

NORTHWESTERN UNIVERSITY

Elucidating the Role of Heat Shock Factor 2 (HSF2) in Cancer

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Driskill Graduate Training Program in Life Sciences

By

Roger S. Smith

EVANSTON, ILLINOIS

June 2022

© Copyright by Roger S. Smith, 2022

All Rights Reserved

ABSTRACT

Heat shock factor 1 (HSF1) is well known for its role in the heat shock response (HSR), where it drives a transcriptional program comprising heat shock protein (HSP) genes, and in tumorigenesis, where it drives a program comprising HSPs and many noncanonical target genes that support malignancy. Here, we find that HSF2, an HSF1 paralog with no substantial role in the HSR, physically and functionally interacts with HSF1 across diverse types of cancer. HSF1 and HSF2 have notably similar chromatin occupancy and regulate a common set of genes that include both HSPs and noncanonical transcriptional targets with roles critical in supporting malignancy. Loss of either HSF1 or HSF2 results in a dysregulated response to nutrient stresses in vitro and reduced tumor progression in cancer cell line xenografts. Together, these findings establish HSF2 as a critical cofactor of HSF1 in driving a cancer cell transcriptional program to support the anabolic malignant state.

ACKNOWLEDGMENTS

This work would not be possible without the generous mentorship and support of so many people. I would first like to thank my thesis advisor, Dr. Marc Mendillo, for cultivating a fun, inspiring research environment. He always challenged and encouraged me to pursue questions I found most interesting. Importantly, he believed in me during the times I struggled to believe in myself. Being able to assist with the Cold Spring Harbor Gene Expression course with Dr. Mendillo, will always be a special memory. I thank the members of the Mendillo Lab, past and present, for the shared learning experiences and riding the highs and lows of scientific research together. First, I thank Seesha Takagishi for being a great colleague and friend, for all her work laying the foundation of this project, and teaching me how to perform RNA-sequencing. I would like to thank Dr. Milad Alasady for all the guidance and wisdom she shared over the years. I'm very grateful to have worked alongside David Amici. As a friend and colleague, I learned so much from him, especially when troubleshooting data analysis pipelines, and brainstorming answers to "what's the mechanism" when either of us had fresh data to interpret.

I thank the members of my thesis committee, Dr. Sarki Abdulkadir, Dr. Dai Horiuchi, and Dr. Ben Singer for all their support, encouragement, and scientific guidance. I look forward to seeing results from productive ongoing collaborations begun with the laboratories of Dr. Abdulkadir and Dr. Horiuchi. I especially thank Dr. Ben Singer for his roles as a personal and professional mentor, and always being available for questions or advice. I would like to thank my past thesis committee member, Dr. David Guis, for valuable lessons in grantsmanship and helping me craft my fellowship application.

I'm grateful for the support of the Medical Scientist Training Program (MSTP) leadership and administrative team. First, I thank Dr. Dane Chetkovich and Dr. Jayms Peterson for recruiting

me to the Northwestern MSTP, and all their advice and guidance during the first few years. I would like to thank Dr. Hossein Ardehali for his leadership of the program, championing student input on programming, and serving as a professional mentor. I thank Dr. Melissa Brown for assistance in fellowship application writing. I thank Dr. Lindsey Martin and Joyce Tamanio for their counsel and tireless work to keep the MSTP running smoothly and improving. I thank the National Institutes of Health for prioritizing and expanding the funding of MD-PhD programs.

I would not be completing a PhD without pivotal early support from mentors in undergraduate research and as a research technician. I thank Dr. Mike Ferdig for letting me join his lab as a freshman and being my first example of how fun, if sometimes stressful, being a principal investigator can be. I'm grateful that Lisa Checkley taught me so many fundamentals of doing science well, and that she paired me with then graduate student, Dr. Geoffrey Siwo, when I shared that I was "interested in DNA". Dr. Siwo was an endlessly patient mentor as I began to grasp the ropes of conducting biomedical research and made my first foray into gene expression measurement and analysis. Dr. Siwo modeled the power of passionately following questions of interest, incorporating new technologies and ways of thinking as often as possible. I am also grateful to other professors and mentors at Notre Dame that believed in my pursuit of MD-PhD training and opened doors to research and networking opportunities, especially Dr. Michelle Whaley, Dr. Greg Crawford, and Dr. Norbert Wiech. I'm grateful to Dr. Marc Ladanyi for hiring me as a research technician at Memorial Sloan Kettering Cancer Center, during which time I experienced immense personal and professional growth, due in large part to Dr. Romel Somwar, who remains a very close mentor and dear friend.

Most importantly, my family and friends have been an enduring source of support, inspiration, and learning. My parents, Amy and Bernie Smith, always encouraged curiosity and fiercely supported my personal, academic, and professional development. They demonstrated their unconditional love, the importance of family, and a strong work ethic, all of which created the foundation upon which I have been able to pursue the medical and scientific training I desired. I am eternally grateful for all they have done to support me along the way. My sister, Kelly Fitzsimons, has been a great friend, and she continues to teach me so much about life, especially since the birth of my niece, Etta Grace. I am also extremely grateful for the friends I have made during my time so far in the MSTP, especially Zachary Chalmers, Maggie Coats-Thomas, Zachary Gaertner, Cooper Hayes, Mihai Truica, and Karthik Vasam. They have enriched my life in countless ways in and beyond the classroom, lab, or clinic. I am excited to see how you decide to employ your passions, skills, and enthusiasm. I look forward to collaborating on any topic with you all, but especially science, food, and travel.

To all the people who helped me become the person I am becoming. Thank you.

*To the people my physician-scientist training serves now, and in the future, so
they may heal to better share their unique gifts with the world.*

TABLE OF CONTENTS

ABSTRACT	3
ACKNOWLEDGMENTS	4
DEDICATION	7
TABLE OF CONTENTS.....	8
LIST OF RESULTS FIGURES	10
LIST OF TABLES	12
CHAPTER 1: INTRODUCTION	13
1.1 The Heat Shock Transcription Factors	13
1.2 Evolution of Transcription Factor Family Functions	20
1.3 Heat Shock Factors in Development	25
1.4 Heat Shock Factors in Neurodegenerative Disease	27
1.5 Heat Shock Factors in Cancer	29
CHAPTER 2: RESULTS.....	34
2.1 HSF2 interacts with HSF1 in diverse cancers.	34
2.2 HSF2 shares chromatin occupancy sites with HSF1 in cancer cells	39
2.3 HSF2 drives a pro-tumorigenic transcriptional program.....	43
2.4 Long-term loss of HSF2 and HSF1 results in sustained suppression of proteostasis gene expression conserved across cancers	51
2.5 HSF2 and HSF1 promote the transcriptional response to cancer-associated stress perturbations	55
2.6 HSF2 is required for tumor progression in cell line xenografts of prostate and breast cancer	58
CHAPTER 3: DISCUSSION.....	62
CHAPTER 4: MATERIALS AND METHODS.....	66
4.1 Multiple Sequence Alignment.....	66

4.2	Developmental Gene Expression of HSFs	66
4.3	Cell culture	66
4.4	Mass spectrometry (M/S) and analysis	67
4.5	LUMIER assay	67
4.6	Electrophoretic mobility shift assay (EMSA).....	68
4.7	Immunoprecipitation.....	70
4.8	Immunoblot	70
4.9	Protein purification.....	71
4.10	Quantitative immunoblot	72
4.11	HSF overexpression constructs.....	73
4.12	Gene silencing.....	74
4.13	CRISPR/Cas9 knockout generation.....	75
4.14	RNA harvesting, library prep, sequencing.....	77
4.15	RNA-seq data processing and analysis	78
4.16	ChIP-seq sample preparation and analysis	79
4.17	Coessentiality analysis	81
4.18	Xenograft experiments and analysis	81
4.19	Xenograft RNA sequencing and analysis.....	82
4.20	Data and materials availability.....	83
	REFERENCES.....	84
	VITA.....	102

LIST OF RESULTS FIGURES

Figure 1. HSF expression across the developmental lifespan.	14
Figure 2. Domain architecture of HSF family of transcription factors.	16
Figure 3. HSF2 interacts with HSF1 in diverse cancers.	35
Figure 4. Further characterization of HSF2 expression and function.	38
Figure 5. HSF2 shares chromatin occupancy sites with HSF1 in cancer cells.	40
Figure 6. Further characterization of ChIP-seq data.	42
Figure 7. HSF2 drives a pro-tumorigenic transcriptional program.	44
Figure 8. Validation of siRNA knockdown, HSF signature, and HSF quantitation.	47
Figure 9. Further characterization of dual siRNA treatment and additivity of HSF1 and HSF2 gene regulation.	49
Figure 10. Overexpressing the reciprocal HSF does not prevent effects of short-term depletion.	50
Figure 11. Long-term loss of HSF2 and HSF1 results in sustained suppression of proteostasis gene expression conserved across cancers.	52
Figure 12. Long-term loss of HSF1 or HSF2 results in cell-type specific adaptations beyond the consensus HSF cancer signature.	54

Figure 13. HSF2 knockout cancer cells have an intact heat shock response.	55
Figure 14. HSF2 and HSF1 promote the transcriptional response to cancer-associated stresses. 57	
Figure 15. HSF2 is required for tumor progression in cell line xenografts of prostate and breast cancer.	59
Figure 16. Further characterization of MDA-MB-231 xenograft tumors and RNA-seq.	61

LIST OF TABLES

Table 1. Oligo sequences for EMSA experiments.....	70
Table 2. siRNA Sample Notes	74
Table 3. sgRNA sequence for CRISPR/Cas9 knockout generation.....	77

CHAPTER 1: INTRODUCTION

1.1 The Heat Shock Transcription Factors

A cellular response to elevated temperature can be found universally throughout the tree of life (Lindquist, 1986). This heat shock response (HSR) results in the increased expression of heat shock proteins (HSPs) that serve as molecular chaperones to restore protein homeostasis, or proteostasis in the cell (Akerfelt et al., 2010a; Gomez-Pastor et al., 2018; Lindquist and Craig, 1988; Wu, 1995). The universal presence and remarkable conservation of some of the constituent genes highlights the biological significance of this response pathway. The first description of the HSR came in 1962 from observations of “puffs” induced in the chromosomes of salivary gland tissue in the fruit fly, *Drosophila brucekii*, treated with heat, dinitrophenol, or sodium salicylate (Ritossa, 1962). The rapid, easily induced nature of the HSR made it a model system for gene structure and transcription regulation (Guertin and Lis, 2010; Lindquist, 1986; Mahat et al., 2016; Vihervaara et al., 2018; Vihervaara et al., 2017). In 1984, two groups identified the first evidence for a transcriptional regulator of this response, the heat shock factor (HSF), by studying DNA-protein interactions in *Drosophila* nuclear extracts (Parker and Topol, 1984; Wu, 1984). These works demonstrated the binding of a protein, whose activity increased with heat, to a consensus sequence of DNA, the heat shock element (HSE), upstream of HSP genes to promote their expression (Parker and Topol, 1984; Wu, 1984). Subsequent studies revealed that invertebrates and ancestral organisms have a single, essential heat shock factor (HSF), but that vertebrates and plants have evolved a more complex family of HSFs (Fujimoto and Nakai, 2010).

The human genome encodes four, well-expressed HSFs, including HSF1, HSF2, HSF4, and HSF5. Additional protein-coding genes similar to HSFs exist on the sex chromosomes (*HSFX1*, *HSFX2*, *HSFY1*, *HSFY2*), but these genes are not expressed according to data from the

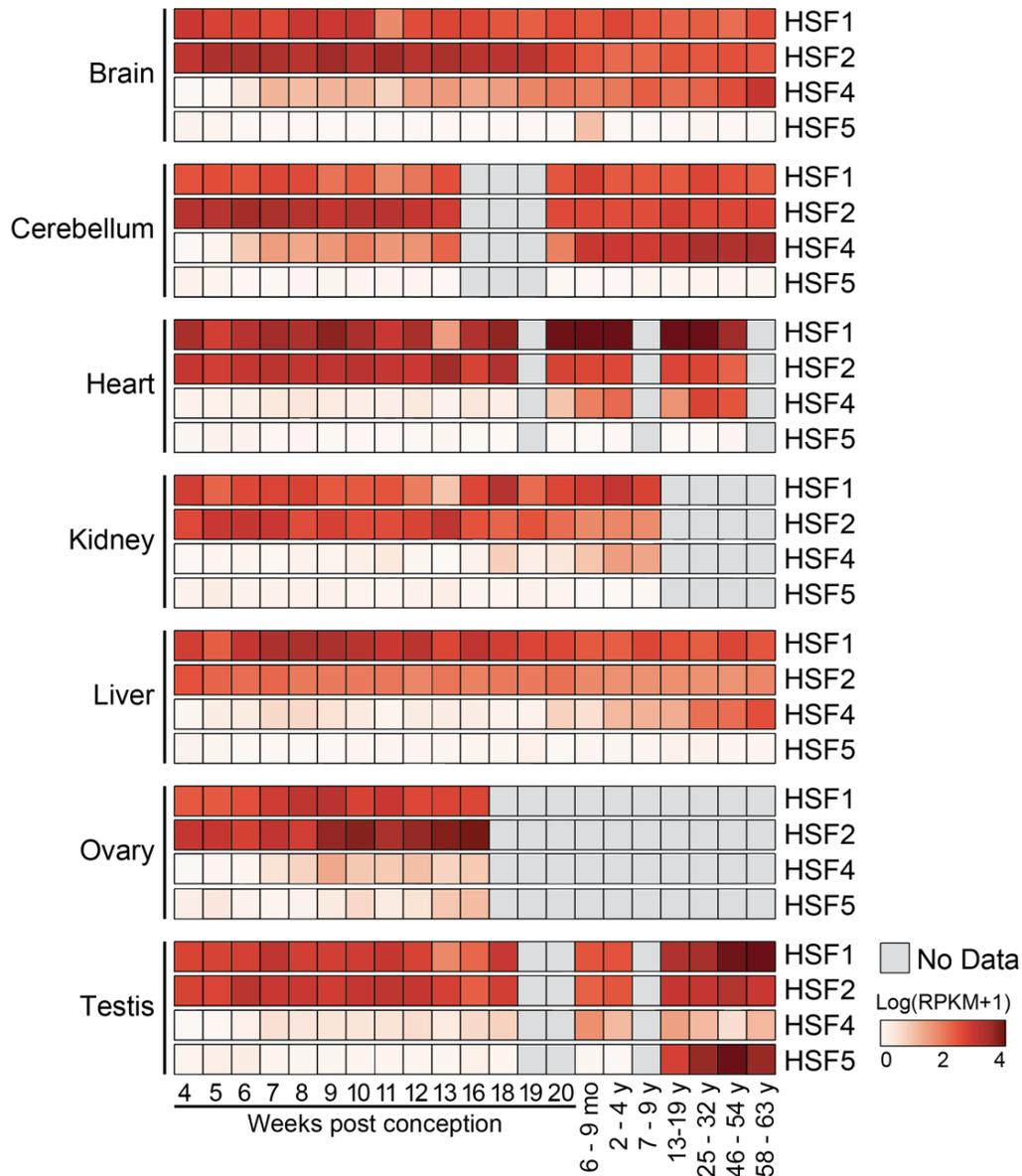


Figure 1. HSF expression across the developmental lifespan.

Human gene expression data obtained from the tissues and developmental timepoints indicated. RPKM gene expression data for human samples for each tissue and developmental timepoint were downloaded from <https://apps.kaessmannlab.org/evodevoapp/>. Replicates ranged from one to four for each tissue and timepoint. Data plotted represent the log normalized, mean of replicate RPKM values + 1.

Genotype-Tissue Expression (GTEx) Project. HSF1, HSF2, and HSF4 exhibit ubiquitous expression in human tissues, while HSF5 expression is restricted to the testis (GTEx Analysis Release V8) (Duchateau et al., 2020). A recent study obtained gene expression data from developing human tissues beginning four weeks post conception, through adulthood for cerebrum, cerebellum, heart, kidney, liver, ovary, and testis (Cardoso-Moreira et al., 2019). These data reveal a striking similarity in expression of HSF1 and HSF2 across tissues and developmental time, beginning at the earliest time points measured (Figure 1). HSF4 prenatal expression is predominantly observed in the brain and cerebellum, with more ubiquitous expression after birth. HSF5 expression is highly restricted to the testis and does not begin until around the time of puberty.

HSF1 is the most well-characterized of the family, owing to its ubiquitous expression and unique role as the master transcriptional regulator of the HSR (Akerfelt et al., 2010a; McMillan et al., 1998). Absent cell stress, HSF1 is sequestered in the cytoplasm, where its monomeric, inactive state is maintained through interactions with chaperone complexes composed of HSP40, HSP70, HSP90, and the cytosolic chaperonin TCP1 ring complex (TRiC) (Gomez-Pastor et al., 2018; Neef et al., 2014; Shi et al., 1998). Upon stress induction, HSF1 proteins translocate to the nucleus and trimerize to achieve the DNA-binding and transcriptionally competent conformation (Neudegger et al., 2016). While it is believed that diverse mechanisms may regulate this activation, *in vitro* trimerization in response to elevated temperature suggests that HSF1 may possess intrinsic temperature sensing capacity (Hentze et al., 2016).

