5'UTR

Table 4 Predicted consequences of validated nonsynonymous variants

mRNA location (nt) ¹	Nucleotide substitution	AA no. ^a	AA substitution	Protein domain ^b	PolyPhen-2 ^c	SIFT ^d
477	T > A	103	C > S	DNA-binding	Possibly damaging (0.677)	Tolerated (0.45)
555	G > T	129	D > T		Probably damaging (1.0)	Damaging (0)
826	C > T	219	A > V		Possibly damaging (.947)	Damaging (0.01)
930	G > A	254	A > T	Regulatory	Benign (0.002)	Tolerated (0.21)
931	C > T	254	A > V	Regulatory	Benign (0.064)	Tolerated (0.07)
1,038	T > C	290	S > P	Regulatory	Probably damaging (0.999)	Tolerated (.13)
1,167	T > G	333	S > A		Possibly damaging (0.718)	Tolerated (0.09)
1,203	G > A	345	V > I		Benign (0.287)	Tolerated (0.12)
1,220	C > A	350	D > E		Benign (0.001)	Tolerated (0.61)
1,249	G > A	360	R > Q		Benign (0)	Tolerated (0.25)
1,263	C > A	365	P > T		Benign (0)	Tolerated (0.24)
1,320	C > T	384	L > F	HR-C, TAD	Probably damaging (1.0)	Damaging (0.01)
1,389	T > C	407	F > L	HR-C, TAD	Benign (0.107)	Tolerated (0.08)
1,428	C > T	420	P > S	TAD	Possibly damaging (0.727)	Tolerated (0.4)
1,536	A > G	456	N > D	TAD	Benign (0.003)	Tolerated (0.6)
1,546	C > T	459	P > L	TAD	Benign (0.003)	Damaging (.01)

^a Based on GenBank file NM_005526.2^b Protein domains from Swiss-Prot accession no. Q00613^c PolyPhen-2 prediction program (Adzhubei et al. 2010); predicted functional effect and (raw score) listed^d SIFT prediction program (Kumar et al. 2009); predicted functional effect and (raw score) listed

variant, C47T, compared with wild type (Fig. 2). The estimated minimum free energy of the optimal secondary structure decreased from -70.8 kcal/mol in the wild-type 5'UTR to -88.6 kcal/mol in the 5'UTR containing the C47T variant. By contrast, the estimated minimum free energy of the other 5'UTR variant, G102A, was identical to wild type.

Predicted effects of validated 3'UTR variants on miRNA target sequences were analyzed using a web-based program, MicroSNiper (Barenboim et al. 2010) (Table 5). The search was restricted to miRNA seed sequences and minimal target sequence length of seven nucleotides. Eight of the ten 3'UTR variants analyzed were predicted to modify miRNA target sequences. One 3'UTR variant disrupted one miRNA target sequence, two variants created one or more new miRNA target sequences, and five variants both disrupted and created multiple miRNA target sequences in the HSF1 3'UTR.

Since genetic variation in the 5'UTR (Ringner and Krogh 2005) and 3'UTR (Barenboim et al. 2010) could alter HSF1 protein levels, we analyzed the biological consequences of reducing HSF1 protein level on resistance to heat-induced cytotoxicity using the A549 human lung adenocarcinoma cell line. We established clones stably expressing an HSF1-targeting shRNA or a scrambled control shRNA. We demonstrated that HSF1 knockdown was stable in these clones with clones exhibiting 36 to 81 % reduction in HSF1 protein levels as determined by Western blotting (Fig. 3a). Exposing wild-type A549 cells to 45°C for 0.5 to 2 h followed by 24 h recovery at 37°C caused 24–39 % cell loss (Fig. 3b). Cells expressing the scrambled shRNA had similar HSF1 protein

levels as wild-type cells and similar heat-induced cell loss. Clones A2, B5, A3, and A6 in which HSF1 was knocked down 81, 80, 61, and 36 %, respectively, exhibited increased heat sensitivity with 44–66, 47–56, 41–63, and 31–50 % cell loss after 0.5 to 2 h heat exposure ($p < 0.04$).

