HSF1 Drives a Transcriptional Program Distinct from Heat Shock to Support Highly Malignant Human Cancers

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1016/j.cell.2012.06.031">http://dx.doi.org/10.1016/j.cell.2012.06.031</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Elsevier</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Apr 24 18:00:11 EDT 2016</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/91912">http://hdl.handle.net/1721.1/91912</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
HSF1 Drives a Transcriptional Program Distinct from Heat Shock to Support Highly Malignant Human Cancers

Marc L. Mendillo,1,8 Sandro Santagata,1,2,8 Martina Koeva,1,3 George W. Bell,1 Rong Hu,5,6 Rulla M. Tamimi,5,6 Ernest Fraenkel,2 Tan A. Ince,7 Luke Whitesell,1,* and Susan Lindquist1,4,*
1The Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
2Department of Pathology, Brigham and Women’s Hospital, Boston, MA 02115, USA
3Department of Biological Engineering
4Howard Hughes Medical Institute, Department of Biology
Massachusetts Institute of Technology, Cambridge, MA 02139, USA
5Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
6Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA 02115, USA
7Department of Pathology, Braman Family Breast Cancer Institute and Interdisciplinary Stem Cell Institute, University of Miami Miller School of Medicine, Miami, FL 33136, USA
8These authors contributed equally to this work
*Correspondence: whitesell@wi.mit.edu (L.W.), lindquist_admin@wi.mit.edu (S.L.)
http://dx.doi.org/10.1016/j.cell.2012.06.031

SUMMARY
Heat-Shock Factor 1 (HSF1), master regulator of the heat-shock response, facilitates malignant transformation, cancer cell survival, and proliferation in model systems. The common assumption is that these effects are mediated through regulation of heat-shock protein (HSP) expression. However, the transcriptional network that HSF1 coordinates directly in malignancy and its relationship to the heat-shock response have never been defined. By comparing cells with high and low malignant potential alongside their nontransformed counterparts, we identify an HSF1-regulated transcriptional program specific to highly malignant cells and distinct from heat shock. Cancer-specific genes in this program support oncogenic processes: cell-cycle regulation, signaling, metabolism, adhesion and translation. HSP genes are integral to this program, however, many are uniquely regulated in malignancy. This HSF1 cancer program is active in breast, colon and lung tumors isolated directly from human patients and is strongly associated with metastasis and death. Thus, HSF1 rewires the transcriptome in tumorigenesis, with prognostic and therapeutic implications.

INTRODUCTION
A wide variety of environmental stressors can damage proteins. These include elevated temperatures, oxidative agents, heavy metals, and low pH. Organisms respond by inducing heat-shock proteins (HSPs), which act as molecular chaperones to restore protein homeostasis (Shamovsky and Nuclider, 2008; Whitesell and Lindquist, 2005). This powerful adaptive mechanism, known as the heat-shock response, is unleashed by the heat-shock transcription factor HSF1. Upon heat shock, HSF1 is phosphorylated, trimerizes, and translocates to the nucleus. There, it induces chaperone gene expression by binding to DNA sequence motifs known as heat-shock elements (HSEs) (Pelham, 1982; Sakurai and Enoki, 2010). Major aspects of this classic response are conserved from yeast to humans and are vital in many stressful environments. HSF1 also functions as a critical regulator of longevity in some organisms (Chiang et al., 2012; Volovik et al., 2012). Consistent with this, recent work indicates that HSF1 helps cells accommodate the complex pathophysiological derangements in protein homeostasis that underlie many human diseases, especially those associated with aging (Morimoto, 2008).

We have previously shown in mice that HSF1 is co-opted by tumor cells to promote their survival, to the detriment of their hosts. The importance of HSF1 in supporting carcinogenesis, at least in model systems, is demonstrated by the dramatically reduced susceptibility of Hsf1-knockout mice to tumor formation. This has been established for cancers driven by oncogenic RAS, tumor suppressor p53 mutations, and chemical carcinogens (Dai et al., 2007; Jin et al., 2011; Min et al., 2007). In addition to its role in tumor formation in mice, HSF1 fosters the growth of human tumor cells in culture. Depleting HSF1 from established human cancer lines markedly reduces their proliferation and survival (Dai et al., 2007; Meng et al., 2010; Min et al., 2007; Santagata et al., 2012; Zhao et al., 2011).