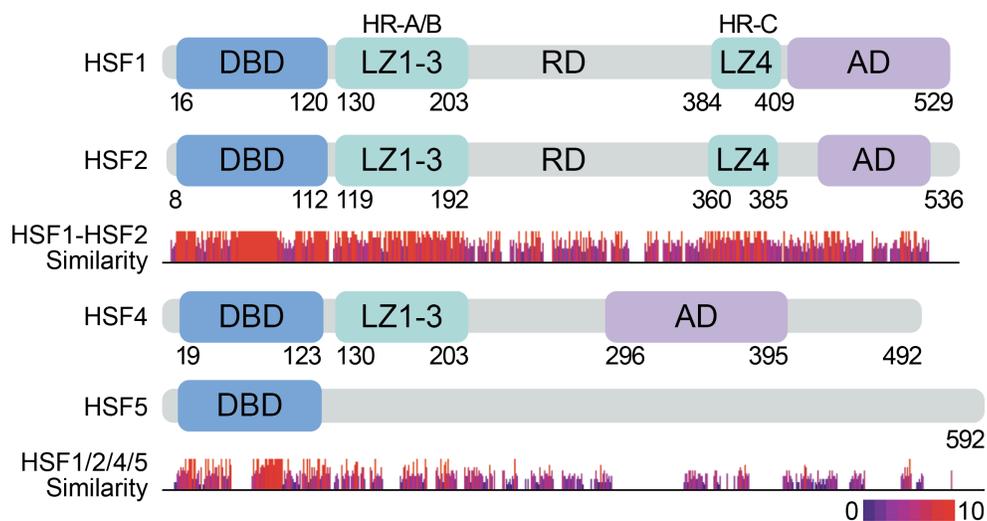


Figure 2. Domain architecture of HSF family of transcription factors.

HSF proteins contain an amino-terminal, winged-helix-turn-helix DNA-binding domain (DBD). Leucine Zipper domains (LZ) contain heptad repeat (HR) regions, with LZ1-3 required for HSF oligomerization. LZ4, containing HR-C, suppresses HSF activity through intramolecular interactions with LZ1-3 in HSF1 and HSF2. Transcriptional activity for some HSFs requires the activation domain (AD). Sequence similarity was determined using Clustal Omega Multiple Sequence Alignment with protein amino acid sequences obtained from Ensembl. Sequence similarity is scored from 0 to 10, as indicated by the height and color of the bars under the respective HSF diagrams.

Trimerization of HSF1 and other HSF family members occurs through interaction of leucine zipper domains (LZ), containing two heptad repeats (HR-A and HR-B) of hydrophobic and charged amino acid residues (Gomez-Pastor et al., 2018; Littlefield and Nelson, 1999; Sorger and Nelson, 1989; Zuo et al., 1994). A third heptad repeat region (HR-C) in the carboxy-terminal end of the protein forms inhibitory, intramolecular interactions disrupting oligomerization of HSF1 (Farkas et al., 1998; Rabindran et al., 1993). Importantly, these HR domains are highly conserved between HSF1 and HSF2 (Figure 2). Transcriptional activation functions map to the carboxyl terminal domain in HSF1, HSF2, and the HSF4 β isoform (Green et al., 1995; Nakai et al., 1997; Tanabe et al., 1999; Yuan et al., 1997; Zhu and Mivechi, 1999). In addition to protein-intrinsic mechanisms of regulation, HSF1 is subject to myriad post translational modifications including phosphorylation, acetylation, and sumoylation, with a significant proportion of these modifications lying in the central, regulatory domain (Gomez-Pastor et al., 2018). The contexts and consequences governing these regulatory modifications are only beginning to be elaborated.

In 1991, heat shock factor 2 (HSF2) was first identified in human HeLa cells and shown to bind an HSE and induce transcription (Schuetz et al., 1991; Skloot, 2010). Phylogenetic analyses of HSF evolution suggest the HSF2 gene appeared with the evolution of jawed vertebrates around 473 million years ago (Cunningham et al., 2022; Fujimoto and Nakai, 2010). HSF2 harbors functional domains that are highly similar to those found in its paralog HSF1, including a DNA-binding domain (DBD) and heptad repeat domains (HR-A/B and HR-C) (Fujimoto and Nakai, 2010; Gomez-Pastor et al., 2018). Consistent with nearly identical DBD sequences, HSF2 can bind the consensus HSE originally defined for HSF1, consisting of alternating, inverted repeats of *nGAAn* DNA sequences, where *n* denotes any DNA base (Gomez-Pastor et al., 2018). Indeed

crystal structures of human HSF1 and HSF2 proteins bound to DNA reveal highly similar HSF-DNA binding interactions (Jaeger et al., 2016; Neudegger et al., 2016). Interestingly, these studies also revealed that when bound to DNA, HSF1 and HSF2 expose biochemically distinct protein sequences (Gomez-Pastor et al., 2018; Jaeger et al., 2016). Moreover, crystal structures of HSF1 and HSF2 heterocomplexes bound to DNA revealed different flexibility of the wing domain and differences in DNA conformational changes induced by binding (Feng et al., 2021). These regions could provide the substrate for paralog-specific regulation by post-translational modification or recruitment of other proteins, despite nearly identical DNA binding (Gomez-Pastor et al., 2018; Jaeger et al., 2016).

Elucidating the primary functions of HSF2 has been less clear than for HSF1, despite their structural similarities. Initially, HSF2 was reported to have only a limited role in promoting the heat shock response despite an ability to bind consensus heat shock elements (HSEs) *in vitro* and *in vivo* (Jaeger et al., 2016; Mahat et al., 2016; Mathew et al., 2001; Sistonen et al., 1994; Sistonen et al., 1992). The earliest functional studies of HSF2 defined its activation of HSP gene expression in response to hemin-induced differentiation of K562 erythroleukemia cells (Sistonen et al., 1992). This study also suggested that compared to HSF1, HSF2's contribution to HSP expression in response to heat stress was much weaker (Sistonen et al., 1992) and dependent on its interplay with HSF1 (Ostling et al., 2007; Sistonen et al., 1994). Consistent with a role for HSF2 beyond heat stress, HSF2 has been reported to function in spermatogenesis (Wang et al., 2004; Widlak and Vydra, 2017), and modulates the response to select proteotoxic stresses including proteasome inhibition (Joutsen et al., 2020; Mathew et al., 1998), ethanol, and febrile-range thermal stress (El Fatimy et al., 2014; Shinkawa et al., 2011).

Intriguingly, despite HSF2's lesser role in heat stress, exogenously expressed human HSF2 is sufficient to rescue viability of yeast depleted of their single, essential HSF, whereas human HSF1 was unable to rescue (Liu et al., 1997). However, disruption of the carboxy-terminal regulatory domain that maintains HSF1 in an inactive, monomeric state rescued yeast viability. This result suggests that as the ancestral HSF evolved into additional factors, HSF1 gained a layer of regulation while HSF2 retains ancestral mechanisms of activation. These results in the context of crystal structure data supporting the possibility for distinct regulation despite similar DNA binding, suggest that the interaction and regulation of HSF1 and HSF2 activity are highly context dependent. In fact, HSF1 and HSF2 can form heterotrimers when bound to DNA (Ostling et al., 2007; Rabindran et al., 1993; Sandqvist et al., 2009), and the composition and post translational modifications of these heterocomplexes may distinguish the heat responsive activity from the developmental functions of HSFs (Sandqvist et al., 2009). Much remains to be learned about how the stoichiometry of the active HSF heterocomplex (e.g., one HSF1 and two HSF2 molecules or vice versa) influences transcriptional output in specific cellular contexts.

The remaining members of the HSF family of proteins have been less well-characterized. HSF4 contains a similar DBD, but lacks properties of a transcriptional activator (Nakai et al., 1997). HSF4 is highly expressed in the eye lens where it is required for growth and differentiation, and mutations in HSF4 result in cataract development (Fujimoto et al., 2004; Fujimoto et al., 2008). HSF5, HSF6, and HSF7 all share sequence similarity to the conserved HSF DBD, but the function of these proteins remain largely uncharacterized (Gomez-Pastor et al., 2018). Together the HSF family represents a clear example of a single protein's function being adapted to specific biochemical environments. The striking evolutionary conservation of core components of the

original HSF is a testament to the fundamental role these proteins play in cellular processes, as well as human health and disease.

1.2 Evolution of Transcription Factor Family Functions

Exploring the evolutionary patterns of paralogous genes can inform expectations about the functions of transcription factors from the same family. Genes are considered paralogs if they result from a gene duplication event (Koonin, 2005). Duplication can occur through a whole genome duplication (WGD) or a small-scale duplication (SSD) (Conant and Wolfe, 2008). Gene duplications occur relatively frequently on evolutionary time scales, with reports citing rates of 0.1 duplications per gene per one million years (Lynch and Conery, 2000). As early as 1918, it was proposed that the duplication of genes provides a robust mechanism for the diversification of function (Bridges, 1935; Taylor and Raes, 2004). In theory, with a second copy of a critical gene, there would be less selective pressure on the ‘backup’ copy, such that mutations spawning novel functions could arise in the duplicate without disrupting the original function (Conant and Wolfe, 2008; Taylor and Raes, 2004). Impressively, 50% of prokaryotic and 90% of eukaryotic genes are products of duplication events (Teichmann and Babu, 2004). Indeed, as evolutionary genetics began to reveal the patterns of gene duplications more broadly across the tree of life, evidence supporting duplications as the engines of functional innovation for speciation events grew (Koonin, 2005; Ohno, 1970; Taylor and Raes, 2004). Even without the tools of modern genetics, Charles Darwin reasoned that “natural selection...should have seized on a certain number of the primordially similar elements, many times repeated, and have adapted them to the most diverse purposes” (Darwin, 1859).

The development of a new, selectively beneficial function not existing prior to the duplication is termed neofunctionalization (Conant and Wolfe, 2008). While there are undoubtedly a number of prominent examples of this process, evolutionary genomic analyses reveal that most duplicate genes do not confer new abilities or adaptations (Conant and Wolfe, 2008; Koonin, 2005; Lynch and Katju, 2004). Indeed, many duplicate genes are silenced or lost within a few million years of their creation (Lynch and Conery, 2000). Understanding why some duplicates are retained rather than lost can reveal potential adaptive characteristics of those genes.

For a duplicated gene to be retained in the genome, the combination of selective advantages and random genetic drift need to outweigh any negative consequences on organismal fitness. The major patterns of paralog activity include subfunctionalization, co-option of function, and dosage selection. Subfunctionalization occurs when each paralog performs only a subset of the functions performed by the ancestral gene (Conant and Wolfe, 2008). Co-option occurs when a gene's previous function, such as enzymatic activity, is applied to a new biochemical process. Importantly, co-option can occur for both the major and minor functions of a given gene, such that minor functions become selectively advantageous in a changed environment (Conant and Wolfe, 2008). Dosage selection can result in retention of duplicate genes in situations where extra expression provides a selective advantage in the face of stress or environmental shifts. While this occurs slowly in evolution, the process is analogous to gene amplifications mediating drug resistance in tumors, such as *MET* amplification mediating resistance to EGFR-targeted therapy in lung cancer (Engelman et al., 2007). Interestingly, absent a selective pressure to select for the increased gene dosage after a duplication, having too much of a protein can be disadvantageous. In response, an early cellular adaptation to duplication is down regulation of expression of both

paralogs (Lan and Pritchard, 2016). Since gene expression changes occur rapidly, cells or organisms can avoid immediate, adverse consequences. Because mutations accrue over far longer time scales, the dosage sharing of paralog expression allows both copies of the gene to persist in the genome while mutations (either advantageous or deleterious) accumulate and are selected for or against (Lan and Pritchard, 2016). These mutations can then lead to the raw material for neo- or sub-functionalization and co-option of ancestral functions (Conant and Wolfe, 2008). This is consistent with patterns observed in fungi where the most common divergence between paralogs is in their respective regulation rather than biochemical function (Conant and Wolfe, 2008; Lan and Pritchard, 2016; Wapinski et al., 2007). Taken together, these models of paralog evolution support a rapid adaptation to a second gene copy of identical function, which creates the opportunity for evolutionarily slow adaptive or deleterious functions.

Although many retained paralogs experience similar evolutionary trajectories, the types of protein duplicates that are retained is not evenly distributed across biochemical functions. Studies assessing the functions of genes retained long after whole genome duplications in yeast, plants, insects, bacteria, and humans, have consistently revealed an enrichment for transcription factors (Aury et al., 2006; Blanc and Wolfe, 2004; Conant and Wolfe, 2008; Harbison et al., 2004; Taylor and Raes, 2004). Considering the ability of transcription factors to regulate hundreds or thousands of target genes, retaining duplicate transcription factors could enable more diverse adaptations than an individual protein with a narrow biochemical function. Another category of genes experiencing more volatile gains or losses of duplicates over evolutionary time is those involved in stress responses (Wapinski et al., 2007). Intriguingly, studies of the effects of whole genome duplication on yeast functional networks for duplicated genes revealed that partitioning of function amongst

duplicate genes can result in one gene holding a more predominant stress-responsive role than its paralog (Conant and Wolfe, 2006). In the context of HSF1 being the predominant heat shock responsive member of the HSF family, this transcription factor family may indeed be a quintessential example of paralogous transcription factor evolution as has been suggested (Jaeger et al., 2016).

While evolutionary processes may partition or optimize subsets of a paralog's function, protein sequence and structure are strong predictors of function. Specifically for transcription factors, most transcription factors within the same family exhibit highly similar patterns of DNA binding specificity (Berger et al., 2008; Jolma et al., 2013; Rogers and Bulyk, 2018). Thus, the null hypothesis when studying the function of two paralogous transcription factors should be that they share genome occupancy and potentially function. Of course, transcription factors are composed of regulatory domains beyond the DNA-binding domain (DBD) that can influence multimer formation or recruitment of additional proteins (Rogers and Bulyk, 2018). Moreover, small changes to the amino acid sequence of the DBD between paralogs can modify the affinity with which each factor binds DNA. As a result, both paralogs may bind similarly at high-affinity DNA recognition sites, while diverging in their ability to bind minor, lower-affinity sites (Rogers and Bulyk, 2018). Indeed, this phenomenon has been demonstrated for mouse Lhx family homeodomain transcription factors (Berger et al., 2008). It is believed that retaining function at high affinity sites and permitting divergence at low-affinity sites would be more permissive to new environments, without sacrificing core functions (Rogers and Bulyk, 2018). This pattern as a mechanism of evolution in paralogous transcription factors has been demonstrated in yeast and *E. Coli* (Teichmann and Babu, 2004).

The similarity of binding activity of transcription factor paralogs has been validated with high-throughput molecular screens. As an example, expression of hundreds of human full-length transcription factors, human DBDs, and mouse DBDs, enabled the systematic assessment of the interactions of these proteins or domains with random oligonucleotide sequences (14 – 40 base pairs in length) to determine DNA recognition capacity (Jolma et al., 2013). Comparing the binding preferences of full-length and DBD-only constructs revealed that analyzing the DBD alone is sufficient for determining transcription factor binding for all assayed proteins but one (ELK1) (Jolma et al., 2013). These experiments revealed that HSF1, HSF2, and HSF4 clustered uniquely by family, for their binding specificity to the canonical HSE. This experimental system also had the power to assess evolutionary conservation of DBD sequence and binding specificity between mouse and human. They found that evolution occurs very slowly in DBDs but do observe some divergence in transcription factor dimer orientation and spacing preferences. The authors propose that DNA sequences flanking core transcription factor recognition motif can influence these characteristics of binding and thus may evolve more rapidly than binding motifs themselves (Jolma et al., 2013). These results are consistent with the notion of first preserving core functions while allowing for fine-tuning or slight modifications in the duplicated protein domain.

Gene duplication is a critical substrate for adaptations to new environments and refining protein function. Studying the features of paralogs that have been maintained and modified across evolutionary time is a powerful lens for considering shared or divergent protein function, especially for classes of protein preferentially retained after duplication events, such as transcription factors (Taylor and Raes, 2004). For the HSF family of transcription factors, similar DBD architecture and DNA-binding preferences anchor predictions of shared functions. However,

significant deviation in other surface features raises the possibility of unique regulation. Thus, elucidating the contexts in which their shared ancestry dominates from environments where partitioned adaptations are arising will be exciting frontiers for future studies of HSF biology.

1.3 Heat Shock Factors in Development

Beyond their namesake role in response to thermal stress, HSFs serve critical functions in organismal development. This role was first observed when knockout of the single HSF in *Drosophila melanogaster* disrupted oogenesis and larval development (Jedlicka et al., 1997). These effects did not result from significant changes in HSP gene expression (Jedlicka et al., 1997), consistent with observations of embryogenesis in the first *Hsf1* knockout mice (Xiao et al., 1999). Nonetheless, depletion of HSF1 and HSF2 in some genetic backgrounds of mouse models have revealed profound consequences on development, especially of brain and reproductive tissues (Akerfelt et al., 2010a; Duchateau et al., 2020). The range of phenotypes observed in different model systems likely reflects redundancy functional redundancy. Understanding how HSFs regulate these developmental processes in greater detail will shed light on the breadth of their biological activity.