To begin to determine whether any of the HSF1 3'UTR SNPs discovered could modify HSF1 translational efficiency, we compared luciferase activity in HEK 293 cells transfected with pmirGLO containing the entire 391 nt wild-type HSF1 3'UTR or the HSF1 3'UTR sequence containing one of four of the 3'UTR SNPs discovered, g1853a, g2012a, c2066t, and g2099c (Fig. 3c). Introducing the wild-type HSF1 3'UTR sequence reduced expressed luciferase activity levels by 65 %. Introducing SNPs g1853a and g2012a into the wild-type HSF1 3'UTR sequence further reduced luciferase expression by an additional 27 and 44 %. In contrast, introduction of g2099c into the HSF1 3'UTR sequence increased luciferase expression by 39 % while c2066t had no effect on luciferase expression.

Discussion

We used a search strategy to catalogue variants in the human HSF1 exomic sequence that combined searches of available databases and direct exomic sequencing of DNA obtained from healthy volunteers. All variants identified were either found in the dbSNP database or discovered during sequencing of DNA from normal volunteers. To avoid analyzing spurious

Fig. 2 Predicted effect of 5'UTR variants on secondary RNA structure of HSF1 using the RNAFold (Gruber et al. 2008) prediction program using both the minimum free energy (Zuker and Stiegler 1981) and centroids (Ding et al. 2005) methods. 5' cap sites are indicated by arrows. The free energy for each predicted structure is noted

Table 5 Predicted effects of validated 3'UTR variants on miRNA target sites

mRNA location (nt) ^a	Nucleotide substitution	miRNAs disrupted ^b	miRNAs created
1,783	G > A	None	None
1,785	T > C	1,207	1,228-5p
1,853	G > A	1,825	4,301; 6,080
1,880	G > A	676-3p; 2,355-3p	550-3p; 200c-5p
2,012	G > A	1,205	None
2,066	C > T	4,781-5p; 4,749-5p; 4,706	6,510-5p, 6,081
2,071	G > C	None	None
2,089	T > C	None	3,924; 297
2,099	G > C	892a; 6,509-3p	199a-5p; 4,302; 199b-5p
2,144	T > G	None	6,720-3p

^a Based on GenBank file NM_005526.2

^b Prediction using MicroSNiPer analysis program (dbSNP build 137) (Barenboim et al. 2010) using human seeds set and minimal 7-bp binding

data from the databases, we focused on variants validated in the dbSNP database by multiple independent observations and those identified in our own sequencing analysis. To avoid false discovery in our sequencing analysis, we used a high-fidelity polymerase and bidirectional sequencing. We only

included variants identified in bidirectional sequencing and confirmed by repeat amplification and sequencing.

Of the variants identified in the dbSNP database, C1263A had the highest reported minor allele frequency, 4 %. Our own sequencing analysis identified C1263A exclusively in 3 of 26

Fig. 3 Effect of HSF1 knockdown on thermal sensitivity of A549 cells. Cells were transfected with HSF1-targeting or scrambled (scr) shRNA-expressing plasmids and cloned by serial dilution. **a** Western blot of HSF1. A representative of three blots is shown and the mean percentage knockdown of three blots is indicated below. **b** Wild-type (WT) A549 and each clone were exposed to 45 °C for 30 to 120 min, then recovered at 37 °C for 24 h. The monolayers were then washed, and the residual cells were stained with crystal violet, solubilized in SDS, and absorption at

570 nm was measured and expressed as percentage of unheated control for each clone. **c** HEK-293 cells were transfected with control pmirGLO without insert (CTRL) or with pmirGLO containing the 391 nt of wild-type HSF1 3'UTR insert (WT) or the HSF1 3'UTR with the indicated SNP and 36 h later, renilla and luciferase activities were measured, the ratio calculated, and the results expressed as a percent of the wild-type HSF1 values. Mean±SE; * $p < 0.05$

samples from African American subjects, suggesting that the minor allelic frequency of this variant might be somewhat higher in African Americans than other populations. Similarly, we only found the SNP C1220A, in Asian subjects, and its presence in 3 of 48 chromosomes in our Asian subjects suggests a minor allelic frequency of approximately 6 %. Although most of the other variants identified in dbSNP and by our own sequencing have much lower minor allele frequencies, it is still possible that their presence alone or in combination may modify susceptibility to certain diseases (Fijal et al. 2001).