In mouse models, HSF1 enables adaptive changes in a diverse array of cellular processes, including signal transduction, glucose metabolism, and protein translation (Dai et al., 2007; Khaleque et al., 2008; Lee et al., 2008; Zhao et al., 2009, 2011). The commonly held view is that HSF1 exerts this broad influence in cancer simply by allowing cells to manage the
RESULTS

HSF1 Is Activated in Highly Tumorigenic Cells

We first asked whether HSF1 expression differed in the highly malignant BPLER and the much less malignant HMLER breast cancer cells (Ince et al., 2007). We used two sets of such cells, each pair derived independently from a different donor. In both, HSF1 protein expression was higher in the more malignant member of the pair, the BPLER cells (Figure 1A). The BPLER cells
also had more phosphoserine-326-HSF1, a well established marker of HSF1 activation (Guettouche et al., 2005), than the HMLER cells (Figure 1A).

To determine whether these differences in HSF1 were simply an artifact of growth in cell culture, we implanted the cells into immunocompromised mice and allowed them to form tumors. HSF1 immunostaining was weak in the HMLER tumors. Moreover, it was largely restricted to nonmalignant, infiltrating stroma and to tumor areas bordering necrosis (Figure 1B), indicating that microenvironmental stress can influence the activation of HSF1. In BPLER tumors, however, HSF1 staining was strong, nuclear localized and very uniform thoughout (Figure 1B; Figure S1A available online). Thus, the dramatic difference in HSF1 expression we observe between BPLER and HMLER cells is due to stable, cell-autonomous factors intrinsic to these distinct cell types (Ince et al., 2007).

Given this evidence for the activation of HSF1 in the BPLER cell type, we asked whether such cells were more dependent on HSF1 than HMLER cells for growth and survival. Neither cell type was affected by negative control shRNA. With two independent shRNAs that knockdown HSF1 expression, however, cell growth and viability were far more strongly reduced in the BPLER than the HMLER cells (Figure S1B).

**HSF1 Genome Occupancy in Cancer Is Distinct from Heat Shock**

To determine whether the transcriptional program driven by HSF1 in highly malignant cells differs from that driven by a classical thermal stress, we used chromatin immunoprecipitation coupled with massively parallel DNA sequencing (ChIP-Seq), characterizing HSF1-binding sites genome wide. We first assessed the immortalized nontransformed progenitor cells, HME and BPE, grown at 37°C or following a 42°C heat shock (Figure 1C). We then related the genome-wide distribution of HSF1-binding sites to those of the oncogenically transformed HMLER and BPLER cells grown at 37°C.

In the HME and BPE parental cell lines, a limited number of genes were bound by HSF1 in the absence of heat shock, and these were bound weakly (Figure 1D; Table S1). Heat shock drove robust binding of HSF1 to ~800 genes in HME cells and to ~1,100 genes in BPE cells (Figure 1D; Table S1). These observations are consistent with a previous report that a large number of genes are bound by HSF1 in the mammalian heat-shock program, we asked whether the genes comprising this program differ. For example, the strongly heat-shock inducible HSPA6 gene (encoding HSP70B) was highly bound in parental lines upon heat shock but only weakly bound in BPLER cells at 37°C (Figures 1F, S1G and S1H). Conversely, PROM2, which encodes a basal epithelial cell membrane glycoprotein, was weakly bound by HSF1 in parental lines following heat shock, but highly bound in BPLER cells (Figure 1F). Thus, HSF1 engages a regulatory program in the highly malignant state that is distinct from the classic heat-shock response.

To assess the functional significance of the HSF1 cancer program, we asked whether the genes comprising this program played a significant role in malignancy and used unbiased data from an independent investigation. The Elledge lab recently conducted a whole-genome siRNA screen to identify genes that are required to maintain growth when cells are transformed with a malignantly activated Ras gene (Luo et al., 2009). Among the ~1,600 genes identified in this screen, our HSF1-bound gene set was very strongly enriched (73 gene overlap; p value = 7.95 × 10^{-15}, Table S2). The HSF1-bound genes we identified as unique to the malignant state were more strongly enriched (49 gene overlap; p value = 1.1 × 10^{-12}) than those shared with heat-shocked cells (24 gene overlap; p value = .0004), but both sets of genes were important in supporting the malignant state.

**HSF1 Regulates the Transcription of Genes It Binds in Malignant Cells**

To investigate the consequences of HSF1 occupancy on gene expression, we compared RNA profiles in HMLER and BPLER cells transduced with control shRNA hairpins to those transduced with hairpins that knockdown HSF1. As we previously
reported, the growth and survival of malignant cells is compromised by prolonged depletion of HSF1 (Dai et al., 2007). Therefore, we only analyzed mRNA expression in the early stages of shRNA inhibition, where HSF1 knockdown was still incomplete (Figure S2), but cell viability was unimpaired. This provides a conservative assessment of the effects of HSF1 on gene expression in malignant cells.