HSF1 and HSF2 are critical regulators of development in reproductive tissues. Although HSF1-deficient mice survive to adulthood, they experience increased prenatal lethality, growth deficits, and female infertility (Xiao et al., 1999; Zhang et al., 2002). This results, at least in part, from HSF1's regulation of HSP90 α , which is required for completion of oocyte meiosis (Metchat et al., 2009). Subsequent work revealed mouse oocytes lacking HSF1 exhibited mitochondrial damage and an increased susceptibility to oxidative stress, resulting in increased cell death (Bierkamp et al., 2010). HSF2 null female mice similarly suffer from meiotic defects, resulting in

reduced ovarian follicles, abnormal eggs, and hemorrhagic cystic follicles (Kallio et al., 2002). In male mice, HSF2 knockout results in smaller testes, reduced sperm count, and misshapen spermatid heads (Kallio et al., 2002; Wang et al., 2003). Indeed, HSF2 is expressed at high levels in mouse testis, where it binds chromatin at genes responsible for sperm quality (Akerfelt et al., 2008; Sarge et al., 1994). Dual knockout of *Hsf1* and *Hsf2* in male mice resulted in even more pronounced phenotypes including the complete lack of mature sperm and resultant sterility (Wang et al., 2004). Studies demonstrating HSF1-HSF2 heterotrimers in testes that synergistically regulate sex chromosomal multi-copy genes in post-meiotic round spermatids provide a mechanistic basis for the more drastic consequences of the double knockout (Akerfelt et al., 2010b; Sandqvist et al., 2009), and underscore the importance of studying these factors in tandem.

Disruption of HSF1 and HSF2 function also has severe consequences for brain and central nervous system development in mice. HSF1 depletion results in enlarged ventricles, astrogliosis, accumulated ubiquitinated proteins, and myelin loss (Homma et al., 2007; Santos and Saraiva, 2004). In addition, olfactory epithelium atrophies and olfactory sensory neurons die, absent HSF1's cytoprotective role (Takaki et al., 2006). HSF2 is highly expressed actively binding DNA in early brain development (Duchateau et al., 2020). Depletion of HSF2 disrupts neuronal migration resulting in abnormal cortical lamination, or layering of neurons in the cerebral cortex (Chang et al., 2006; El Fatimy et al., 2014; Kallio et al., 2002; Wang et al., 2003). These effects are mediated through HSF2's direct regulation of p35 which regulates cortical migration signaling (Akerfelt et al., 2010a; Chang et al., 2006). These data support a role for HSFs regulating genes beyond the canonical HSPs. Moreover, they elaborate a complex interplay between HSF1 and HSF2 in diverse cellular contexts to support proper organismal development and stress resilience.

1.4 Heat Shock Factors in Neurodegenerative Disease

A hallmark of neurodegenerative diseases is protein misfolding and aggregation. Not surprisingly, HSF1 and its target genes have a well-established link to neurodegenerative disease (Anckar and Sistonen, 2011). While mutation or loss of HSF1 is not causative of these diseases, disruption of protein folding machinery and chaperones exacerbates neurodegenerative phenotypes and disease progression (Gomez-Pastor et al., 2017; Gomez-Pastor et al., 2018). Reduced HSF1 activity and protein levels are observed in ageing and many protein-folding based neurodegenerative diseases. These diseases include polyglutamine expansion diseases, such as Huntington's (HTT), as well as Parkinson's, Alzheimer's, and amyotrophic lateral sclerosis (ALS) (Gomez-Pastor et al., 2018).

In a mouse model of Huntington's disease, *Hsf1* knockout led to increased accumulation of mutant huntingtin protein and shortened lifespan (Hayashida et al., 2010). This result could be reversed in cell and mouse models with the expression of constitutively active HSF1, reducing aggregate formation and improving lifespan (Fujimoto et al., 2005). In Parkinson's disease, α -synuclein (α -syn) aggregates lead to the progressive loss of dopaminergic neurons in the substantia nigra. Modeling this disease by over-expressing α -syn in human cells and mouse models is sufficient to result in HSF1 depletion, further exacerbating disease progression (Kim et al., 2016). Alzheimer's disease results in the accumulation of amyloid- β , resulting in the loss of neurons in the hippocampus and cerebellum (Gomez-Pastor et al., 2018; Mavroudis et al., 2010). In both patients with Alzheimer's and mouse models, reduced levels of HSF1 and chaperones are observed in the cerebellum (Jiang et al., 2013). Aggregates of another protein, TDP43, are seen in up to 50% of patients with Alzheimer's disease and nearly all cases of familial ALS (Gomez-Pastor et al.,

2018). HSF1 knockout mice exhibit TDP43 accumulation and ALS-like symptoms and increased susceptibility to other aggregating proteins in ALS (Chen et al., 2016). Importantly, restoring HSF1 activity genetically or pharmacologically or overexpressing the chaperone proteins to augment the protein folding milieu has demonstrated promise in alleviating protein aggregates and neurodegenerative phenotypes in these diverse model systems (Fujimoto et al., 2005; Gomez-Pastor et al., 2017; Gomez-Pastor et al., 2018; Neef et al., 2011; Neef et al., 2010; Westerheide et al., 2004). These observations fuel optimism for restoring protein homeostasis in these disease contexts via HSF and its protein quality control target genes.

Each of these diseases exhibit enhanced degradation of HSF1 during disease progression. In Huntington's disease models, it has been shown that elevated activity of casein kinase 2 (CK2) can promote the degradation of HSF1 through phosphorylation at Ser303 and Ser307 (Gomez-Pastor et al., 2017; Gomez-Pastor et al., 2018). Inhibition of CK2 led to stabilized HSF1 levels, increased chaperone gene expression, and a concomitant reduction in huntingtin protein aggregation (Gomez-Pastor et al., 2017). CK2 levels are also elevated and correlate with disease progression in Parkinson's, Alzheimer's, and ALS patient samples and mouse models (Gomez-Pastor et al., 2018). Mechanistic studies of CK2 activity in settings beyond Huntington's disease will refine therapeutic strategies and identify other contributing factors.

Studies of other HSF family members are notably sparse in the neurodegeneration literature. Knockout of HSF2 in mouse cell models of Huntington's disease revealed increased protein aggregation and reduced lifespan, as previously described for HSF1 (Shinkawa et al., 2011). However, no studies to-date have assessed interplay of HSF1, HSF2 and other HSFs in the setting of neurodegenerative disease. Considering the shared DNA occupancy and regulation of

chaperones, HSF2 could play an important role here too. For example, in a study of HSF1 genome-wide activity in Huntington's disease, changes in gene expression for cytoskeletal and adhesion genes were observed, despite no changes in HSF1 DNA binding (Riva et al., 2012). Interestingly, recent reports have revealed a role for HSF2 in regulating cell-cell adhesion genes, which could explain the findings in Huntington's disease (Joutsen et al., 2020). Both HSF2 and HSF4 exhibit expression throughout brain tissues, especially in the cerebellum where neurodegenerative pathology can manifest (GTEx data) (Carithers et al., 2015; Duchateau et al., 2020). These observations highlight the need for more comprehensive evaluation of HSFs in the context of neurodegeneration which will inform the understanding of HSF biology and potentially reveal new approaches for therapy in these diseases.

1.5 Heat Shock Factors in Cancer

Evidence accumulating over the last decade has revealed that the gene expression program driven by HSF1 can vary dramatically depending on the context in which it is activated. One of the most profound and well-established examples of this is the discovery of HSF1 supporting tumor initiation and progression. *Hsf1*-knockout mice have a decreased incidence of tumors when subjected to oncogenic stimuli including RAS mutations, p53 loss, or chemical carcinogens (Dai et al., 2012; Dai et al., 2007; Jin et al., 2011; Kourtis et al., 2018; Meng et al., 2010; Min et al., 2007; Xi et al., 2012). Patient tumors of diverse histopathological origin exhibit elevated levels and nuclear localization of HSF1 (Alasady and Mendillo, 2020; Dai and Sampson, 2016; Mendillo et al., 2012). Indeed, HSF1 activity serves as predictive biomarkers for patient outcomes, with elevated levels correlating with poor prognosis (Alasady and Mendillo, 2020; Bjork et al., 2018; Dai, 2018; Dai et al., 2007; Mendillo et al., 2012; Santagata et al., 2011). Tumor gene expression

data demonstrate elevated expression of HSF2 and HSF4 in lung, breast, liver, esophageal and colorectal cancers (Meng et al., 2017; Puustinen and Sistonen, 2020; Zhong et al., 2016), but decreased expression of HSF2 in others such as prostate cancer (Bjork et al., 2016; Chen et al., 2022). However, few studies have evaluated the functional consequences of HSF2 and HSF4 expression differences in tumors.

The mechanisms by which expression of HSFs differs in tumors compared to normal and the consequences of those differences in activity are only beginning to emerge. In cancer, HSF1 promotes gene expression not only of canonical HSP target genes, but also many non-canonical target genes with roles in diverse biological processes including cell-cycle regulation, translation, adhesion, and metabolism (Alasady and Mendillo, 2020; Mendillo et al., 2012). In breast, lung, and colon cancers, this “HSF1 cancer signature” is associated with metastasis and patient death (Mendillo et al., 2012; Santagata et al., 2011). In this manner, HSF1 is an example of a “non-oncogene addiction”. While it does not cause cancer itself, its elevated activity allows cancer cells to address the heightened demand on protein quality control in a rapidly proliferating cell subjected to the additional stresses of aneuploidy and nutrient restriction (Dai and Sampson, 2016; Luo et al., 2009; Solimini et al., 2007). As an example, hyperactive HSF1 in estrogen receptor alpha positive (ER α +) breast cancer leads to upregulation of HSP90, which chaperones the folding of ER α along with kinases and other proteins essential for rapid proliferation (Vydra et al., 2019; Whitesell et al., 2014; Xi et al., 2012). Indeed, elevated HSF1 potentially mediates resistance to antiestrogen therapies in breast cancer (Silveira et al., 2021).

HSF1’s impact on cancer biology extends beyond cancer-cell autonomous effects. Cancer associated fibroblasts (CAFs) are essential to cultivating a tumor microenvironment permissive to

growth and metastasis. HSF1 is activated in CAFs of patient tumors and correlates with poor prognosis (Scherz-Shouval et al., 2014). Depleting HSF1 from CAFs reduces xenograft tumor proliferation. These effects result from HSF1 directing a transcriptional program that is distinct, but complementary, to that previously defined in cancer cells themselves (Scherz-Shouval et al., 2014). Considering the pleiotropic roles of HSF1 in cancer cells as well as the supporting stroma, it will be an attractive target for cancer therapy with promising initial results in leukemia and prostate cancer mouse xenograft models (Dong et al., 2020; Santagata et al., 2013).

Studies of HSF2 function in cancer are limited to only a few recent reports and suggest that HSF2 can both promote and suppress cancer cell growth in different cancer models (Bjork et al., 2016; Meng et al., 2017; Zhong et al., 2016). Recent work has described HSF2 enhancing breast cancer tumorigenesis through an interaction with ZEB1 to promote an epithelial to mesenchymal transition associated with aggressive cancers (Li et al., 2014). HSF2 regulation of ALG3 has also been shown to promote the proliferation and migration of MCF-7 breast cancer cells (Yang et al., 2018). Interestingly, each of HSF1, HSF2, and HSF4 can regulate hypoxia-inducible-factor-1 α (HIF-1 α) (Chen et al., 2011; Gabai et al., 2012), which increases vascular endothelial growth factor (VEGF), a potent signaling molecule for new blood vessel formation and target of cancer therapies. These studies of HSF2 and HSF4, along with most studies of HSF1 in cancer, have focused on one or the other HSF without regard to any potential interplay, but understanding the potential cooperative or antagonistic interactions of these transcription factors is crucial (Roos-Mattjus and Sistonen, 2021). Indeed, whether these disparate effects of HSF2 are due to a role directly regulating transcription or through a functional link to HSF1's well-established pro-tumorigenic functions is unclear (Puustinen and Sistonen, 2020). Therefore, rigorous characterization of

HSF2's role in cancer and its mechanistic interplay with HSF1 are required to shed light on how HSFs support cancer.

In the following chapters, studies undertaken to increase our understanding of HSF1's role in cancer reveal a prominent cooperation with its paralog, HSF2 (Smith et al., 2022). We demonstrate that HSF2 physically interacts with HSF1 in cancer cells, resulting in an indistinguishable pattern of genome occupancy. Across breast, prostate, colon, and lung cancers, HSF2 regulates expression of HSF-bound genes, including both canonical HSPs and non-canonical factors involved in cell metabolism and proliferation. Regulation of these genes by HSF2 is critical for the cellular response to cancer-associated stressors in a manner distinct from its dispensable role in the HSR. Lastly, HSF2 is required for cell line xenograft progression, and as a result, HSF2 loss extends tumor-bearing mouse survival. These findings identify HSF2 as a critical accomplice of HSF1 in driving pro-tumorigenic gene expression programs and indicate that the role of HSF2 in cancer biology is far greater than previously appreciated.

*The work in the following chapters, including figures, is reproduced with minor modifications from: Smith RS, Takagishi SR, Amici DR, Metz K, Gayatri S, Alasady MJ, Wu Y, Brockway S, Taiberg SL, Khalatyan N, Taipale M, Santagata S, Whitesell L, Lindquist S, Savas JN, Mendillo ML. HSF2 cooperates with HSF1 to drive a transcriptional program critical for the malignant state. *Sci Adv.* 2022;8(11):eabj6526. Epub 20220316. doi: 10.1126/sciadv.abj6526.*

CHAPTER 2: RESULTS

2.1 HSF2 interacts with HSF1 in diverse cancers.

To identify protein interacting partners of HSF1 in cancer, we stably expressed an HSF1 tagged with a 3xFLAG-V5 epitope in a panel of four cancer cell lines and used anti-FLAG beads to perform immunoprecipitation (IP) followed by liquid chromatography tandem mass spectrometry (MS)-based proteomic analysis. HSF2, the highest expressed HSF1 paralog in human tumors (Figure 4A), emerged as a high confidence interacting protein in each of the four cell lines tested (Figure 3A). We subsequently performed the reciprocal experiment and recovered HSF1 from HSF2-FLAG in two of the three lines tested. Although we were unable to detect HSF1 in HSF2 IP in ZR-75-1 cells, we identified lower amounts of HSF2 itself, perhaps due to a mechanism limiting HSF2 expression or IP efficiency in that cell line. Additional interacting proteins include those involved in cellular metabolism and cytoskeletal organization, but none of these interacting proteins are observed as consistently as HSF2 in the cancer cell lines tested (Figure 4B).

As an orthogonal strategy to identify HSF1-interacting proteins, we adopted LUMIER, a quantitative, high-throughput protein-protein interaction assay previously used to study chaperone-client interactions (Taipale et al., 2012; Taipale et al., 2014). Briefly, we transiently expressed a collection of FLAG-tagged plasmids (“the bait”) in a 293T cell line that stably expressed HSF1 tagged with Renilla luciferase (“the prey”). The bait collection consisted of 2853 unique clones comprising most transcription factors along with many other genes involved in chromatin regulation, protein synthesis, and post translational modification (Table S2). After lysis, the bait proteins were individually captured with anti-FLAG coated 384-well plates, and luminescence (i.e., HSF1 concentration) was measured. Of the 2853 clones, HSF2 – a ubiquitously expressed HSF1 paralog – was the top HSF1-interacting transcription factor (Figure 3B; Figure

4C, Table S2). Overall, only the HSF1 control (due to homo-oligomerization) and two *HSP70*-family proteins, known to strongly interact with HSF1 (Alasady and Mendillo, 2020; Pincus, 2020), scored higher. We further validated these findings using co-IP of endogenous HSF1 and immunoblot analysis of endogenous HSF2 in a panel of cancer cell lines including breast, prostate, and lung cancer (Figure 3C). Thus, HSF1 and HSF2 physically interact in cancer cells.

To explore the functional significance of this protein interaction in cancer cells, we performed an electrophoretic mobility shift assay (EMSA) to test whether HSF2 and HSF1 can

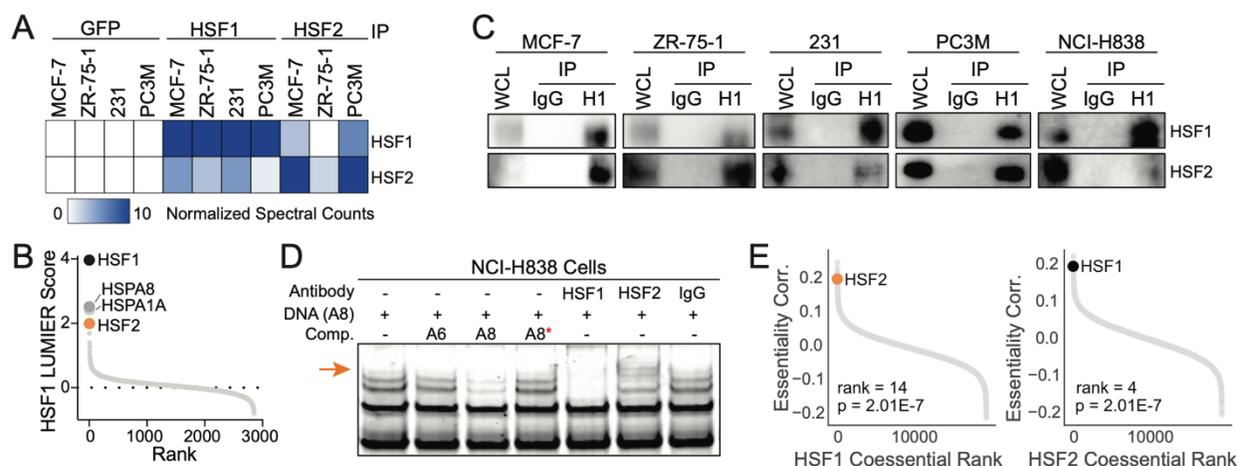


Figure 3. HSF2 interacts with HSF1 in diverse cancers.