Protein prediction programs suggest that some of the nonsynonymous *HSF1* variants may alter DNA binding, oligomerization, or transactivating functions of HSF1. However, these programs are only predictive, and definitive determination of functional consequences will require empiric studies. Since 5'UTR SNPs can alter translational efficiency by modifying secondary structure (Ringner and Krogh 2005), we used two methods to analyze how introducing each of the 5' UTR variants modifies secondary structure. The minimum free energy method utilizes an algorithm developed by Zuker and Stiegler that produces a molecular structure with the lowest predicted folding energy (Zuker and Stiegler 1981). The centroid method produces a structure that is the most representative of the set of predicted structures (Ding et al. 2005). Similarity in secondary structure predicted by the two algorithms increases the confidence in the prediction accuracy (Gruber et al. 2008). The centroid and minimum free energy (MFE) methods generated very similar 5'UTR structures for the wild-type HSF1 and 5'UTR variants. The calculated free energy of the MFE structures was 18 kcal/mol lower for the C47T variant than either the wild-type HSF1 or G102A variant, indicating a much more stable 5'UTR and a lower translation rate for the C47T variant. Ringner and Krogh demonstrated a genome-wide inverse correlation between stability of 5'UTR folding and protein expression level in yeast (Ringner and Krogh 2005).

We utilized the MicroSNiPer prediction program (Barenboim et al. 2010) to predict how each of the 3'UTR variants might alter miRNA targets. We elected a conservative approach by limiting the predicted RNA:RNA interactions to known miRNA seed sequences with binding over at least 7 sequential nucleotides. That 8 of the 10 identified 3'UTR variants were predicted to both create and disrupt miRNA targets suggests that the variants could potentially increase or decrease levels of HSF1 protein depending on the relative availability of miRNAs. We restricted our analysis of miRNA targets to the 3'UTR. While miRNAs can bind to the coding region of genes, the impact of such interactions on translational efficiency is usually much less than the interaction of miRNAs with target sequences in the 3'UTR (Fang and Rajewsky 2011). Future empiric studies are planned to analyze the impact of the 3'UTR variants on HSF1 protein level.

Complete loss of HSF1 has been shown to impact survival of cells (Page et al. 2006) and entire organisms (Sagi and Kim 2012). To begin to understand whether partial reductions in HSF1 protein levels that may be caused by variants in the 5' and 3'UTRs could have biological consequences, we analyzed heat resistance in A549 cells with varying levels of HSF1 knockdown. We used a heat exposure previously shown to cause cell death in A549 cells (Armour et al. 1993) and a simple cytotoxicity assay developed for adherent cells (Ruff and Gifford 1981). We found that knocking down HSF1 levels between 36 and 61 % increased susceptibility to heat-induced cell death. The mechanisms of cell death and HSF1-dependent protection have not been addressed in this study, but the results of this experiments suggest that variants in the 5' and 3'UTR that affect HSF1 levels may have consequences for cell survival, especially in the face of stress. Using an expression system developed to evaluate the translational effects of miRNA, we found that inserting the entire wild-type HSF1 3'UTR sequence repressed translation by 65 %. Two of the four 3'UTR SNPs tested reduced translational efficiency between 27 and 44 % more than wild type. One of these SNPs, g1853a, was predicted using a computer-based analysis to inactivate on miRNA target and create two new targets. The other active SNP, g2012a, was predicted to inactivate a miRNA target but not create a new target. Another SNP, g2099c, repressed the expression plasmid less than wild-type and was predicted to disrupt two and create three potential miRNA targets. These findings demonstrate the complexity in predicting the effects of 3' UTR mutations on gene translation but support a potentially important influence of SNPs in the human HSF1 3'UTR sequence in determining basal levels of HSF1 protein and the capacity of cells to withstand stress.

In summary, we have described 34 variations in the human exomic sequence of *HSF1*, several of which have the potential to alter the function or level of HSF1 protein. We showed that partial reduction in HSF1 protein levels as might occur with some of these variants is sufficient to increase sensitivity to heat-induced cytotoxicity. These studies underscore the importance of empiric studies to more completely define the biological consequences of HSF1 genetic variation and the association of these SNPs with human diseases.

Acknowledgments The authors would like to thank the NHLBI GO Exome Sequencing Project and its ongoing studies which produced and provided exome variant calls for comparison: the Lung GO Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926), and the Heart GO Sequencing Project (HL-103010) and Dr. Braxton Mitchell for his help in preparing this manuscript.