Control hairpins that did not reduce HSF1 levels (Scr and GFP; Figure S2) had minimal effects on the expression of HSF1-bound genes (Figure 2A; Table S3). Targeted hairpins that did reduce HSF1 had a minor impact in HMLER cells but markedly changed expression in BPLER cells. The expression of many genes decreased indicating that they were negatively regulated by HSF1. Genes unique to the malignant state and genes shared by heat-shocked cells were affected equivalently. For example, expression of the malignancy-associated genes CKS2 and RBM23 and the heat-shock protein genes HSPA8 (HSC70) and HSP90AA1 (HSP90) were all reduced (by ~50%) following HSF1 knockdown (Table S3).

Relating the effects of HSF1 knockdown on gene expression to our earlier ChIP-Seq analysis, ~70% of genes positively regulated by HSF1 were bound at the promoter, whereas only ~30% of these genes were bound in distal regions (Figure 2B). Genes that were negatively regulated by HSF1 showed the opposite pattern (Figure 2B). This observation (p value = 0.00004) suggests that the direction of regulation (positive versus negative) in these cells is influenced by the location of the HSF1-binding site.

Next we examined the effects of HSF1 knockdown on gene expression in a cell line that had not been deliberately engineered. The MCF7 line was established from a human breast cancer metastasis. Moreover, as an estrogen receptor positive (ER+) line, its biology is fundamentally distinct from the hormone-receptor-negative HMLER and BPLER cell lines. Despite these differences, the changes in gene expression caused by HSF1 knockdown was very similar in BPLER cells and MCF7 cells (Figure 2A).

**HSF1 Gene Occupancy Is Conserved across a Broad Range of Common Human Cancer Cell Lines**

Next we used ChIP-qPCR to monitor HSF1 binding to a representative set of the HSF1-target genes in cell lines derived from patients with breast cancer. We used nine well-studied cancer lines (including MCF7 cells) representing all three major categories of breast cancer: ER+, HER2+ and Triple-Negative (TN). Under basal conditions (at 37°C), we detected HSF1 binding in each of the major breast cancer subtypes (Figure S3A). A range of binding intensities was observed. Most notably, however, the distinct pattern of HSF1 gene occupancy we had identified in the highly malignant BPLER cells was also present in these naturally-arising malignant cells. This included genes unique to malignant cells, such as CKS2 and RBM23, and genes shared by heat-shocked cells, such as HSPA1/E1. Again, the gene most strongly inducible by heat shock, HSPA6, was minimally bound across this entire panel of cancer lines under basal conditions (37°C; Figures S3A, S3B and S3C). We also analyzed HSF1 binding in the nonmalignant breast cell line MCF10A. Comparable to the low malignancy HMLER cells, MCF10A cells had low levels of HSF1 occupancy across all genes examined (Figures S3A and S3C).

These ChIP-qPCR data spurred us to employ ChIP-Seq to generate additional genome-wide high-resolution maps of HSF1 occupancy. We performed ChIP-Seq analysis on the non-tumorigenic MCF10A cell line grown either at 37°C or following a 42°C heat-shock. We compared these data with our prior data from the nontumorigenic cell lines HME and BPE and weakly tumorigenic HMLER cells. We then assessed HSF1 binding in a panel of human tumor lines that extended to other types of malignancy: duplicate samples of four breast, three lung and three colon cancer cell lines (Figures 3A and S3D), thus covering the human cancers with the highest total mortality in the developed world.

After heat shock, MCF10A cells exhibited an HSF1-binding profile that was comparable to that of heat-shocked HME and BPE cells. In the absence of heat shock, the overall magnitude
Figure 3. Genome-wide patterns of DNA occupancy by HSF1 across a broad range of common human cancer cell lines

(A) Heat map of ChIP-Seq read density for all HSF1 target regions (union of all HSF1-bound regions in all data sets). Genomic regions from −1kb to +1kb relative to the peak of HSF1 binding are shown. Regions are ordered the same in all data sets. Read density is depicted for nontumorigenic cells at 37°C (green), cancer lines at 37°C (black) and nontumorigenic (nt) lines following heat shock at 42°C (red). Asterisks indicate data sets also used for the analysis in Figure 1E.

(B) Principal component analysis (PCA) of HSF1 binding in heat-shocked parental lines (red) and cancer lines (black). Tumorigenic lines (HME, BPE, MCF10A) and both cancer and heat-shocked nontumorigenic lines. See also Figure S3 and Table S1.

(C) ChIP-Seq density heat map of genomic regions differentially bound by HSF1 in cancer lines at 37°C/C14.