(A) IP-MS of green fluorescent protein (GFP), HSF1, or HSF2 in four cancer cell lines, indicated. Data for HSF1 and HSF2 are plotted as the number of spectral counts normalized to GFP control. (B) HSF1 LUMIER assay reveals HSF2 as the top HSF1-interacting transcription factor. Plot shows ranked HSF1 LUMIER score (GFP-normalized, log₂-transformed luminescence values). (C) IP for HSF1 (H1) followed by immunoblot for HSF1 and HSF2. WCL, whole-cell lysates as input; 231, MDA-MB-231. (D) EMSA assay using HSPA8 (A8) as bait. Reactions were incubated with HSF1, HSF2, or control IgG antibody as indicated. A6 and A8 represent promoter sequences used for HSPA6 or HSPA8, respectively. A8* indicates a mutated HSE in the competing HSPA8 DNA promoter sequence. An arrow is included to highlight the level of supershift in HSF2 antibody lane. (E) Coessentiality rank (x axis) and correlation coefficient (y axis) for all genes with HSF1 (left) or HSF2 (right). The position of HSF2 or HSF1 in the coessentiality plot is indicated (orange or black dot, respectively).

each bind the canonical HSE found within the promoter of *HSPA8*, a site bound by HSF1 with high affinity in cancer cells (Figure 3D). A DNA duplex harboring the *HSPA8* HSE was incubated with cancer cell lysates in the presence or absence of a control IgG, HSF1, or HSF2 antibodies. Proteins in cell lysates interacted with the *HSPA8* promoter, causing a specific electrophoretic band shift that could be competed away with a 20x excess of unlabeled *HSPA8* HSE, but not with a mutant *HSPA8* HSE containing three substitutions in nucleotides critical for HSF1 binding or with the HSE of *HSPA6* (an HSF1 target gene in the HSR that is minimally bound by HSF1 in cancers contrasting heat shock) (Mendillo et al., 2012). Moreover, incubation with HSF1- or HSF2-specific antibodies but not IgG resulted in a super shift, suggesting that the interacting proteins HSF1 and HSF2 can both bind to this same target DNA. The HSF1-HSF2-HSE complex formed using lysates from cancer cells grown under basal conditions was far more robust than that observed for non-tumorigenic mouse embryonic fibroblasts (MEFs), which required treatment with proteasome inhibitor, MG132, to stimulate HSF-HSE binding (Figure 4D). These results suggest that HSF2 forms an active complex with HSF1 capable of binding DNA in proliferating cancer cells grown.

Physically interacting proteins with cooperative roles in the same process are often critical for optimal cell growth in the same cellular contexts, and thus frequently display shared knockout fitness profiles in large-scale genetic screening studies (Boone et al., 2007; Pan et al., 2018; Wang et al., 2017). To investigate the extent to which HSF1 and HSF2 have similar fitness requirements in cancer cells, we applied a bias-adjusted, rank-based co-essentiality approach to whole-genome dependency data from over 700 diverse cancer cell lines (Amici et al., 2021; Meyers et al., 2017; Tsherniak et al., 2017). Our analysis revealed that HSF1 and HSF2 have strongly correlated

patterns of cancer cell essentiality ($r = 0.184$, $p = 2.01e-7$, Figure 3E). Indeed, HSF2 is the 14th ranked co-essential gene with HSF1, and the highest ranked transcription factor (Figure 3E, left). Similarly, HSF1 is HSF2's 4th rank most similarly essential gene (Figure 3E, right). Taken together, the physical interaction, cancer cell HSE-binding activity, and correlated cancer cell essentiality profiles of HSF1 and HSF2 suggest that these factors have a critical shared function in cancer.

MBP-HSF2 (HSF2) was analyzed by SDS-PAGE with Coomassie stain. **(F–H)** Gels are loaded with equal amounts (1 μ g) of purified HSF1 or HSF2 as indicated and immunoblotted with the indicated antibodies for HSF2 (F), HSF1 (G), or HIS-tag (H).

2.2 HSF2 shares chromatin occupancy sites with HSF1 in cancer cells

For a more global view of HSF2 function in cancer, we assayed its chromatin occupancy using chromatin immunoprecipitation followed by sequencing (ChIP-seq) in two aggressive cancer cell lines: a triple negative breast cancer cell line (MDA-MB-231) and a metastatic prostate cancer cell line (PC3M). We first validated our HSF2 antibody by performing immunoblot analysis using purified recombinant polyhistidine-tagged maltose binding protein (HIS-MBP)-HSF1 and HIS-MBP-HSF2 to demonstrate antibody specificity and lack of paralog cross-reactivity (Figure 4E-H). Using a previously validated HSF1 antibody (Mendillo et al., 2012) and the validated HSF2 antibody for ChIP-Seq, we identified 677 regions bound by HSF2 or HSF1, with a strong pattern of co-occupancy (Figure 5A, 6C). Motif analysis demonstrated a strong enrichment for the consensus HSE ($p = 1e-217$) (Figure 5B). Among bound regions observed in MDA-MB-231, most occurred at intergenic or intronic loci (Figure 6B). PC3M cells yielded similar results with 794 regions bound by either HSF2 or HSF1, many of which were co-occupied (Figure 5F). Additionally, at regions called peaks only for HSF1, HSF2 was present at lower levels as demonstrated in all heatmaps which plot the union of HSF1 and HSF2 peaks and depict indistinguishable occupancy patterns (Figure 5F, 6G, 6I). Motif analysis of these regions similarly revealed the HSE as the most significantly enriched motif ($p=1e-31$, Figure 5G). Most binding occurred at promoter sites in PC3M, similar to other cell lines with high levels of HSF1 activation (Figure 6H, (Mendillo et al., 2012)). Co-occupied genes in both cell lines included HSP targets

(e.g., *HSP90AB1*, *HSPA8*, *HSPA1B*) as well as non-HSPs (e.g., *RBM23*, *CKS2*) (Figure 5C and 5H).

To determine the ability of either HSF2 or HSF1 to bind chromatin absent its paralog, we performed ChIP-seq for each factor in cells with the other factor knocked out (e.g., HSF2 ChIP in HSF1 knockout cells) (immunoblots in 6F and 6J). In MDA-MB-231 cells, loss of HSF1 resulted in a global reduction in HSF2 chromatin occupancy (Figure 5D, 6D). This result is consistent with previous data demonstrating that HSF2 requires HSF1 for maximal *HSP70* promoter binding (Ostling et al., 2007). In HSF2 knockout cells, there was an increase of HSF1 bound to chromatin compared with controls, perhaps reflecting a replacement of HSF2 in active heteromeric complexes (Figure 5E, 6E). Importantly, we do not observe changes in the genomic loci where

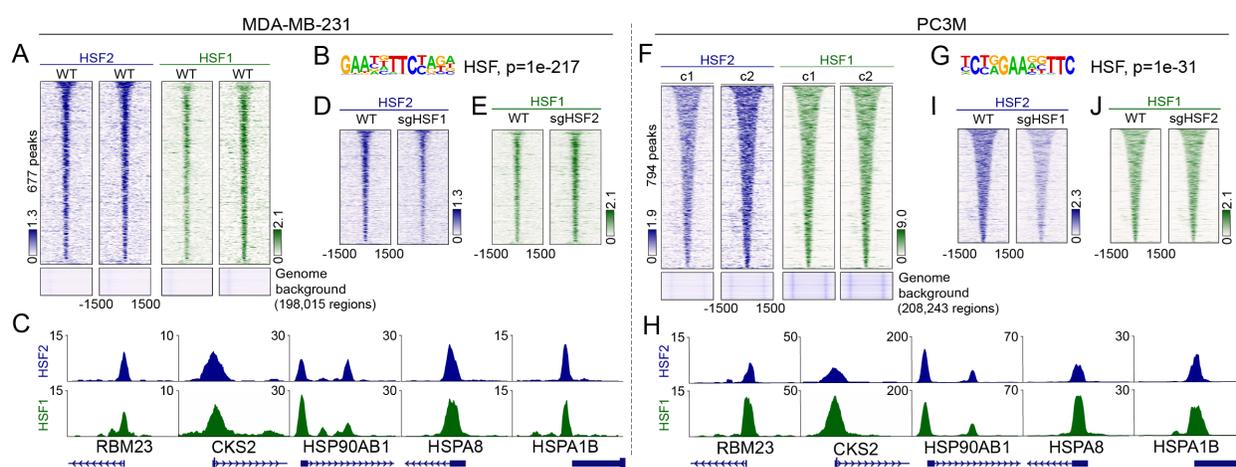


Figure 5. HSF2 shares chromatin occupancy sites with HSF1 in cancer cells.

(A, F) ChIP-seq for HSF2 or HSF1 in the indicated cell line. Dual tracks indicate biological replicates. (B, G) Enriched motifs of ChIP peaks. (C, H) Example ChIP-seq tracks for HSF2 and HSF1 at five loci implicated in HSF cancer programs. (D, E and I, J) ChIP-seq tracks for HSF1 or HSF2 in HSF knockout cells as indicated. These data merged all replicates for ease of visualization. WT = wild-type, parental cell lines. c1 and c2 represent cell lines expanded from a single wild-type cell clone as control for clonal CRISPR knockouts.

this HSF1 homo-oligomer binds when HSF2 is depleted. In PC3M cells, loss of HSF1 also resulted in reduced HSF2 chromatin occupancy (Figure 5I). In this cell line, however, some loss of HSF2 chromatin occupancy is attributable to reduced total HSF2 protein levels (Figure 6J). HSF1 chromatin occupancy in HSF2-depleted cells was not noticeably changed (Figure 5J).

Taken together, these results suggest that while HSF2 is dispensable for HSF1 chromatin occupancy, HSF1 can promote HSF2 chromatin occupancy through either its effects on HSF2 protein stability or on HSF2 DNA binding (Figures 8A and 11C) (Mathew et al., 1998; Ostling et al., 2007; Sandqvist et al., 2009; Santopolo et al., 2021). More broadly, the nearly identical HSF2 and HSF1 chromatin occupancy landscape in two distinct types of cancer cell lines with different levels of HSF1 and HSF2 activity, along with their highly correlated effects on cell growth across hundreds of human cancer cell lines, supports the likelihood that these factors regulate transcription of similar genes in cancer cells.

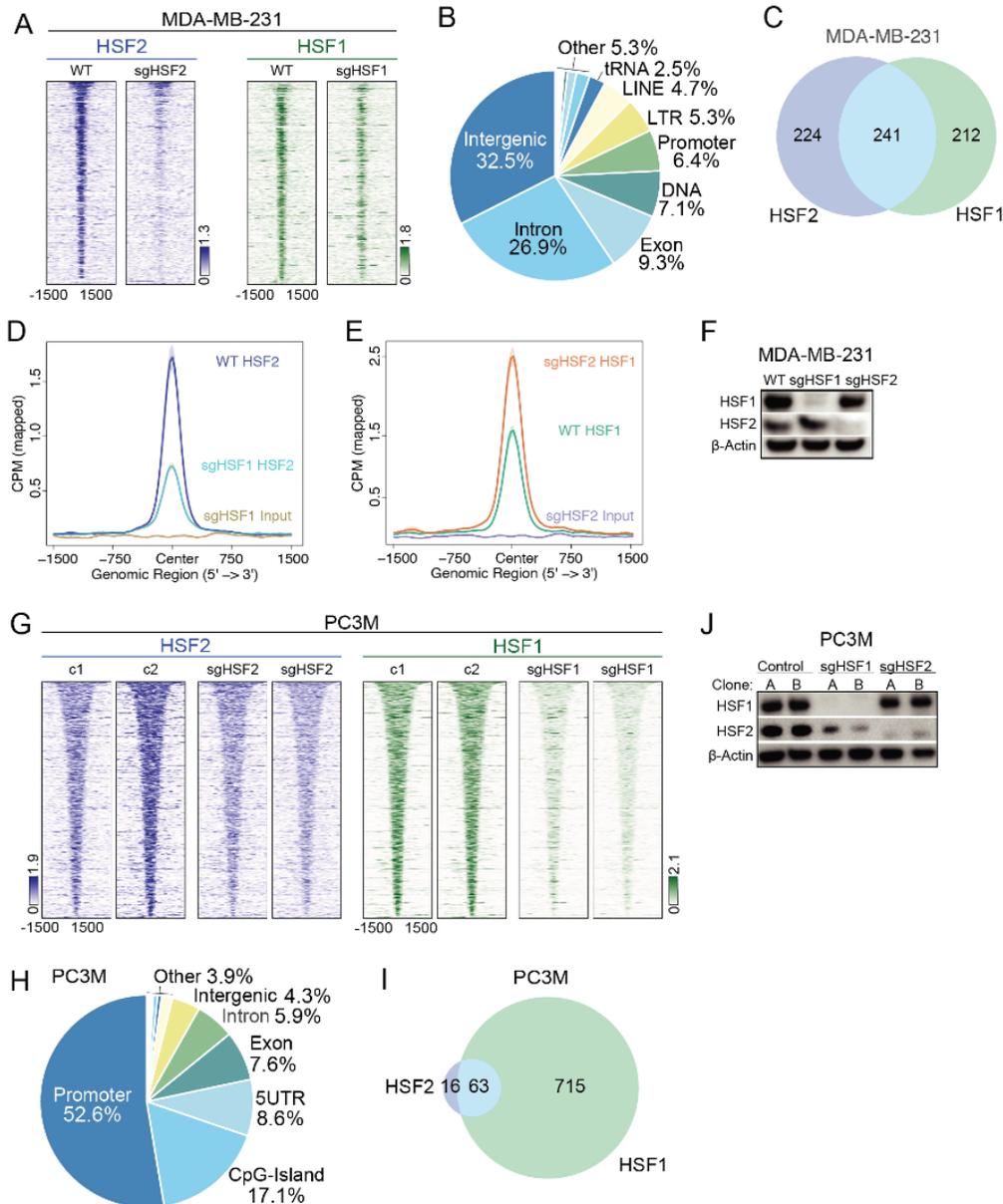


Figure 6. Further characterization of ChIP-seq data.

For both MDA-MB-231 and PC3M ChIP-seq experiments, bound peaks were called relative to knockout cell lines shown (**A** and **G**). Genomic annotations of ChIP peaks for MDA-MB-231 and PC3M are shown in panels **B** and **H**. **C**) Depicts the overlap of peaks called for HSF2 and HSF1 in MDA-MB-231 cells, with significant enrichment of shared peaks. **D**) Metagene plot for HSF2 ChIP in WT and sgHSF1 MDA-MB-231. **E**) Metagene plot for HSF1 ChIP in WT and sgHSF2 MDA-MB-231 cells. **F**) Immunoblot of MDA-MB-231 WT and knockout cells. **I**) Venn diagram depicting unique and shared peaks for HSF2 and HSF1 in PC3M. **J**) Western blot analysis of PC3M clonal knockout cells used in ChIP-seq experiments and xenografts (Figure 15).

2.3 HSF2 drives a pro-tumorigenic transcriptional program

To investigate how HSF2 affects cancer cell gene expression, and how this relates to HSF1 activity, we performed RNA sequencing (RNA-Seq) to measure gene expression in a panel comprising eleven cancer cell lines treated with short-interfering RNA (siRNA) pools targeting either HSF1, HSF2, or non-targeting (NT), control siRNAs. The cancer cell lines profiled cover the most common and lethal cancers: luminal and basal breast cancers (Neve et al., 2006), along with prostate, lung and colon cancers. We focused our assessment of gene expression changes on those most significant (FDR-adjusted p-value < 0.05) and conserved across these diverse cell line backgrounds (see Methods). Using these stringent parameters, knockdown of HSF2 resulted in 32 or 10 genes with either reduced or elevated levels, respectively. HSF1 knockdown resulted in 167 or 164 genes with either reduced or elevated levels, respectively.