The work was supported by the National Institutes of Health grants HL69057 and HL085256, and the Veterans Administration Merit Review grant BX002143 with additional partial support provided by the Mid-Atlantic Nutrition and Obesity Center (P30 DK072488)

References

- (2003) The International HapMap Project. *Nature* 426: 789–796
- Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249
- Armour EP, McEachern D, Wang Z, Corry PM, Martinez A (1993) Sensitivity of human cells to mild hyperthermia. *Cancer Res* 53: 2740–2744
- Barenboim M, Zoltick BJ, Guo Y, Weinberger DR (2010) MicroSniPer: a web tool for prediction of SNP effects on putative microRNA targets. *Hum Mutat* 31:1223–1232
- Charette SJ, Lavoie JN, Lambert H, Landry J (2000) Inhibition of Daxx-mediated apoptosis by heat shock protein 27. *Mol Cell Biol* 20: 7602–7612
- Dai C, Whitesell L, Rogers AB, Lindquist S (2007) Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. *Cell* 130:1005–1018
- Ding Y, Chan CY, Lawrence CE (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. *RNA* 11: 1157–1166
- Exome Variant Server NGSPE. <http://evs.gs.washington.edu/EVS>
- Fang Z, Rajewsky N (2011) The impact of miRNA target sites in coding sequences and in 3'UTRs. *PLoS One* 6:e18067
- Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu Rev Physiol* 61:243–282
- Fijal BA, Kim LL, Buxbaum SG, Witte JS (2001) Predicting quantitative trait levels by modeling SNP interaction. *Genet Epidemiol* 21(Suppl 1):S608–S613
- Fujimoto M, Hayashida N, Katoh T, Oshima K, Shinkawa T, Prakasam R, Tan K, Inouye S, Takii R, Nakai A (2010) A novel mouse HSF3 has the potential to activate nonclassical heat-shock genes during heat shock. *Mol Biol Cell* 21:106–116
- Goldring CE, Reveneau S, Chantome A, Pance A, Fleury C, Hume DA, Sester D, Mignotte B, Jeannin JF (2000) Heat shock enhances transcriptional activation of the murine-inducible nitric oxide synthase gene. *FASEB J* 14:2393–2395
- Green M, Schuetz TJ, Sullivan EK, Kingston RE (1995) A heat shock-responsive domain of human HSF1 that regulates transcription activation domain function. *Mol Cell Biol* 15:3354–3362
- Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL (2008) The Vienna RNA website. *Nucleic Acids Res* 36:W70–W74
- Huang AJ, Yu KD, Li J, Fan L, Shao ZM (2012) Polymorphism rs4919510: C > G in mature sequence of human microRNA-608 contributes to the risk of HER2-positive breast cancer but not other subtypes. *PLoS One* 7:e35252
- Inouye S, Fujimoto M, Nakamura T, Takaki E, Hayashida N, Hai T, Nakai A (2007) Heat shock transcription factor 1 opens chromatin structure of interleukin-6 promoter to facilitate binding of an activator or a repressor. *J Biol Chem* 282:33210–33217
- Izu H, Inouye S, Fujimoto M, Shiraishi K, Naito K, Nakai A (2004) Heat shock transcription factor 1 is involved in quality-control mechanisms in male germ cells. *Biol Reprod* 70:18–24
- Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4:1073–1081
- Lasky-Su J, Anney RJ, Neale BM, Franke B, Zhou K, Maller JB, Vasquez AA, Chen W, Asherson P, Buitelaar J, Banaschewski T, Ebstein R, Gill M, Miranda A, Mulas F, Oades RD, Roeyers H, Rothenberger A, Sergeant J, Sonuga-Barke E, Steinhausen HC, Taylor E, Daly M, Laird N, Lange C, Faraone SV (2008) Genome-wide association scan of the time to onset of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 147B:1355–1358