(D) Motif analysis of 100 bp regions surrounding HSF1-binding peaks for genes enriched in cancer lines (BT20, NCIH838 and SKBR3), heat-shocked nontumorigenic lines (HME, BPE, MCF10A) and both cancer and heat-shocked nontumorigenic lines.
of HSF1 binding in all of the nontumorigenic cell lines (nt) was uniformly very weak, and the total number of bound genes was small (Figure 3A; Table S1). In contrast, in the cancer lines a range of HSF1 binding was observed at 37°C (Figure 3A). For example, robust binding was observed in the lung adenocarcinoma line NCI-H838 and in the TN breast carcinoma line BT20. Less pronounced overall binding was seen in others lines such as the weakly malignant HMLER. Binding in BPLER cells was intermediate.

Irrespective of the level of binding, the distribution of HSF1 occupancy on a genome-wide scale was remarkably similar among the cancer cell lines and distinct from the pattern of binding in the heat-shocked cells (Figure 3A). The differences between the heat-shocked and malignant states were further probed using principal component analysis (PCA; Figure 3B). This unsupervised method of clustering sets of data clearly distinguished one cluster containing all cell lines exposed to heat-shock and a second cluster containing all cancer cell lines. This analysis confirmed the global nature of the differences in the HSF1-binding profiles.

Data from these multiple cell lines allowed us to confidently identify regions of HSF1 binding that were strong in cancer cells but not in heat-shocked cells, weak in cancer but strong in heat shock, or similarly strong in both (Figure 3C). Examples of genes that were strongly bound in cancer but not in heat shock included CKS2, LY6K, RBM23, CCT6A, CKS1B, ST13, EIF4A2 (Figures S3E; Figure 3D). Genes that were weakly bound in cancer lines but strongly bound in heat shock included HSPA6 and DNAJC7 (Figure 3D). Genes that were strongly bound in both cell types included HSPA4L and HSP90A1B1 (Figure 3D).

We performed motif analysis to evaluate the 100 bp genomic regions surrounding the peaks of HSF1 binding in each of these groups. The HSE, comprised of adjacent inverted repeats of 5'-nGAA-n3', was the most enriched motif in all three groups (Figure 3E). The regions strongly bound in cancer but not in heat shock were enriched in HSEs that had three such repeats (p value = 8.8 × 10^{-106}). They were also enriched in binding elements for YY1, the so-called “ying-yang” transcription factor that is involved in activating and repressing a broad range of genes (p value = 3.7 × 10^{-7}). The regions strongly bound in heat-shocked cells but not cancer were enriched for expanded HSEs with a fourth 5'-nGAA-n3' repeat (p value = 4.6 × 10^{-128}). They were also enriched in an AP1/Fos/NRF2 (NF2EL2) binding site (p value = 1.4 × 10^{-24}) as previously reported for mammalian heat-shock genes. This variation in binding motifs suggests the involvement of distinct coregulators in establishing differential patterns of HSF1 occupancy. The regions strongly bound by HSF1 in both cancer and in heat shock, had features of both groups. They were enriched for HSEs with three inverted repeats (p value = 1.3 × 10^{-125}). They were not enriched for the YY1 sites but were enriched for the AP1/Fos and NRF2 binding site (p value = 5.2 × 10^{-7}).

HSF1-Bound Genes form Distinct, Coordinately Regulated Modules

Integrating our diverse data sets revealed a direct and pervasive role for HSF1 in cancer biology (Figure 4A). Extending far beyond protein folding and stress, HSF1-bound genes were involved in many facets of tumorigenesis, including the cell cycle, apoptosis, energy metabolism, and other processes. To gain a more global view of the relationship between the genes most strongly bound by HSF1 in cancer cell lines, we generated an RNA expression correlation matrix through meta-analysis of pre-existing data sets (Figure 4B). We used the UCLA Gene Expression Tool (UGET) (Day et al., 2009) to query the extent to which the expression of each HSF1-bound gene correlated with every other HSF1-bound gene across all of the ~12,000 human expression profiles that have been generated with Affymetrix HG U133 Plus 2.0 arrays and made available through the Celsus database (Day et al., 2009). Hierarchical clustering of this gene-gene correlation matrix revealed five major transcription modules (Figure 4B).

The largest module was enriched for protein folding, translation, and mitosis. Genes within this dominant module showed the strongest positive correlation with the expression of HSF1 mRNA itself. Many of these genes had indeed proven to be regulated by HSF1 in our HSF1 shRNA knockdown experiments (Figures 2, 4A and S4). A second, smaller module was positively correlated with the first and strongly enriched for RNA-binding genes. Many of these genes, too, were positively regulated by HSF1 in our knockdown experiments (Figures 2, 4A, and S4). The remaining three modules (center to lower right of the matrix) were enriched for processes involved in immune functions, insulin secretion, and apoptosis. All three of these modules were negatively correlated with the largest module, suggesting negative regulation by HSF1.