K-means clustering of the union of differentially expressed genes upon acute depletion of either HSF revealed five major clusters (Figure 7A). Clusters 1 and 2 contained genes whose expression decreased upon HSF2 and HSF1 depletion, indicating that they are positively regulated by both HSFs. Gene set enrichment analysis (GSEA) of this group of genes revealed a strong enrichment for genes annotated as cell cycle (FDR-adjusted Q-value (Q) = 4.3E-18), HSF1 activation (Q = 1.0E-11), RNA binding (Q = 7.1E-11), and protein folding (Q = 1.3E-07). Importantly, these clusters reveal similar regulation of gene expression by HSF2 and HSF1. Not all genes in these clusters are significantly differentially expressed for HSF2 depletion across all cell lines using an FDR-adjusted p-value less than 0.05, however HSF2 depletion still reduces expression of most target genes in these clusters (Figure 7A, 3B). Cluster 3 contained genes generally positively regulated by HSF1 but not HSF2 and enriched for cell cycle processes (Q = 1.4E-37) and protein folding (Q = 3.8E-07). Clusters 4 and 5 contained many genes that had

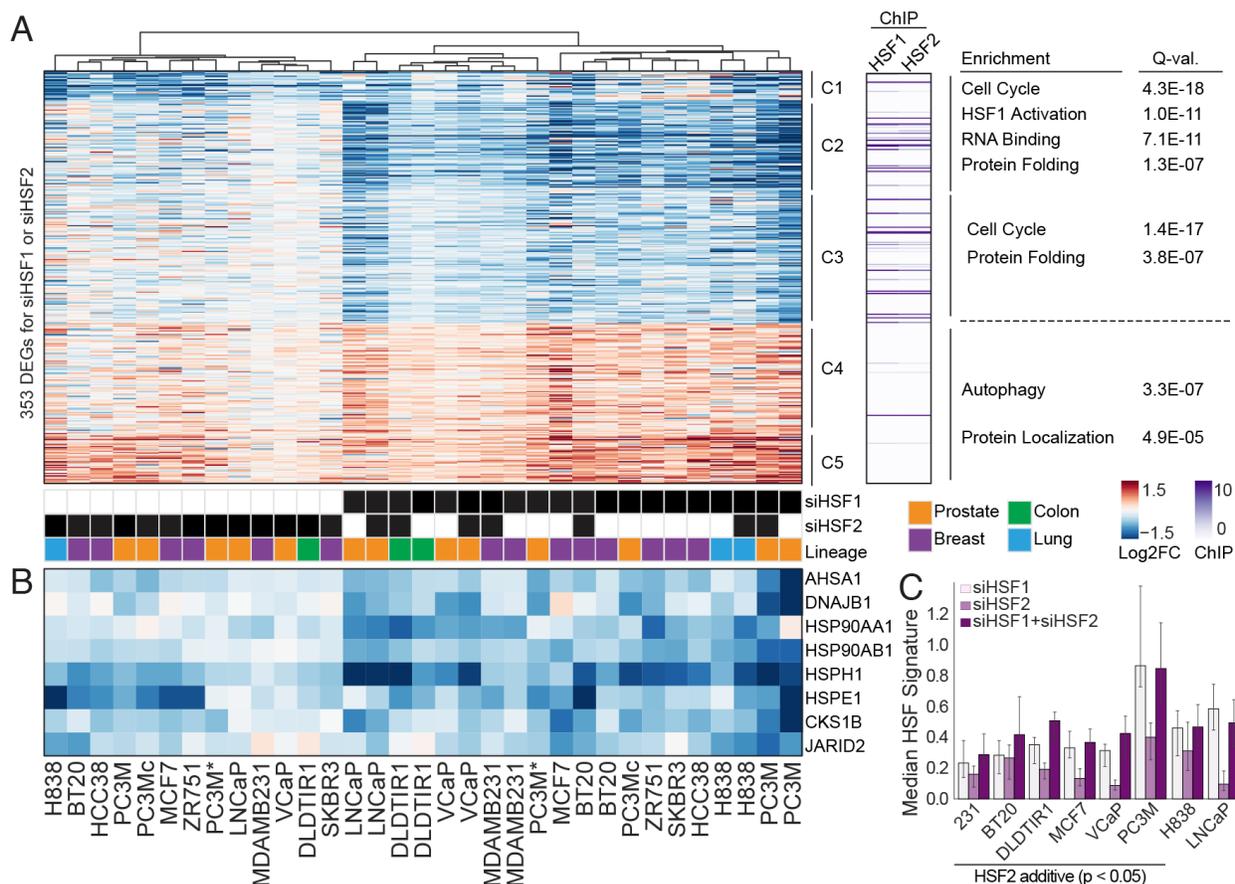


Figure 7. HSF2 drives a pro-tumorigenic transcriptional program.

A) RNA sequencing data of 11 cancer cell lines treated with either siRNA targeting HSF1 or HSF2 or both. Data are expressed as \log_2 fold change (L2FC) versus the non-targeting control in the respective cell line. The heatmap contains genes called significant for either siHSF1, siHSF2, or both by EdgeR. Relative ChIP-seq binding intensity is plotted to the right of the heatmap for both HSF1 and HSF2. Gene set enrichment analysis was performed and select enriched GO terms are displayed along with their FDR-adjusted p-value (Q-val.). **(B)** Select target genes from (A) are highlighted. **(C)** Median absolute L2FC signature strength for ChIP-bound genes in (A) for cell lines receiving single and double siRNA treatment. Data are plotted as median \pm 95% confidence interval. Additivity was tested (see Methods) and indicated where $p < 0.05$, see also Figures 9B and 9D. PC3Mc = clonally derived, wild-type PC3M. PC3M* = wild-type population with On-Target Plus siRNA as opposed to siGenome siRNA for “PC3M”. 231 = MDA-MB-231.

increased expression upon HSF2 or HSF1 depletion. These genes were enriched in processes that include autophagy ($Q = 3.3E-07$), and protein localization ($Q = 4.9E-05$). Notably, HSF2 regulates both canonical HSP targets (e.g., *HSP90AB1*, *HSPH1*, *HSPE1*) and non-HSP (e.g., *CKS1B*, *JARID2*) targets previously attributed to HSF1 in cancer gene expression programs highlighted in Figure 7B. Thus, the highly concordant changes in the cancer cell transcriptome resulting from loss of either HSF are, consistent with a model of cooperative gene expression regulation.

To identify direct transcriptional targets of HSF2 and HSF1, we integrated our ChIP-seq and RNA-seq data (Figure 7A, right heatmap). Importantly, many regulated genes in clusters 1 through 4 are also bound by HSF2 and HSF1. In clusters 4 and 5, fewer genes are bound, suggesting that most of the increases in gene expression conserved across cancers occur as a downstream consequence of HSF loss. Indeed, focusing only on HSF-bound genes highlights an extraordinarily similar program of cancer cell gene regulation by HSF2 and HSF1, with many bound genes exhibiting reduced expression upon depletion of either HSF, despite some variation in the strength of regulation (Figure 8B). Taken together, the chromatin binding and gene expression data demonstrate that HSF2 activates many of the same targets as HSF1 in cancer. Thus, considering the previously established link between HSF1 activity in cancer cells, its pro-tumorigenic phenotypes in animal models, and the expression of HSF1's transcriptional targets as a coherent program in human tumors associated with poor clinical outcomes in patients (Dai et al., 2007; Kourtis et al., 2018; Mendillo et al., 2012; Santagata et al., 2011), these data suggest HSF2 likewise promotes this transcriptional program critical for the malignant state.

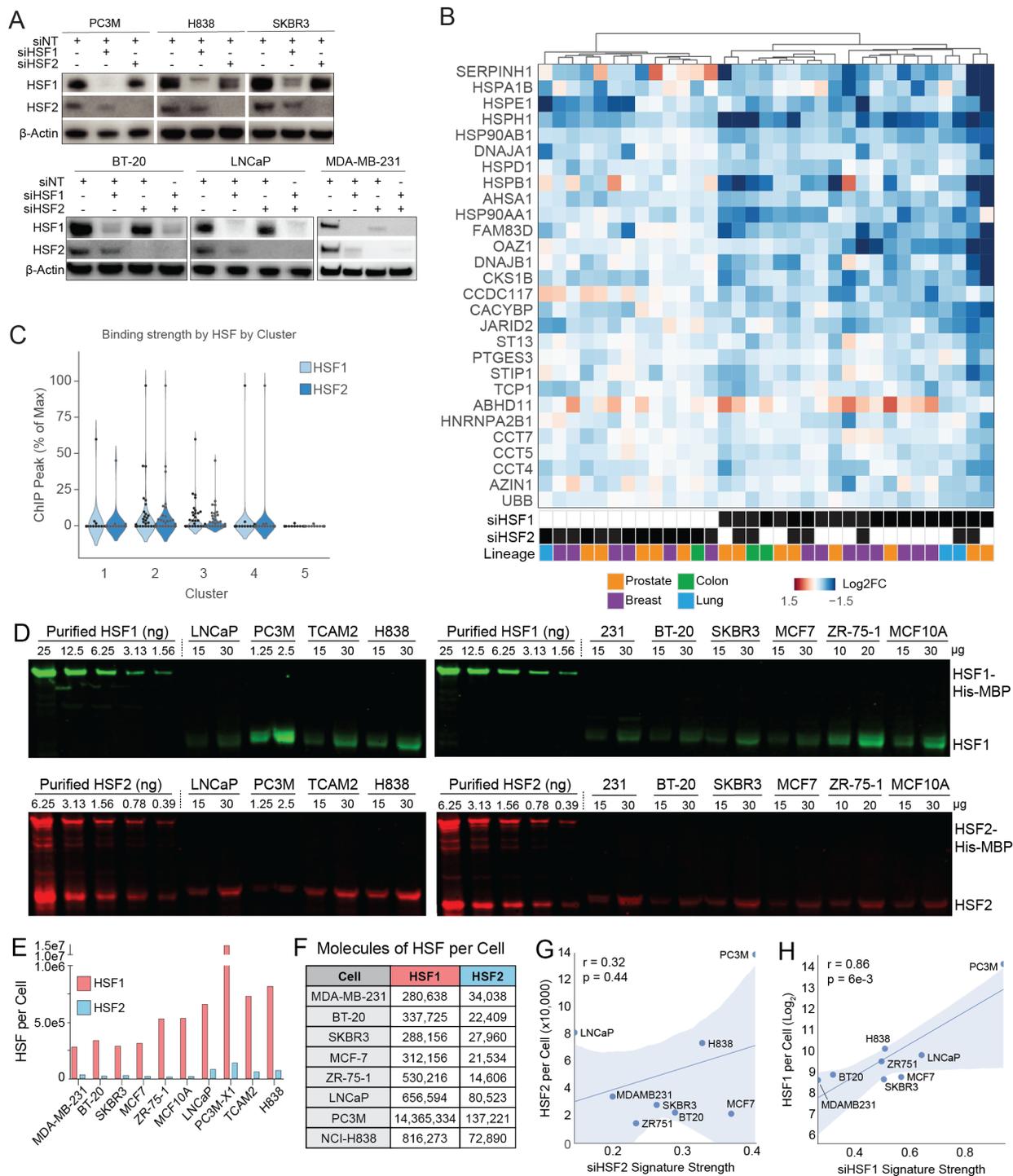


Figure 8. Validation of siRNA knockdown, HSF signature, and HSF quantitation.

(A) Western blot for HSF1, HSF2, or beta-actin in select cell lines treated with siRNA targeting either HSF1, HSF2, or both. (B) Same data as in Figure 7A, but only showing genes identified as bound by both HSF1 and HSF2 by ChIP-sequencing from Figure 5. Cell lines are indicated below heatmap with siRNA and cell line lineage indicated by boxes of shading and colored key, respectively. (C) Relative ChIP signal strength for genes in each of the major clusters in Figure 7A (C1 – C5, at right of heatmap), expressed as a percentage of the largest peak for each antibody for ease of comparison across antibodies. (D) We generated a standard curve of Purified His-MBP-tagged HSF1 or HSF2 for interpolating absolute HSF levels in each cell line given a known amount of protein loaded from a known number of cells. Absolute quantitation of HSF1 and HSF2 molecules per cell for each cancer line reveals markedly higher levels of HSF1 than HSF2 for all cell lines. PC3M, H838, LNCaP are the three lines with the highest levels of HSF2, followed by TCAM2 (a seminoma), included as a positive control known to express high levels of HSF2 protein. (E) Molecules of HSF1 (red) or HSF2 (blue) in each cell line, calculated from immunoblot in (D) and quantified in (F). (G) Correlation of HSF2 per cell with HSF signature strength for siHSF2 treated cell lines shown. (H) Correlation of HSF1 per cell with HSF signature strength for siHSF2 treated cell lines shown.

While HSF2 and HSF1 drive expression of similar genes, the magnitude of changes observed with HSF2 depletion was generally less than that of HSF1 depletion for many genes, which could be influenced by differences in chromatin binding or protein levels. We did not find differences in relative binding of these genes between HSF1 and HSF2 in any cluster that would explain this difference (Figure 8C). To better quantify the effect of each HSF on gene regulation, we used the absolute value of \log_2 -normalized fold-change data for “HSF cancer signature”, defined as ChIP-bound genes that are differentially expressed across cancer cell lines. We next assessed how the expression levels of each HSF paralog in a cell might affect the strength of the “HSF cancer signature” by quantifying total HSF1 and HSF2 in cell lysates using an immunoblot with purified protein standards (Figure 8D-F). Indeed, the total protein level of HSF1 is far greater than that of HSF2 in each cell line. However, the protein expression levels of HSF2 did not

significantly correlate with the strength of the siHSF2 gene expression signature ($r = 0.3$, $p = 0.48$, Figure 8G), in contrast to HSF1 protein expression levels (Figure 8H).

Another possible explanation for the stronger effect of HSF1 depletion on transcription in cancer cells is that HSF2 protein levels were also reduced after HSF1 depletion (Figure 8A), making HSF1- depleted cells functionally closer to an HSF1/HSF2 double knockdown. This idea is also supported by our observation that HSF1 depletion reduced global HSF2 chromatin occupancy (Figure 5D, 5I). To address this possibility, we simultaneously depleted HSF1 and HSF2 in seven cell lines of our original panel. In each cell line tested, HSF1/HSF2 double knockdown tightly clusters with that cell line's HSF1 knockdown sample (Figure 7A). Comparing the strength of the HSF cancer signature for single and double knockdown samples revealed an additive effect of depleting HSF2 in combination with HSF1 depletion (double knockdown) in 6 of the 8 cell lines and an epistatic relationship in LNCaP and PC3M (Figure 7C, statistics in 9B and 9D). Some genes exemplifying this pattern of additivity included *HSPH1*, *HSP90AA1*, *CKS1B* and *JARID2* (Figure 9A and 9C). This result suggests HSF1 and HSF2 cooperate for maximal target gene expression.

Considering HSF2 and HSF1 bind and regulate target genes similarly, we asked whether expressing each HSF paralog could serve to rescue gene expression in the reciprocal knockout. We first engineered MCF7 cells, to express a control GFP-FLAG, HSF1-FLAG, or HSF2-FLAG construct. Next, we treated each cell line with siRNA targeting HSF2, HSF1, or a non-targeting control (Figure 10A) and performed RNA-sequencing. We examined the high confidence HSF bound and transcriptionally regulated genes defined as the HSF cancer signature (Figure 8C) and found HSF1 overexpression failed to rescue all but a few genes (Figure 10B). These genes included

AZIN1, *DNAJB1*, and *HSPH1* which HSF1 overexpression modestly induced in HSF2 knockout cells, but they were not induced in the HSF2 rescue of HSF1-depleted cells. These data suggest that neither HSF can fully function without an intact HSF1:HSF2 complex. Taken together, these experiments reveal the intertwined roles of HSF1 and HSF2 in cancer, where formation of an HSF complex drives concordant, direct regulation of many genes critical to the cancer gene expression programs regulated by HSF1 and HSF2.

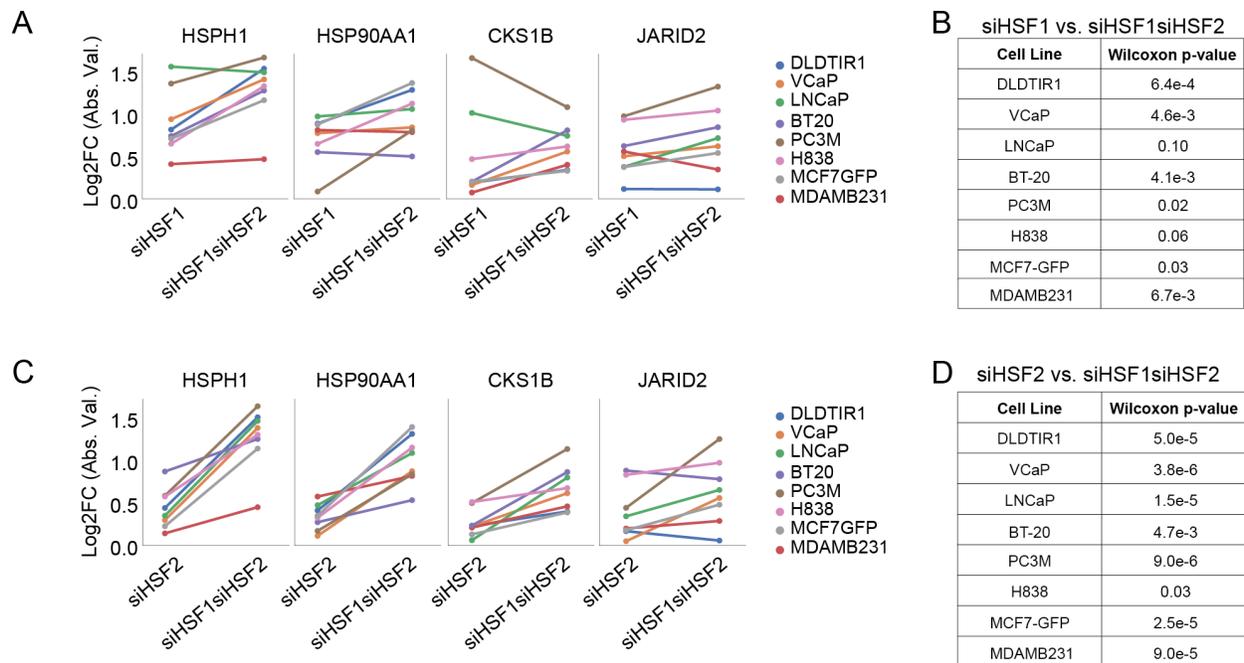


Figure 9. Further characterization of dual siRNA treatment and additivity of HSF1 and HSF2 gene regulation.