- Lee YJ, Lee HJ, Lee JS, Jeoung D, Kang CM, Bae S, Lee SJ, Kwon SH, Kang D, Lee YS (2008) A novel function for HSF1-induced mitotic exit failure and genomic instability through direct interaction between HSF1 and Cdc20. *Oncogene* 27: 2999–3009
- Li QL, Ju ZH, Huang JM, Li JB, Li RL, Hou MH, Wang CF, Zhong JF (2011) Two novel SNPs in HSF1 gene are associated with thermal tolerance traits in Chinese Holstein cattle. *DNA Cell Biol* 30:247–254
- Liggett SB (2003) Polymorphisms of adrenergic receptors: variations on a theme. *Assay Drug Dev Technol* 1:317–326
- McMillan DR, Xiao X, Shao L, Graves K, Benjamin IJ (1998) Targeted disruption of heat shock transcription factor 1 abolishes thermotolerance and protection against heat-inducible apoptosis. *J Biol Chem* 273:7523–7528
- Mendillo ML, Santagata S, Koeva M, Bell GW, Hu R, Tamimi RM, Fraenkel E, Ince TA, Whitesell L, Lindquist S (2012) HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. *Cell* 150:549–562
- Murray JI, Whitfield ML, Trinklein ND, Myers RM, Brown PO, Botstein D (2004) Diverse and specific gene expression responses to stresses in cultured human cells. *Mol Biol Cell* 15:2361–2374
- Nakai A, Tanabe M, Kawazoe Y, Inazawa J, Morimoto RI, Nagata K (1997) HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. *Mol Cell Biol* 17:468–481
- Nusbaum C, Mikkelsen TS, Zody MC, Asakawa S, Taudien S, Garber M, Kodira CD, Schueler MG, Shimizu A, Whittaker CA, Chang JL, Cuomo CA, Dewar K, FitzGerald MG, Yang X, Allen NR, Anderson S, Asakawa T, Blechschmidt K, Bloom T, Borowsky ML, Butler J, Cook A, Corum B, DeArellano K, DeCaprio D, Dooley KT, Dorris L 3rd, Engels R, Glockner G, Hafez N, Hagopian DS, Hall JL, Ishikawa SK, Jaffe DB, Kamat A, Kudoh J, Lehmann R, Lokitsang T, Macdonald P, Major JE, Matthews CD, Mauceli E, Menzel U, Mihalev AH, Minoshima S, Murayama Y, Naylor JW, Nicol R, Nguyen C, O'Leary SB, O'Neill K, Parker SC, Polley A, Raymond CK, Reichwald K, Rodriguez J, Sasaki T, Schilhabel M, Siddiqui R, Smith CL, Sneddon TP, Talamas JA, Tenzin P, Topham K, Venkataraman V, Wen G, Yamazaki S, Young SK, Zeng Q, Zimmer AR, Rosenthal A, Birren BW, Platzer M, Shimizu N, Lander ES (2006) DNA sequence and analysis of human chromosome 8. *Nature* 439:331–335
- Page TJ, Sikder D, Yang L, Pluta L, Wolfinger RD, Kodadek T, Thomas RS (2006) Genome-wide analysis of human HSF1 signaling reveals a transcriptional program linked to cellular adaptation and survival. *Mol Biol* 2:627–639
- Perisic O, Xiao H, Lis JT (1989) Stable binding of drosophila heat shock factor to head-to-head and tail-tail repeats of a conserved 5 bp recognition unit. *Cell* 59:797–806
- Peteranderl R, Nelson HC (1992) Trimerization of the heat shock transcription factor by a triple-stranded alpha-helical coiled-coil. *Biochemistry* 31:12272–12276
- Pirkkala L, Alastalo TP, Zuo X, Benjamin IJ, Sistonen L (2000) Disruption of heat shock factor 1 reveals an essential role in the ubiquitin proteolytic pathway. *Mol Cell Biol* 20:2670–2675
- Rabindran SK, Giorgi G, Clos J, Wu C (1991) Molecular cloning and expression of a human heat shock factor, HSF1. *Proc Natl Acad Sci U S A* 88:6906–6910
- Rabindran SK, Haroun RI, Clos JW, Sisniewski J, Wu C (1993) Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. *Sci* 259:230–234
- Ringner M, Krogh M (2005) Folding free energies of 5'-UTRs impact post-transcriptional regulation on a genomic scale in yeast. *PLoS Comput Biol* 1:e72