Activation of HSF1 in a Broad Range of Cancer Specimens Taken Directly from Patients

Recently, we evaluated HSF1 expression and localization in a cohort of breast cancer patients culled from the Nurses’ Health Study (NHS) (Santagata et al., 2011). In that work, HSF1 was cytoplasmic and expressed at low levels in normal breast epithelial cells, but it accumulated in the nucleus of the majority of tumor specimens. Here, we confirm that finding (Figures 5A, S5), combining samples from two independent breast cancer collections representing all three major clinical subtypes (see Extended Experimental Procedures).

Next, because our ChIP-Seq analysis showed that the HSF1 cancer program is engaged not just in breast cancer cell lines but also in colon and lung cancer cell lines, we examined more than 300 formalin-fixed surgical specimens taken directly from patients. We included not only colon and lung cancer but also a wide variety of other tumor types.

Normal cells adjacent to the tumor demonstrated low HSF1 levels and cytoplasmic localization of the protein. In contrast, high-level expression of HSF1 and nuclear localization was common (Figure 5C) across every cancer type we examined, including carcinomas of the cervix, colon, lung, pancreas, and prostate as well as mesenchymal tumors such as meningioma. HSF1 staining was negative or weak in some tumors from each cancer type (Figure 5C). However, in those tumors, where expression was high, it was remarkably uniform across the sample, with nearly all tumor cells expressing similar levels of nuclear HSF1.

To determine whether the high-level nuclear localization of HSF1 detected by immunostaining was truly indicative of its
Figure 4. Distinct, Coordinately Regulated Modules of HSF1-Bound Genes

(A) Graphical representation of the HSF1 cancer program integrating information on gene binding, regulation and function. The peak height is reflected in the diameter of the circle (log2 peak height: range ~3 to 9) and color intensity reflects gene regulation (average of log2 fold change in BPLER and MCF7 cells upon HSF1 knockdown; red, positively regulated; green, negatively regulated; gray, no data available). Well-bound, differentially regulated genes as well as several genes of biological interest are displayed.

(B) Gene-gene expression correlation matrix of HSF1-bound genes. Pair-wise correlation map is presented of the genes that were bound by HSF1 in at least two of the three cancer cell lines (BT20, NCIH38, and SKBR3). The Pearson correlation coefficient relating normalized mRNA expression data for each gene pair was assessed in nearly 12,000 expression profiles. Enriched GO (gene-ontology) categories for each module are shown. See also Figure S4.
Figure 5. HSF1 Is Activated in a Broad Range of Human Tumors
(A) IHC shows strong nuclear HSF1 staining in human breast tumor cells (top) with adjacent normal breast epithelial cells (bottom) showing a lack of nuclear HSF1. 
(B) Images of HSF1 IHC on breast cancer tissue microarray (TMA) cores. Heat map shows scoring of three TMAs. The top panel depicts data from two TMAs (BRC1501 and BRC1502), containing 138 breast tumors of all major breast cancer subtypes. Progesterone receptor (PR), ER, and HER2 were also evaluated. The middle panel shows data from 161 triple-negative (TN) breast cancer cases. The bottom panel shows the lack of HSF1 nuclear expression in 16 normal mammary tissue sections. A summary is provided in the bar graph (right). 
(C) HSF1 IHC showing high-level nuclear staining in indicated tumors; T, Tumor; N, Normal adjacent tissue. A summary is provided in the bar graph.
activation, we obtained human tumor samples from breast and colon adenocarcinomas that had been cryopreserved and were of a quality suitable for ChIP-Seq analysis (Figures 5D and S5). Obtaining and processing such human tumor specimens for a technique as demanding as genome-wide ChIP-Seq is highly challenging. In addition, many potentially confounding factors are unavoidable (e.g., cell-type heterogeneity due to the presence of blood and stromal elements, areas of necrosis and micro-environmental stress). Despite these difficulties, the distinct HSF1-binding profile we had established with cultured cancer cell lines was clearly conserved in those tumors that expressed high levels of HSF1. Genes (such as ST13 and EIF4A2) that were strongly bound by HSF1 in cancer lines but weakly bound after heat shock in nontransformed cells, were also strongly bound in these human tumor samples taken directly from patients (Figure 5E). Genes that were weakly bound by HSF1 in cancer lines but strongly bound after heat shock in nontransformed cells (such as HSPA6 and DNAJC7) were also weakly bound in patient tumor samples (Figure 5E). These global similarities in HSF1-binding profiles between cancer cell lines and tumor samples, as well as their divergence from heat-shock profiles, were validated by principal component analysis (Figure 5F).