(A) and (C) Log₂FC signature value for example genes in the HSF signature between siHSF1 or siHSF2 and double knockdown (siHSF1siHSF2), respectively. (B) and (D) Results of Wilcoxon signed-rank test for each cell line comparing siHSF1 or siHSF2 to double knockdown (siHSF1siHSF2), respectively.

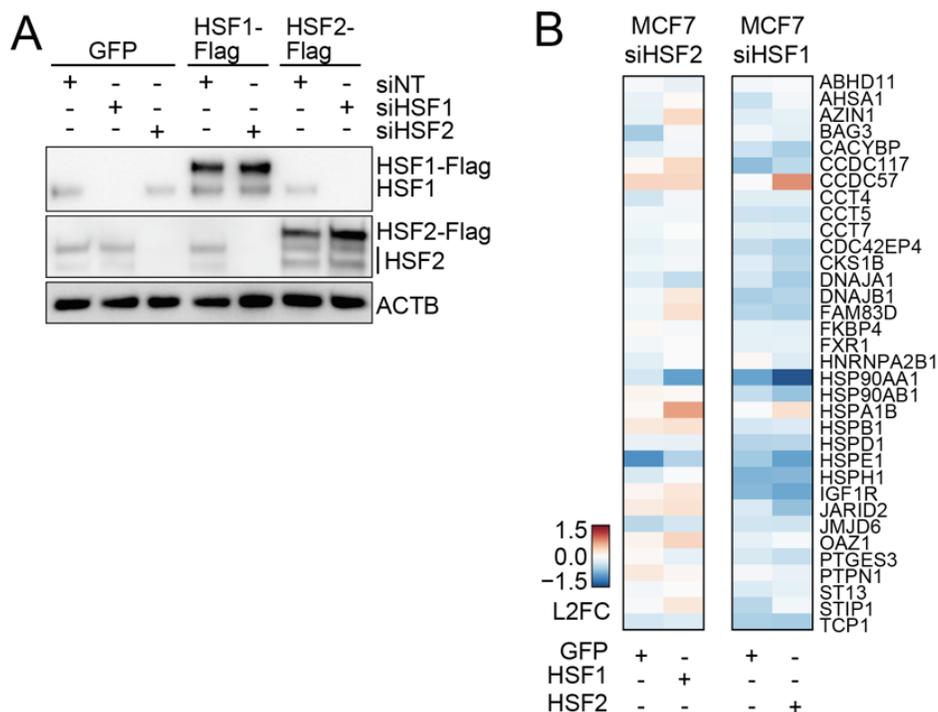


Figure 10. Overexpressing the reciprocal HSF does not prevent effects of short-term depletion.

MCF7 cells were engineered to overexpress either GFP-Flag, HSF1-Flag, or HSF2-Flag and treated with siRNA targeting the reciprocal HSF. **(A)** Western blot validation of siRNA knockdown and HSF-Flag overexpression. **(B)** Log₂ fold change values relative to the vector-specific, non-targeting siRNA are displayed. Genes included are those defined in the HSF cancer signature. Overexpression of HSF1 does not compensate for gene expression changes observed with siHSF2 (heatmap column 1 versus 2). Similarly, exogenous HSF2 expression does not alter siHSF1 regulated genes (heatmap column 3 versus 4).

2.4 Long-term loss of HSF2 and HSF1 results in sustained suppression of proteostasis gene expression conserved across cancers

While investigating the effects of short-term HSF loss provides insight into their role in gene regulation, it was unknown whether compensatory mechanisms could reestablish equilibrium in a cancer cell lacking HSF2 or HSF1 for a longer period. Considering the similarities of these HSF paralogs in regulating cancer cell gene expression and their strong correlation in effects on cancer cell line growth over a longer duration in CRISPR screening datasets (Fig 1D), we hypothesized that long-term loss of HSF1 and HSF2 would result in concordant effects on gene expression in a manner similar to short-term depletion. To test this hypothesis, we transduced six cancer cell lines with an sgRNA targeting HSF2 or HSF1 for CRISPR-Cas9 mediated knockout or a non-targeting sgRNA control and performed RNA-seq (Figure 11A). To preserve heterogeneity within cell lines, we generated population knockouts where possible (population knockouts failed for PC3M, so we engineered and selected two independent control and knockout clones). While acute (3 days) depletion of HSF1 or HSF2 by siRNA resulted in broadly correlated gene expression profiles such that neither siRNA nor cell line dominated hierarchical clustering (Figure 11B, left), long-term (> 14 days) HSF loss by CRISPR resulted in distinct cell-type-specific gene expression changes (Figure 11B, right). That is, the cell line background was the dominant variable between clusters, rather than which HSF was depleted. This suggests that loss of either HSF provokes an adaptive change which is dependent on the cellular context.

Despite the strong lineage-specific adaptation, we were able to identify a core module of genes with persistent reduced expression following from HSF knockout across cell lines (Figure 11C). Knockout of HSF2 resulted in 17 down-regulated and 9 up-regulated genes. HSF1 knockout resulted in 78 down-regulated and 30 up-regulated genes for a total of 186 unique genes (adjusted

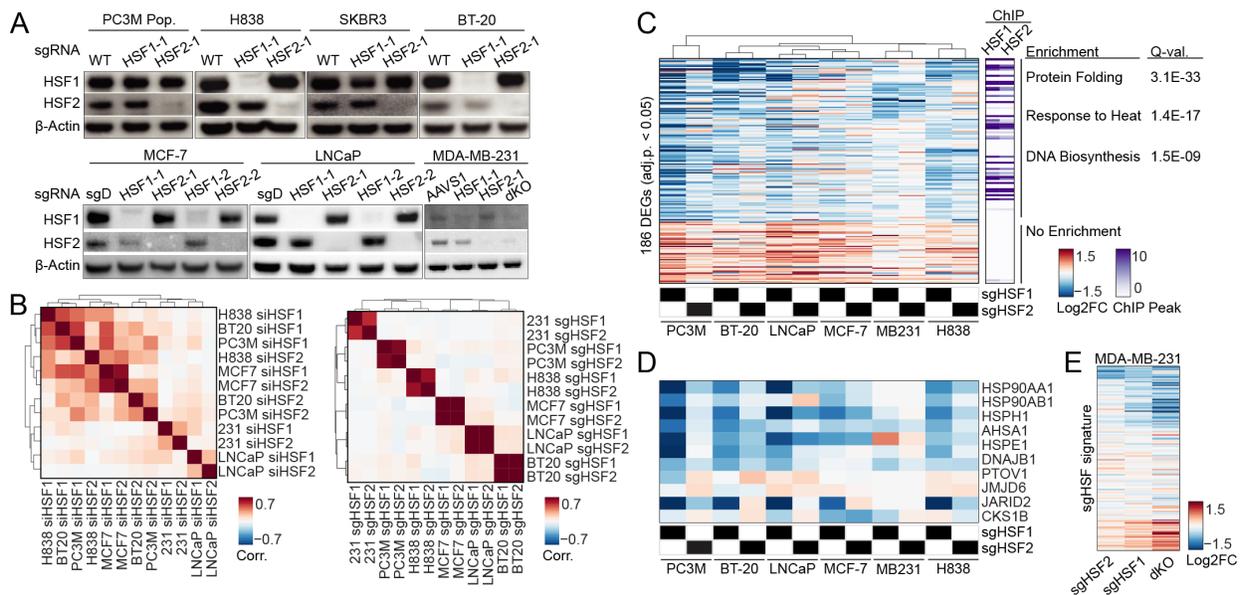


Figure 11. Long-term loss of HSF2 and HSF1 results in sustained suppression of proteostasis gene expression conserved across cancers.

(A) Knockout of HSF1 or HSF2 was confirmed by immunoblot. (B) Correlation plots of L2FC values for the union of differentially expressed genes for each siRNA-treated sample and each sgRNA knockout sample. 231 = MDA-MB-231 (C) Differentially expressed genes in either sgHSF1 or sgHSF2 knockout cells for the indicated cell lines. Relative ChIP-seq binding intensity is plotted to the right of the heatmap for both HSF1 and HSF2. Gene set enrichment analysis was performed and select enriched GO terms are displayed along with their FDR-adjusted p-value (Q-val.). (D) Select target genes from (A) are highlighted. (E) Heatmap of single and double knockout (dKO) MDA-MB-231 cells, plotting genes from (C).

p-value < 0.05 for either HSF1- or HSF2 knockout) (Figure 11A). K-means clustering of genes revealed patterns of down regulated genes upon knockout enriched in biological processes including protein folding (Q=5E-25), DNA biosynthesis (Q=3.3E-11), and stress response (7.8E-09). Notably, HSFs directly bind many of these genes in ChIP-seq experiments, especially genes down-regulated with HSF loss (Figure 11C, right heatmap). These genes included many canonical HSPs (e.g., *HSP90s*, *HSPH1*, *AHSA1*) but also included non-HSPs (e.g., *JMJD6*, *PTOVI*) (Figure 11D). Interestingly, long-term HSF loss did not substantially affect the proliferation gene

expression (e.g., *CKS1B*), as observed with short-term depletion (Figure 7). To assess the extent to which depletion of both HSF1 and HSF2 affects gene expression in CRISPR-edited cells, MDA-MB-231 double knockout (dKO) cells were engineered, and gene expression was measured (4E). Similar to the many cell lines treated with siRNA targeting both HSF1 and HSF2, MDA-MB-231 dKO

cells exhibited stronger down regulation of genes in the HSF cancer signature than depletion of one HSF or the other. These results are consistent with cooperative activity of these factors as observed in short-term siRNA experiments (Figure 7).

In addition to the effects on proteostasis gene expression, many cell-type specific changes were observed with HSF loss, suggesting that these transcription factors may directly or indirectly regulate cell state more broadly (Figure 12). These data demonstrate that there are many more differentially expressed genes for each cell line, extending beyond the consensus HSF signature defined in Figures 3 and 4, highlighting the long-term consequences HSF loss can have on tumor cell biology. These results demonstrate that loss of either HSF has durable consequences on gene expression programs with concordant decreases in proteostasis gene expression programs, and further supports the inability of either factor to fully compensate for loss of the other HSF.

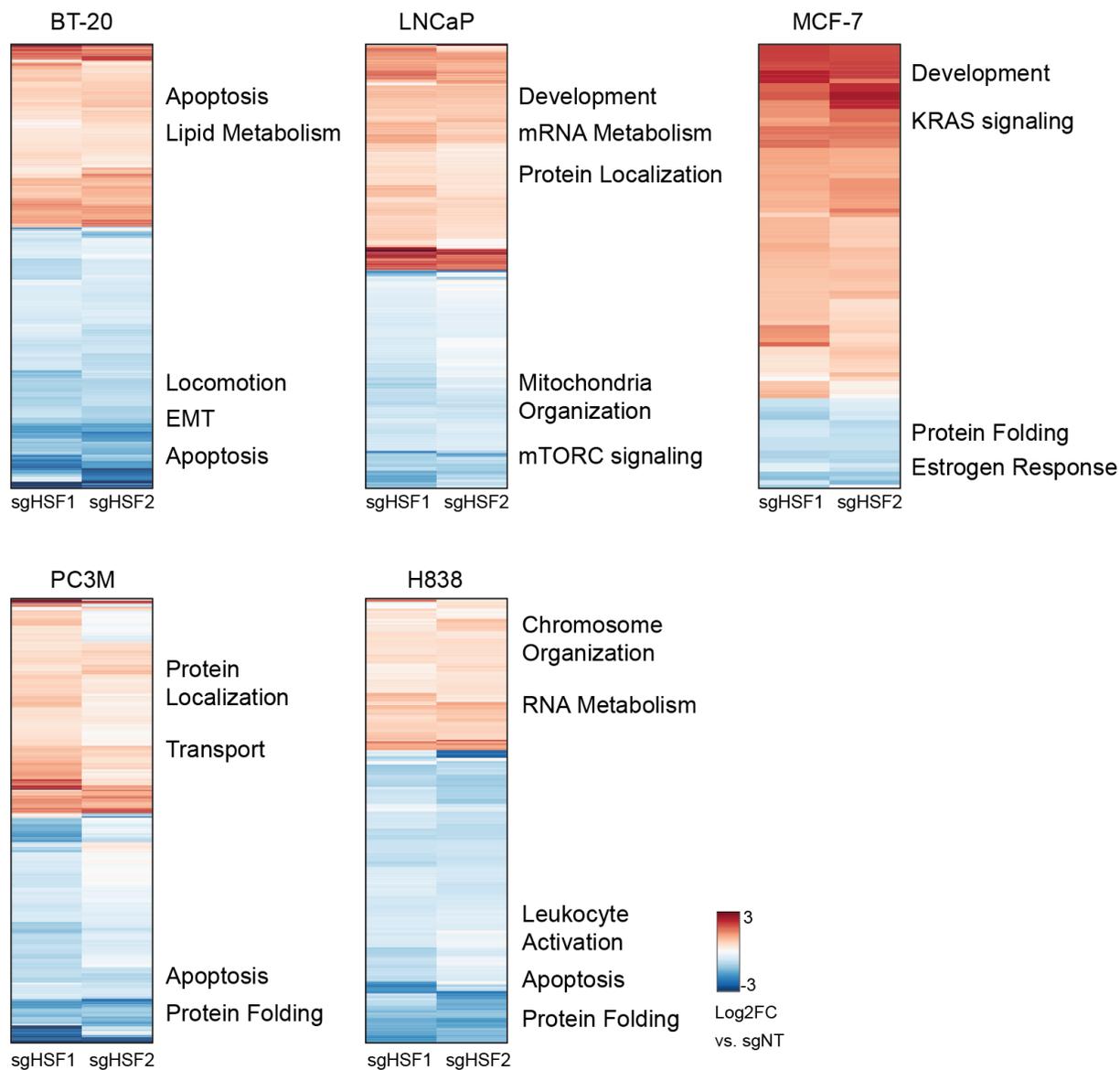


Figure 12. Long-term loss of HSF1 or HSF2 results in cell-type specific adaptations beyond the consensus HSF cancer signature.

Each heatmap represents log₂ fold change gene expression relative to the sgNT control for each cell line, respectively. All genes with an EdgeR adjusted p-value less 0.05 are plotted. Enriched GO terms are displayed next to heatmaps.

2.5 HSF2 and HSF1 promote the transcriptional response to cancer-associated stress perturbations

Previous studies using HSF2 knockout MEFs did not identify a significant role for HSF2 in the global response to thermal stress (Mahat et al., 2016; Solis et al., 2016). Because we revealed increased activity for HSF2 in a subset of highly malignant cancer cell lines compared to non-tumorigenic MEFs absent exogenous stress, we wondered whether the HSR in these cell lines might also be more dependent on HSF2. To test this idea, we performed RNA-seq in our HSF2 knockout MDA-MB-231 cell lines grown either at 37° C or following a 42°C heat shock along with wild-type and HSF1 knockout control lines. As expected, we found that even in these cancer cells with high levels of HSF2 basal activity, HSF2 loss did not greatly alter the HSF1-dependent

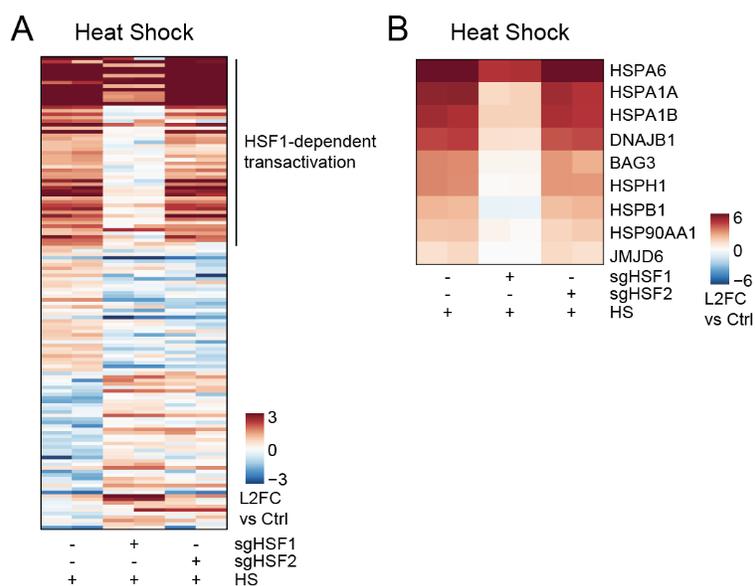


Figure 13. HSF2 knockout cancer cells have an intact heat shock response.