- Rossi A, Coccia M, Trotta E, Angelini M, Santoro MG (2012) Regulation of cyclooxygenase-2 expression by heat: a novel aspect of heat shock factor 1 function in human cells. *PLoS One* 7:e31304
- Rossi A, Trotta E, Brandi R, Arisi I, Coccia M, Santoro MG (2010) AIRAP, a new human heat shock gene regulated by heat shock factor 1. *J Biol Chem* 285:13607–13615
- Ruff MR, Gifford GE (1981) Rabbit tumor necrosis factor: mechanism of action. *Infect Immun* 31:380–385
- Sagi D, Kim SK (2012) An engineering approach to extending lifespan in *C. elegans*. *PLoS Genet* 8:e1002780
- Santos SD, Saraiva MJ (2004) Enlarged ventricles, astrogliosis and neurodegeneration in heat shock factor 1 null mouse brain. *Neuroscience* 126:657–663
- Sarge KD, Murphy SP, Morimoto RI (1993) Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization, acquisition of DNA-binding activity, and nuclear localization and can occur in the absence of stress. *Mol Cell Biol* 13:1392–1407
- Schuetz TJ, Gallo GJ, Sheldon L, Tempst P, Kingston RE (1991) Isolation of a cDNA for HSF2: evidence for two heat shock factor genes in humans. *Proc Natl Acad Sci U S A* 88:6911–6915
- 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
- Singh IS, Calderwood S, Kalvakolanu I, Viscardi R, Hasday JD (2000) Inhibition of tumor necrosis factor- α transcription in macrophages exposed to febrile range temperature: a possible role for heat shock factor-1. *J Biol Chem* 275:9841–9848
- Singh IS, Gupta A, Nagarsekar A, Cooper Z, Manka C, Hester L, Benjamin IJ, He J-E, Hasday JD (2008) Heat shock co-activates interleukin-8 transcription. *Am J Respir Cell Molec Biol* 39:235–242
- Singh IS, Shah NG, Almutairy EA, Hasday JD (2009a) Role of HSF1 in infectious disease. In: Pockley AG, Santoro MG, Calderwood SK (eds) *Prokaryotic and eukaryotic heat shock proteins in infectious disease*. Springer, Dordrecht Heidelberg London, pp 1–32
- Singh U, Bongcam-Rudloff E, Westermarck B (2009b) A DNA sequence directed mutual transcription regulation of HSF1 and NFIX involves novel heat sensitive protein interactions. *PLoS One* 4:e5050
- Singhal SS, Yadav S, Drake K, Singhal J, Awasthi S (2008) Hsf-1 and POB1 induce drug sensitivity and apoptosis by inhibiting Ralbp1. *J Biol Chem* 283:19714–19729
- Skaggs HS, Xing H, Wilkerson DC, Murphy LA, Hong Y, Mayhew CN, Sarge KD (2007) HSF1-TPR interaction facilitates export of stress-induced HSP70 mRNA. *J Biol Chem* 282:33902–33907
- Small KM, Brown KM, Seman CA, Theiss CT, Liggett SB (2006) Complex haplotypes derived from noncoding polymorphisms of the intronless α 2A-adrenergic gene diversify receptor expression. *Proc Natl Acad Sci U S A* 103:5472–5477
- Takaki E, Fujimoto M, Nakahari T, Yonemura S, Miyata Y, Hayashida N, Yamamoto K, Vallee RB, Mikuriya T, Sugahara K, Yamashita H, Inouye S, Nakai A (2007) Heat shock transcription factor 1 is required for maintenance of ciliary beating in mice. *J Biol Chem* 282:37285–37292
- Takaki E, Fujimoto M, Sugahara K, Nakahari T, Yonemura S, Tanaka Y, Hayashida N, Inouye S, Takemoto T, Yamashita H, Nakai A (2006) Maintenance of olfactory neurogenesis requires HSF1, a major heat shock transcription factor in mice. *J Biol Chem* 281:4931–4937
- Trinklein ND, Chen WC, Kingston RE, Myers RM (2004) Transcriptional regulation and binding of heat shock factor 1 and heat shock factor 2 to 32 human heat shock genes during thermal stress and differentiation. *Cell Stress Chaperones* 9:21–28
- Wang J, He H, Yu L, Xia HH, Lin MC, Gu Q, Li M, Zou B, An X, Jiang B, Kung HF, Wong BC (2006) HSF1 down-regulates XAF1 through transcriptional regulation. *J Biol Chem* 281:2451–2459
- Wang KS, Liu XF, Aragam N (2010) A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res* 124:192–199
- Westwood JT, Clos J, Wu C (1991) Stress-induced oligomerization and chromosomal relocation of heat-shock factor. *Nature* 353:822–827
- Xiao X, Zuo X, Davis AA, McMillan DR, Curry BB, Richardson JA, Benjamin IJ (1999) HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. *Embo J* 18:5943–5952
- Xie Y, Chen C, Stevenson MA, Hume DA, Auron PE, Calderwood SK (2002) NF-IL6 and HSF1 have mutually antagonistic effects on transcription in monocytic cells. *Biochem Biophys Res Commun* 291:1071–1080
- Yuan CX, Gurley WB (2000) Potential targets for HSF1 within the preinitiation complex. *Cell Stress Chaperones* 5:229–242
- Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res* 9:133–148
- Zuo J, Baler R, Dahl G, Voellmy R (1994) Activation of the DNA-binding ability of human heat shock transcription factor 1 may involve the transition from an intramolecular to an intermolecular triple-stranded coiled-coil structure. *Mol Cell Biol* 14:7557–7568