An HSF1-Cancer Signature Identifies Breast Cancer Patients with Poor Outcome

In our prior analysis of the Nurses’ Health cohort, HSF1 overexpression and nuclear localization was associated with reduced survival (Santagata et al., 2011). That work, however, was based entirely on HSF1 immunohistochemistry, an approach that is inherently only semiquantitative. To acquire more precise and molecularly defined information about the effects of HSF1 activation in cancer, we asked whether malignant potential and long-term outcomes correlate with the HSF1 transcriptional program identified above. We distilled an “HSF1-cancer signature” of 456 genes that were bound by HSF1 near their transcription start sites (Figure 2). Expression of these genes (Table S4) was interrogated in ten publicly available mRNA data sets derived from breast cancer patients that had been followed for an average of 7.58 years and had known clinical outcomes (referenced in Table S5). In total, these cohorts encompassed nearly 1,600 individuals of diverse national and ethnic origin. We divided each data set into two groups, those with high (top 25%) and those with low (bottom 75%) expression of the HSF1-cancer signature. We performed Kaplan-Meier analysis independently on each data set to assess potential associations between the HSF1-cancer signature and patient outcome: metastasis-free, relapse-free, or overall survival, depending on the reported outcome parameter for that data set. One representative analysis is presented in Figure 6A, the remainder are shown in Figure S6. High expression of our HSF1-cancer signature had a remarkable correlation with poor prognosis (HSF1-CaSig; Figures 6B and S6). In nine of ten independent data sets reported over the past 10 years, the p values ranged from 0.05 to < 0.0001.

Next, we considered a recent finding that many published cancer signatures are not significantly better outcome predictors than random signatures of identical size (Venet et al., 2011). We performed Kaplan-Meier analysis on independent data sets to evaluate associations between 10,000 individual randomly generated gene signatures and patient outcome (compiled data Table S4, example Figure 6C). A meta-analysis of the breast data sets showed that the HSF1-CaSig outperformed all 10,000 random gene signatures (Monte Carlo p value across breast data sets < 0.0001, Table S4).

Our HSF1-cancer signature was more broadly associated with outcome than other well-established prognostic indicators (Figures 6B and S6) including the oncogene MYC, the proliferation marker Ki67 and even MammaPrint, an expression-based diagnostic tool used in routine clinical practice (Kim and Paik, 2010). Because various HSPs have been implicated as prognostic markers for a range of cancers, including breast cancer (Ciocca and Calderwood, 2005), we also tested many individual HSP transcripts for possible association with outcome. None of these genes, or even a panel of HSP genes, was as strongly associated with poor outcome as our broader HSF1-cancer signature (Figures 6B and S6).
Figure 6. An HSF1-Cancer Signature Is Associated with Reduced Survival in Patients with Breast Cancer

(A) Representative data set (Pawitan et al., 2005) is shown from a meta-analysis of 10 publicly available mRNA expression data sets (Table S5) derived from human breast tumors with known clinical outcome and representing a total of 1594 patients. Each column corresponds to a tumor, and each row corresponds to a microarray probe for an HSF1-cancer signature (HSF1-CaSig) gene. Median levels of expression are depicted in black, increased expression in yellow, and decreased expression in blue. Tumors are ordered by average level of expression of the HSF1-cancer signature, from low (blue) to high (yellow). Red bars indicate deaths. Kaplan-Meier (KM) analysis of the tumors with high expression of the HSF1-cancer signature (top 25%, “High HSF1-CaSig,” yellow) versus low expressers (bottom 75%, “Low HSF1-CaSig,” blue) is shown.

(B) Log-rank p values for each of the indicated classifiers were calculated for each data set; results are displayed as a heat map. Corresponding KM curves are provided in Figure S6.

(C) Random gene signature analysis of a representative data set (Pawitan et al., 2005). KM analysis on the data set to evaluate associations between 10,000 individual randomly generated gene signatures and patient outcome. The random signatures are binned and ordered from least significant to most significant by correlation with poor outcome (test statistic).

(D) % of Patients surviving distant metastasis free

(E) % Survival

Note: (E) p = 0.0015

Correlation with poor outcome (test statistic)
HSF1-Cancer Signature Is Associated with Poor Outcome in Diverse Human Cancers

Next, we asked whether the HSF1-cancer signature might have prognostic value beyond breast cancer. Analyzing multiple independent gene expression data sets that include outcomes data, increased expression of the HSF1 cancer program in colon and lung cancers was strongly associated with reduced survival (Figures 7A and 7B). The HSF1-CaSig outperformed all 10,000 random gene signatures in these data sets (Monte Carlo p value across data sets < 0.0001, Table S4). Again, our HSF1-cancer signature was more significantly associated with outcome than any individual HSP transcript or even a panel of HSP genes (Figures 7B and S7). As expected, the MammaPrint expression signature, which was computationally derived by using breast cancers, was a poor indicator of outcome in lung and colon cancers (significant in one of four data sets). Additional HSF1 signatures also comprising positively and negatively regulated genes (HSF1-CaSig2) or containing both positively and negatively regulated genes (HSF1-CaSig3) were also strongly associated with patient outcome across tumor types (Table S4). We conclude that the HSF1 cancer program that we have identified supports the malignant state in a diverse spectrum of cancers because it regulates core processes rooted in fundamental tumor biology that ultimately affect outcome.