(A) Gene expression changes of heat shock (HS) treated (42°C, 1 hour) MDA-MB-231 cells with either HSF1 (sgHSF1) or HSF2 (sgHSF2) knockout, expressed as log₂ fold-change (L2FC) relative to non-targeting sgRNA controls without heat shock. **(B)** A subset of core heat shock response genes highlights from data in (A) (Note scale of L2FC).

HSR (Figure 13A-B). Thus, the increased HSF2 activity and concordant gene expression regulation by HSF2 and HSF1 in cancer highlight a role for HSF2 that is distinct from its dispensable nature in the HSR.

Cancer cells are not only subject to the stresses associated with rapid proliferation, but also the stresses of the tumor microenvironment such as limited nutrient and oxygen availability. Because previous work established a role for HSF1 in nutrient and oxygen stress responses (Bierkamp et al., 2010; Jin et al., 2011; Qiao et al., 2017; Su et al., 2019; Zou et al., 2003), we tested whether HSF2 might also have a role in these other malignancy-associated cellular stresses. To do so, we subjected control MDA-MB-231 cells or cells lacking HSF2, HSF1, or both to glycolytic stress with 2-deoxy-D-glucose (2-DG), serum starvation (SS), or the hypoxia mimetic cobalt chloride (CoCl₂) (Munoz-Sanchez and Chanez-Cardenas, 2019) and performed RNA-seq. In stark contrast to cells subjected to thermal stress, loss of either or both HSFs resulted in similar and broadly dysregulated transcriptional responses to these stresses (Figure 14A, C, E). K-means clustering in each condition revealed six major modules of gene expression regulation patterns: genes down-regulated by stress and either up-regulated (1) or further down-regulated by HSF loss (2), genes little changed in stress but either up-regulated (3) or down-regulated (4) with HSF loss, genes increased with stress or HSF1 loss, but decreased with HSF2 loss (5), and genes up-regulated by stress that are further increased with HSF loss (6) (Figure 14B, D, F).

This transcriptional dysregulation involved genes implicated in cell development and differentiation, adhesion, signaling, oxygen response and proliferation, and included many direct target genes. Select genes highlight distinct patterns of dysregulation (Figures 5B, 5D, 5F). For example, both nutrient stresses induce *AHS1*, a co-chaperone critical for HSP90 function, but

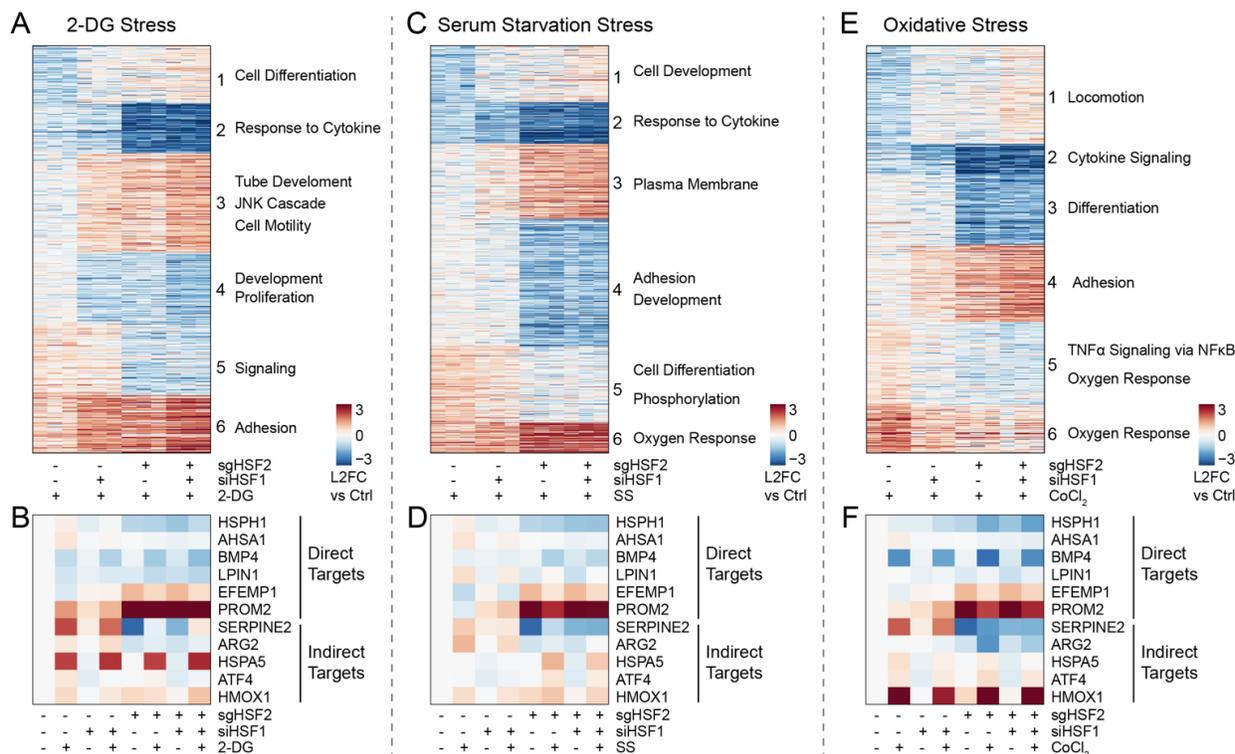


Figure 14. HSF2 and HSF1 promote the transcriptional response to cancer-associated stresses.

We treated HSF2 knockout (sgHSF2) or control (sgNT) MDA-MB-231 cells with siRNA targeting HSF1 (siHSF1) or non-targeting (siNT). Cells were treated with either 10 mM 2-deoxy-D-glucose (2-DG) for 24 hours, serum starvation (SS) for 48 hours, or 250 μ M cobalt chloride (CoCl₂). **(A, C, E)** Differentially expressed genes were determined relative to unstressed sgNT+siNT cells (Ctrl). Enriched GSEA terms are indicated. **(B, D, F)** Select genes from A – C are plotted and labeled as direct or indirect transcriptional targets based on ChIP-seq from Figure 5.

cells that lack HSF1 induce *AHSA1* to a lesser extent and HSF2 knockout cells are unable to induce *AHSA1* altogether. Similar results are observed for *HSPH1*. The expression of another class of genes is induced to a similar degree when stress is applied but is already suppressed in the absence of HSF1 or HSF2 in control conditions. For example, despite induction with serum starvation for each genetic background, *LPIN1*, a gene critical for triglyceride synthesis and implicated in the

pathogenesis of lipodystrophies (Peterfy et al., 2001), fails to increase in expression beyond basal levels in cells lacking HSF1 and HSF2. Some genes, such as *BMP4*, have lower basal levels of expression in the absence of HSF1 or HSF2 and decrease further with the addition of stress. Notably, HSF2 strongly represses a subset of directly bound genes including *EFEMP1* and *PROM2* (i.e., expression increased with loss). Still many other genes are dysregulated indirectly (not bound in ChIP-seq, e.g., *SERPINE2*, *ARG2*) demonstrating a broader transcriptional rewiring that occurs because of long term HSF2 loss and stress. These consequences of HSF2 loss and its similarity to HSF1 loss under conditions of impaired glycolysis, low serum, or oxygen deprivation highlight that while HSF2 and HSF1 have clearly divergent functions in response to thermal stress, they share a similar and pervasive role in regulating gene expression in cells subjected to these metabolic stresses.

2.6 HSF2 is required for tumor progression in cell line xenografts of prostate and breast cancer

Our results demonstrate that HSF2 interacts with HSF1 to bind and regulate a program of genes that support the anabolic state of many cancers, while also promoting the response to stresses that characterize the tumor microenvironment. However, how HSF2's function in regulating cancer cell transcription *in vitro* relates to tumorigenesis *in vivo* is not well understood. To address this question, we first injected our PC3M clonal knockouts or MDA-MB-231 population knockouts of HSF2 or HSF1 subcutaneously into immunocompromised (NOD-*scid* IL2Rg^{null}, NSG) mice and measured tumor volume over time and mouse survival. Knockout of HSF2 in the

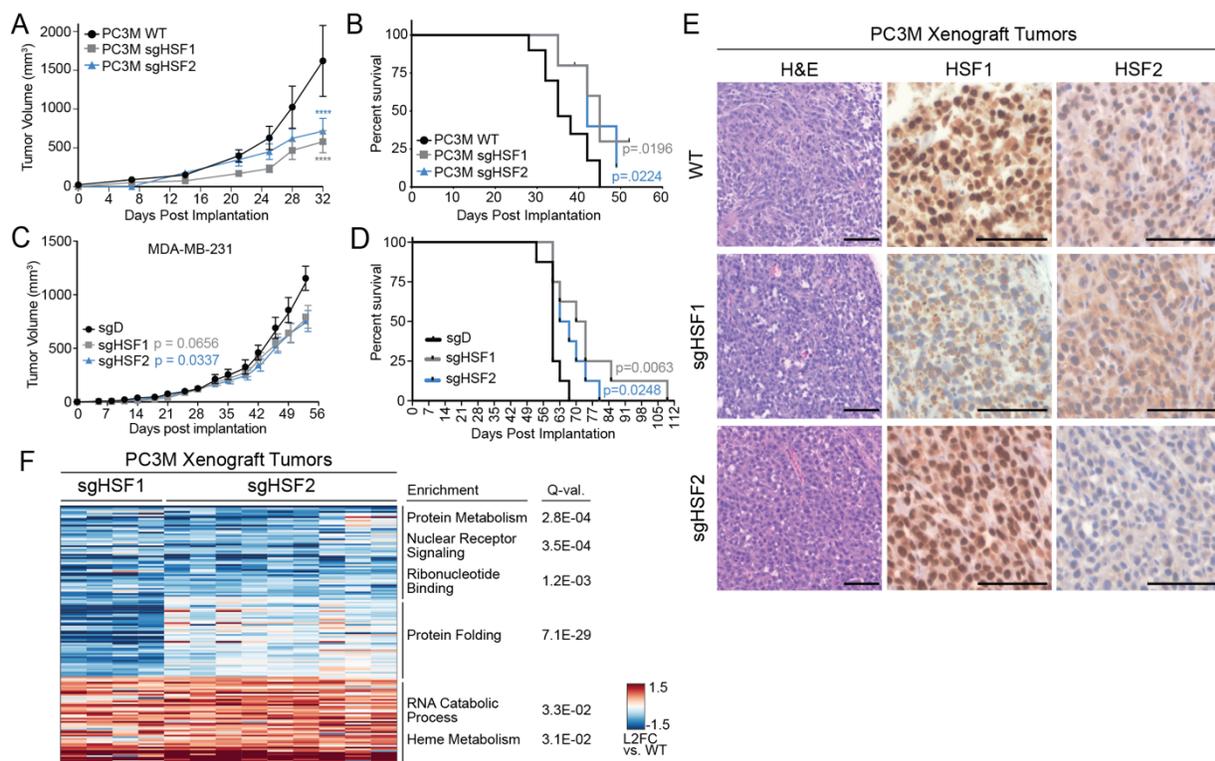


Figure 15. HSF2 is required for tumor progression in cell line xenografts of prostate and breast cancer.

(A) Tumor volume after subcutaneous injection of two independent PC3M cell clones for each of wild type (WT), HSF1 knockout (sgHSF1), or HSF2 knockout (sgHSF2). Data are grouped by genotype, combined across clones. **** = $p < 0.0001$ by Two-way ANOVA at day 32 post inoculation. (B) Survival analysis of mice bearing PC3M tumors. (C) Tumor volume measurements for MDA-MB-231 xenografts. (D) Survival analysis for of mice bearing MDA-MB-231 tumors. (E) Representative H&E and IHC staining for HSF1 or HSF2 for PC3M cell line xenograft tumor. Scale bars indicate 100 μm . Immunoblot validation of knockout cells prior to injection are provided in Figure 6J. (F) RNA sequencing performed on tumors from each knockout or control at experiment end point. Data are expressed as \log_2 fold-change relative to 4 control tumors. Enriched GO terms are indicated.

PC3M cell line resulted in a striking reduction in tumor growth compared to control (Figure 15A), and these effects were indistinguishable from those resulting from HSF1 knockout. The reduced tumor growth corresponded to prolonged mouse survival (Figure 15B). We obtained similar, although blunted, effects in the MDA-MB-231 cell line (Figure 15C, D).

To better understand these results, we excised and characterized the resultant tumors by immunohistochemistry (IHC) and RNA-Seq. Notably, IHC staining for HSF1 in MDA-MB-231 tumors revealed heterogeneous HSF1-positive staining (Figure 16A) that reflects a minority subpopulation of HSF1-expressing cells that had grown out from the initial knockout population confirmed by immunoblot (Figure 6F). This result suggests the observed tumor growth reduction, while not statistically significant, is underestimated. Lower levels of HSF2 expression in MDA-MB-231 tumors precluded assessment by IHC (Figure 16A). Consistent with outgrowth of a wild-type subpopulation of cells, very few genes were differentially expressed between control and knockout tumors (Figure 16B), which contrasts with the many gene expression differences of these lines detected *in vitro* (Figure 16C). Persistent knockout of HSF1 or HSF2 was confirmed for tumors derived from PC3M clonal knockouts by IHC (Figure 15E) and additionally revealed reduced HSF2 staining for HSF1 knockout tumors. Consistent with our *in vitro* data, RNA-seq of HSF2 knockout tumors had similar gene expression to HSF1 knockout tumors (Figure 15F) with decreased expression of protein folding ($Q = 7.1E-29$) and protein metabolism ($Q = 2.8E-04$) genes in common for HSF1- and HSF2-depleted tumors. While these effects have been reported for HSF1, our data establish a significant role for HSF2 in regulating these processes in cancer. Comparing the *in vitro* and *in vivo* RNA-seq for PC3M reveals significant overlap despite expected differences due to the complex contribution of the tumor microenvironment to cancer cell gene

expression. Enriched pathways for sgHSF1 comparisons include HSF1 activation/attenuation ($Q = 1.73e-15$) and regulation of protein metabolic processes ($Q = 1.65e-14$). The overlap for sgHSF2 is enriched for biological adhesion ($Q = 2.3e-3$) and protein localization ($Q = 1.42e-2$) (16D). Altogether, these data establish a critical role for HSF2 in directly regulating cancer gene expression programs, response to metabolic stresses, and tumor progression in cooperation with HSF1.

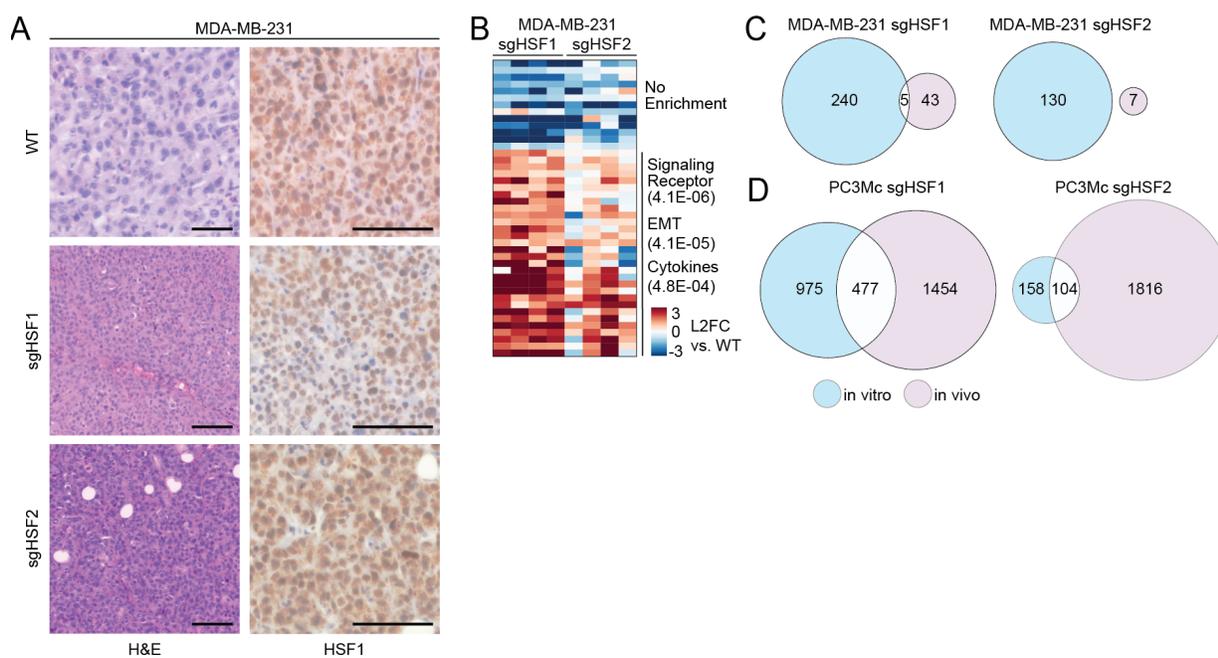


Figure 16. Further characterization of MDA-MB-231 xenograft tumors and RNA-seq.