DISCUSSION

We have defined a distinct genome-wide transcriptional program that HSF1 coordinates in malignancy. This program includes some heat-shock proteins well known to be involved in oncogenic programs, such as HSP90 (Whitesell and Lindquist, 2005). However, it differs fundamentally from the HSF1 program induced by thermal stress, in that it includes many genes that are not induced by heat shock and does not include many that typically are. This cancer program is commonly activated in a wide variety of human malignancies. It is strongly associated with metastasis and death in at least the three cancers responsible for ~30% of all cancer-related deaths worldwide: those of the breast, colon and lung. The very broad range of tumors in which we see immunohistochemical evidence of HSF1 activation suggests it will play a pervasive role throughout tumor biology.

What types of cellular processes does HSF1 regulate in cancer? They constitute an astonishingly diverse group that extends far beyond protein folding and includes energy metabolism, cell cycle signaling, DNA repair, apoptosis, cell adhesion, extracellular matrix formation, and translation. Some of these processes were previously known to be affected by HSF1 (Dai et al., 2007; Jin et al., 2011; Solimini et al., 2007). However, the common assumption had been that HSF1’s effects were mediated primarily by HSP chaperone activities (Jin et al., 2011; Meng et al., 2010; Solimini et al., 2007). The remarkable breadth of the HSF1 cancer program in humans explains why HSF1 is such a powerful modifier of tumorigenesis in multiple animal models (Dai et al., 2007; Jin et al., 2011; Zhao et al., 2009) and why HSF1 was identified as one of only six potent metastasis-promoting genes in a genome-wide screen for enhancers of invasion by malignant melanoma cells (Scott et al., 2011). Not only is the repertoire of HSF1-regulated genes in cancer much larger than just heat-shock genes, but even the manner in which some of the classical heat-shock genes are regulated differs. For example, HSPA6 (HSP70B’), a pillar of the heat-shock response, differs dramatically in these two states. Following heat stress, HSPA6 is the most highly induced of all mRNAs, yet, surprisingly in cancer, HSPA6 is only bound very weakly by HSF1. Its expression is not changed following HSF1 depletion and its transcript level does not correlate with that of HSF1 in our meta-analysis of 12,000 gene expression experiments.
What could account for activation of a distinct HSF1-regulated program in cancer? After many years of investigation, we do not yet fully understand how HSF1 activity is regulated during the classic heat-shock response. Multiple mechanisms have been described. These include the release of HSF1 from its normal sequestration by chaperones when unfolded substrates compete for chaperone binding. But in addition, HSF1 is subject to an extensive array of posttranslational modifications (at least 30) including acetylation, sumoylation, and numerous phosphorylations (Anckar and Sirstonen, 2011).

Some of these heat-shock regulatory mechanisms are likely shared by cancer cells. For instance, impaired protein homeostasis driven by the accumulation of mutant, misfolding-prone oncoproteins, aneuploidy, and the increased rate of translation in cancer could chronically stimulate HSF1 activity by releasing it from sequestration from chaperones (Anckar and Sirstonen, 2011). Dysregulation of signaling pathways in cancer could also drive posttranslational modifications to HSF1. Some of these (such as those responsible for phosphorylation at serine 326) will likely be shared with heat-shocked cells. But others will likely be unique to cancer. Indeed, it seems extremely likely that different mechanisms of activation will operate in different cancers. Several pathways activated in cancer such as EGFR/HER2 axis (Zhao et al., 2009), the RAS/MAPK (Stanhill et al., 2006) or the insulin/IGF-like growth factor system (Chiang et al., 2012) have all been reported to alter HSF1 activity. Additional modes of cancer-specific regulation might include epigenetic states common to cancer and proliferating cells and transcriptional coregulators.

How might the distinct transcriptional program regulated by HSF1 in malignancy have arisen? The association of this program with metastasis and death points to an evolutionary origin distinct from cancer itself. The broad range of cancer types in which we find HSF1 activated suggests that this program originated to support basic biological processes. Indeed, the sole heat-shock factor in yeast (yHSF), even at basal temperatures, binds many genes that are involved in a wide-range of core cellular functions (Hahn et al., 2004). These transcriptional targets allow yeast not only to adapt to environmental contingencies but also to modulate metabolism and maintain proliferation under normal growth conditions (Hahn et al., 2004; Hahn and Thiele, 2004). As a result, yHSF is essential for viability, paralleling the importance of HSF1 for the survival and proliferation of cancer cells (Dai et al., 2007). Activation of HSF1 may also be required in animals in states of high proliferation and altered metabolism such as immune activation and wound healing (Rokavec et al., 2012; Xiao et al., 1999; Zhou et al., 2008). Moreover, in diverse eukaryotes, HSF1 is a well-validated longevity factor; nonclassical activation of this transcription factor could be highly relevant in this context (Chiang et al., 2012; Volovik et al., 2012).

Ironically, the evolutionarily ancient role played by HSF1 in helping cells to adapt, survive, and proliferate is co-opted frequently to support highly malignant cancers. By enabling oncogenesis, the activation of this ancient prosurvival mechanism thereby actually impairs survival of the host. HSF1 activation in a particular tumor may reflect the degree to which accumulated oncogenic mutations have disrupted normal physiology even before overt invasion or metastasis occurs. This interpretation would explain the impressively broad prognostic value of our HSF1-cancer signature across disparate cancers and even at early stages of disease. Clinical implementation will require further refinement of the signature and validation in tissue and RNA samples from multiple clinical cohorts. Such studies are certainly warranted. As just one potential application, it might aid in the identification of indolent tumors that do not require intervention, reducing the burdens of unnecessary treatment (Kalager et al., 2012). In addition to its prognostic value, HSF1 and diverse regulators that activate it might prove useful targets for cancer therapeutics.

Our understanding of the extensive role played by HSF1 in supporting cancers continues to mature. The protein has been defined for decades by its ability to coordinate chaperone protein expression and enhance survival in the face of heat stress (Christians et al., 2002; Ritosoa, 1962). Although appreciating the importance of these classical mechanisms, the role of HSF1 is clearly much broader and deeper.

**EXPERIMENTAL PROCEDURES**

**Cell Culture Methods**

Cell lines were cultured as detailed in Extended Experimental Procedures.

**ChIP-Seq and ChIP-qPCR**

ChIP-qPCR and ChIP-Seq experiments were performed as described previously (Lee et al., 2008), with modifications detailed in Extended Experimental Procedures.

**Gene Expression Analysis**

Lentiviral shRNA methods were described previously (Dai et al., 2007). Gene expression analysis was performed as described in Extended Experimental Procedures. Microarray data were deposited in NCBI Gene Expression Omnibus. RT-PCR analysis, gene-gene correlation analysis of HSF1-bound genes, and correlation of HSF1-bound gene expression with outcome is detailed in Extended Experimental Procedures.

**Immunohistochemistry and The NHS Analysis**

Paraffin sections were stained with HSF1 antibody (Thermo Scientific, RT-629-PABX) as detailed in Extended Experimental Procedures. The NHS is a prospective cohort study initiated in 1976 (Hu et al., 2011; Tamimi et al., 2008). For design and study population, and analysis, see Extended Experimental Procedures.

**Statistical Analysis**

Correlation of gene expression with location of HSF1 occupancy was performed by using a two-tailed Fisher’s Exact Test. P values for significance of overlap between pairs of gene sets were generated by using the hypergeometric distribution. Statistical methods for ChIP-Seq analysis and the Nurses’ Health Study outcome data analysis are detailed in Extended Experimental Procedures. Kaplan-Meier analysis was used to compare outcome events. P values were generated by using the logrank test. For all other data, mean ± standard deviation is reported and statistical significance between means was determined by using a two-tailed t test.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Extended Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2012.06.031.
ACKNOWLEDGMENTS

We thank G. Frampton, I. Barrasa and S. Gupta for bioinformatic assistance. We thank T. Mazor, T. Volkert and the WIBR-GTC for sequencing support. We thank the Lindquist lab and T. Lee for discussion and K. Allendoerfer, B. Bevis, G. Karras, R. Shouval and K. Maticak for comments. The work was supported by the J&J COSAT focused funding program (L.W.) and the Marble Fund (S.L.). S.L. is an investigator of the Howard Hughes Medical Institute. M.L.M. is supported by American Cancer Society New England Division-Spi-

REFERENCES


Hahn, J.S., and Thiele, D.J. (2004). Activation of the Saccharomyces cerevi-
siae heat shock transcription factor under glucose starvation conditions by Snf1 protein kinase. J. Biol. Chem. 279, 5169–5176.


Jin, X., Moskophidis, D., and Mivechi, N.F. (2011). Heat shock transcription factor 1 is a key determinant of HCC development by regulating hepatic stea-

Kalager, M., Adami, H.O., Bretthauer, M., and Tamimi, R.M. (2012). Overdiag-


Received: October 13, 2011
Revised: April 10, 2012
Accepted: June 4, 2012
Published: August 2, 2012


syst. 2, 627–639.