IHC staining of MDA-MB-231 tumors for HSF1. Scale bar indicates 100 μ m. **(B)** Differentially expressed genes (DEGs) of tumors compared to sgNT controls. **(C, D)** Comparison of DEGs between in vitro and in vivo RNAseq for each indicated cell line and sgRNA.

CHAPTER 3: DISCUSSION

A large body of work has clearly established the importance of HSF1 activation across a diverse spectrum of cancers. HSF1 supports the malignant state by directing a transcriptional program of genes involved in many facets of tumorigenesis, extending far beyond protein folding and stress responses to include cell cycle, energy metabolism and other proliferation-associated processes (Mendillo et al., 2012). Indeed, high levels of HSF1 activation and its transcriptional activity correlate with metastasis and death in patients (Mendillo et al., 2012). Here, we establish that HSF2, like its paralog HSF1, is a significant contributor in supporting cancer cell gene expression programs.

Our data identify a prominent role for HSF2 in cancer where it cooperates with HSF1 in a manner that starkly contrasts with their divergent roles in the response to thermal stress. HSF1 has been well-studied for its profound induction of chaperones upon acute heat stress and thus has served as a model for inducible gene expression, yielding significant insight into transcription regulation (Guertin and Lis, 2010; Teves and Henikoff, 2013; Vihervaara et al., 2018; Vihervaara et al., 2017). In part because HSF2 does not play a significant role in promoting the response to heat stress (Gomez-Pastor et al., 2018; Korfanty et al., 2014; Mahat et al., 2016; Solis et al., 2016), its role in transcription regulation – and biology in general – has been far less explored. For example, while HSF1 has an extensive body of literature connecting its activity to diseases that include cancer (Alasady and Mendillo, 2020) and neurodegenerative diseases (Neef et al., 2011), among others (Gomez-Pastor et al., 2018), mechanistic studies of HSF2 biology are largely limited to its role in development and response to mild proteotoxic stresses (Akerfelt et al., 2010a). Here, we demonstrate that HSF2, like HSF1, promotes the expression of HSPs as well as non-HSP transcriptional targets to support the malignant state. Notably, these HSPs do not include those

most inducible in response to elevated temperature, supporting the idea that studying HSF2 function at loci beyond the highly heat shock-inducible *HSP70*-family and *HSP40*-family genes will be critical to reveal a comprehensive picture of its role in regulating gene transcription.

Our study suggests that an important aspect of HSF1 function in cancer is to promote HSF2 abundance and activity, and phenotypes previously attributed solely to HSF1 may reflect the effect of loss of both HSFs. Thus, the effect of HSF1 loss on reducing HSF2 expression and chromatin occupancy may contribute to its more pronounced suppression of cell cycle and proliferation-associated gene expression. On the other hand, it is possible that HSF2 loss alone falls below a threshold of cytotoxicity that fails to result in severe growth arrest because of the intact function of HSF1, explaining the greater absolute fitness cost of HSF1 loss versus HSF2 loss in large-scale genetic dependency screens (Meyers et al., 2017). Because HSF1 has critical roles in many other aspects of biology (Akerfelt et al., 2010a; Alasady and Mendillo, 2020), it will be important to understand the cooperation and interplay of HSF1 and HSF2 across diverse biological contexts.

Why might cancer cells engage both HSF1 and HSF2? Consistent with structural studies that have revealed highly conserved DNA-binding domain conformations (Jaeger et al., 2016), our data demonstrate a nearly identical chromatin occupancy pattern of HSF1 and HSF2, suggesting their specialization does not lie in DNA-binding activity as has been described for other paralog transcription factors (Rogers and Bulyk, 2018). However, their divergence in surfaces that mediate protein-protein interactions could result in differential recruitment of regulatory cofactors. Considering HSF2 is a labile protein sensitive to rates of protein synthesis and degradation (Mathew et al., 1998; Santopolo et al., 2021), another possibility is that it provides cancer cells an additional mechanism to meet the anabolic demands of proliferation and increased protein

synthesis through the HSF cancer program. Indeed, proteins with a shorter half-life allow cells to rapidly adjust expression levels, and this regulatory feature can be observed in other proteins associated with proliferation (e.g. Myc). On the other hand, cells maintain higher HSF1 levels, with a much longer half-life, to appropriately respond to sudden and acute proteotoxic stresses in addition to its cancer supporting functions. Other mechanisms, yet undiscovered, may also stabilize HSF2 during formation and maintenance of tumors. Lastly, previous observations that HSF2 has greater chromatin occupancy in mitotic cells than HSF1 (Vihervaara et al., 2013) suggest cell-cycle dependent differences between these paralogs ensure maintenance of these transcriptional programs in highly proliferative cells.

Our study also sheds light on the biological roots of the roles HSFs play in cancer. Previous work has revealed a striking similarity between HSF1's transcriptional program in cancer and that of the single HSF-1 in *C. elegans* during larval development (Li et al., 2016). Both cancer and larval development are characterized by biomass accumulation and cell proliferation generating increased demand on protein synthesis and quality control machinery. Each of these highly proliferative settings reveals an expansion of HSF1 transcriptional activity beyond its canonical HSR target genes to include anabolic gene expression programs. Considering the intimate link our study reveals between HSF2 and HSF1 in cancer and HSF2's more prominent role in development and mitotic gene expression regulation during mitosis (Vihervaara et al., 2013; Xing et al., 2005), HSF2 serves as a more direct connection between these developmental and proliferative gene expression programs. Thus, cancer cells seem to co-opt the ancient, developmental HSF program (Li et al., 2017) to support the malignant state.

In conclusion, our study has documented a significant role for HSF2 in supporting malignancy. In contrast with heat shock, cancer-relevant stresses such as nutrient deprivation or *in vivo* tumorigenesis invoke an HSF2-driven gene expression program which parallels that driven by HSF1. Thus, HSF2 is a critical HSF1 accomplice, promoting a gene expression program which supports the anabolic malignant state and fuels cancer progression.

CHAPTER 4: MATERIALS AND METHODS

4.1 Multiple Sequence Alignment

Amino acid sequences were obtained from Ensembl.org for the main functional isoform for each of HSF1 (ENSP00000431512), HSF2 (ENSP00000357440), HSF4 (ENSP00000430947), and HSF5 (ENSP00000313243) under genome build GRCh38.p13 (Cunningham et al., 2022). Protein sequence alignment and similarity were determined using the Clustal Omega Multiple Sequence Alignment tool, through EMBL-EBI: <https://www.ebi.ac.uk/Tools/msa/clustalo/> (Madeira et al., 2022).

4.2 Developmental Gene Expression of HSFs

Cardoso-Moreira et al. 2019 performed RNA sequencing of cerebrum, cerebellum, heart, kidney, liver, ovary, and testis tissue across the developmental lifespan of human and other organisms (Cardoso-Moreira et al., 2019). RPKM gene expression data for human samples for each tissue and developmental timepoint were downloaded from <https://apps.kaessmannlab.org/evodevoapp/>. Replicates ranged from one to four for each tissue and timepoint. Data plotted in Figure 1 represent the log normalized, mean of replicate RPKM values + 1. Included transcript IDs for each HSF are as follows: HSF1 (ENSG00000185122), HSF2 (ENSG0000025156), HSF4 (ENSG00000102878), and HSF5 (ENSG00000176160).

4.3 Cell culture

Cell lines were obtained from American Type Culture Collection or a gift from S. Lindquist. PC3M, NCI-H838, SKBR3, BT-20, ZR-75-1, HCC38, LNCaP, VCaP, and DLD1 cells were maintained in RPMI 1640 (Gibco, #11875119), 10% fetal bovine serum (Clontech,

#631106), and 1% penicillin/streptomycin (pen/strep; Gibco, #15140122). MDA-MB-231, MCF7, MEFs, and 293T cells were maintained in Dulbecco's modified Eagle's medium (Gibco, #11995073), 10% fetal bovine serum, and 1% pen/strep. Cell lines were authenticated at the University of Arizona Genetics Core and tested negative for mycoplasma. Cells were lifted for passaging with Accumax (Innovative Cell Technologies, #AM105). Cells were maintained at 37°C and 5% CO₂ in a HeraCell Vios 160i incubator (Thermo Fisher Scientific).

4.4 Mass spectrometry (M/S) and analysis

Immunoprecipitation samples were prepared for LC-MS/MS with chloroform methanol precipitation (Wessel and Flugge, 1984). LC-MS/MS of immunoprecipitated GFP-Flag, HSF1-Flag, or HSF2-Flag was performed as described (Hickox et al., 2017). Unfiltered mass-spectrometry data is available for experiment (see Table S1). The following measurements and information are given for each protein identified: Accession number, Peptide Count, NSAF (normalized spectral abundance factor), emPAI (exponentially modified protein abundance index), Spectral Count, Percent Sequence Coverage, and Description of protein identified. We defined high confidence interacting proteins as those with at least three-fold more spectral counts in our experiments relative to the average of spectral counts for affinity purified mass spectra from publicly available IP-MS experiments, specifically those targeting Flag epitope and agarose beads (crapome.org) (Mellacheruvu et al., 2013).

4.5 LUMIER assay

The LUMIER assay was performed essentially as described previously (Taipale et al., 2014). Briefly, 3X FLAG-tagged bait proteins (Table S2) were transiently transfected in 96-well

format into a 293T cell line stably expressing HSF1 with a codon-optimized *Renilla reniformis* luciferase C-terminal tag. Two days after transfection, cells were rapidly washed in 1× ice-cold phosphate-buffered saline (PBS) and lysed in ice-cold HENG buffer (50 mM HEPES-KOH pH 7.9, 150 mM NaCl, 2 mM EDTA, 5% glycerol, 0.5% Triton X-100 supplemented with protease, and phosphatase inhibitors) (Taipale et al., 2012). The lysate was transferred into 384-well plates coated with monoclonal M2 antibody and blocked with 3% bovine serum albumin/5% sucrose/0.5% Tween 20. Plates were incubated at 4°C for 3 hours, after which plates were rapidly washed in HENG buffer. Last, luciferase-tagged HSF1 was detected by measuring luminescence using the Gaussia FLEX Luciferase Kit [New England Biolabs (NEB)]. An HSF1-LUMIER interaction score was determined from the log₂-transformed bait-FLAG/GFP-FLAG luminescence ratio (reported in Table S2).

4.6 Electrophoretic mobility shift assay (EMSA)

EMSAs were performed as described previously (Mosser et al., 1988) with modifications based on the LICOR Odyssey EMSA kit protocol to accommodate the infrared dye (IRDye) labeling and detection strategy. EMSA protein lysates were prepared either from NCI-H838 cells or from Hsf1^{+/+} or Hsf1^{-/-} MEFs treated with either dimethyl sulfoxide control or 10 μM MG-132. Briefly, cells were washed twice with ice-cold PBS, scraped, centrifuged, flash-frozen, and stored at -80°C. The frozen cell pellets were lysed in buffer containing 20 mM tris (pH 8.0), 25% glycerol, 420 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.5 mM dithiothreitol (DTT), and protease inhibitor cocktail. Samples were centrifuged for 5 min at 100,000g, and supernatants were flash-frozen and stored at -80°C. Protein concentrations were determined using a BCA protein assay (Pierce).

EMSA DNA substrates were prepared as follows. All oligonucleotides were resuspended in 1X TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA) to a final concentration of 20 pmol/ μ L. Then, 5 μ L of forward and reverse oligonucleotides were mixed in a single tube as and annealed by placing in a 100° C heat block for 5 minutes, followed by turning of the heat block and allowing it to slowly cool to room temperature. A working DNA substrate stock was prepared by diluting annealed oligos 1:200 in water. Substrates were prepared that were labeled with DY-782 or DY-672 (where both forward and reverse oligos were end-labeled with the same IRDye) or where unlabeled for use as a cold competitor.

EMSA binding reactions were performed using 8 μ g lysates, 5 nM IRDye labeled oligo (see **Table 1**), 100 nM unlabeled competitor oligo, and 100 nM antibody or control IGG, as indicated in binding buffer containing 10 mM Tris, pH 7.5, 1 mM EDTA, 100 mM NaCl, poly(deoxyinosinic-deoxycytidylic) (50 ng/ μ l), 0.5 mM DTT, and 0.25% Tween 20. Reactions were incubated for 20 min at room temperature after which 1 μ l of 10 \times Orange loading dye (LI-COR, P/N 927-1010 0) was added to each reaction, mixed, and loaded on a Mini-PROTEAN TBE precast gel (Bio-Rad). The gel was run at 10 V/cm for about 30 min in 1 \times Tris/Borate/EDTA (TBE) buffer. Following electrophoresis, gels were imaged directly on an LI-COR Odyssey imager.

Table 1. Oligo sequences for EMSA experiments.

EMSA Oligo Name	EMSA Oligo Sequence (5' to 3')
EMSA-HSPA8-F	CTTATACCCCTATCTTAGAACCTTCCAGAAGGGGCCCCGCC
EMSA-HSPA8-R	GGGCGGGCCCCTTCTGGAAGGTTCTAAGATAGGGTATAAG
EMSA-HSPA6-F	GGAAGGTGCGGGAAGGTGCGGAAAGGTTTCGCGAAAGTTCG
EMSA-HSPA6-R	CGAACTTTCGCGAACCTTTCGCGACCTTCCCGCACCTTCC
EMSA-HSPA8*-F (mut HSE)	CTTATACCCCTATCTTATAACCTTGCACAAGGGGCCCCGCC
EMSA-HSPA8*-R (mut HSE)	GGGCGGGCCCCTTGTGCAAGGTTATAAGATAGGGTATAAG

4.7 Immunoprecipitation

Cells were rinsed twice with cold PBS, removed from plates by scraping, and centrifuged for 4 min at 1000g and 4°C before pellet resuspension in cold lysis buffer. For IPs of HSF1 or HSF2, cells were lysed in buffer containing 1% NP-40, 100 mM NaCl, 50 mM tris (pH 7.5), 0.2 mM EDTA, 5% glycerol, and 1 mM phenylmethylsulfonyl fluoride (PMSF), and lysis was achieved by sonication in a 4°C water bath (10 cycles of 30-s on and 1-min off). After lysis, cells were spun at 21,000g at 4°C for 10 min, and the supernatant was kept for input and IP. FLAG-IPs were performed using M2 affinity agarose (Thermo Fisher Scientific). HSF1 IP was performed with a polyclonal rabbit anti-HSF1 antibody (Cell Signaling Technology, #4356S). HSF2 IP was performed with a monoclonal rat anti-HSF2 (3E2) antibody (Santa Cruz Biotechnology, sc-13517). As a control for nonspecific binding, normal mouse IgG was immunoprecipitated (Santa Cruz Biotechnology, sc-2027).

4.8 Immunoblot

Protein samples were lysed in radioimmunoprecipitation assay buffer [10 mM tris-Cl (pH 8.0), 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, and

140 mM NaCl] containing 1 mM PMSF and a Roche cOmplete Protease Inhibitor Cocktail tablet (catalog no. 11697498001) and passed through a 21-gauge needle. Protein concentration was determined with a BCA assay (Pierce, #23255), denatured in Laemmli sample buffer containing 1% β -mercaptoethanol, and heated at 95°C for 5 min. Electrophoresis used 4 to 20% bis-tris gradient gels unless otherwise specified, with transfer to polyvinylidene difluoride membranes using the iBlot 2 Dry Blotting System (Thermo Fisher Scientific). Membranes were blocked in 5% fat-free milk for 1 hour at room temperature. Primary and secondary antibodies were diluted in 5% fat free milk and exposed to membranes overnight at 4°C. Immunoblots were developed with Immobilon Western Chemiluminescent HRP Substrate (Millipore, #WBKLSO500) and visualized with a ChemiDoc Touch Imaging System (Bio-Rad, #732BR0783), and images were analyzed using ImageLab v6.0.1 (Bio-Rad). HSF1 immunoblot was performed with a polyclonal rabbit anti-HSF1 antibody (Cell Signaling Technology, #4356S). HSF2 immunoblot was performed with a monoclonal rat anti-HSF2 (3E2) antibody (Santa Cruz Biotechnology, sc-13517). Antibody specificity was confirmed with purified proteins. Equal loading of protein was confirmed with a β -actin monoclonal antibody (BA3R, Invitrogen, #MA5-15739).

4.9 Protein purification

Human HSF1 (NM_005526.4) and HSF2 (NM_004506.4) were cloned into the pET His6 MBP TEV LIC cloning vector that was a gift from S. Gradia (Addgene plasmid #29656) and verified by Sanger sequencing. BL21 DE3 (NEB) were transformed with either pMH-HSF1 or pMH-HSF2. A single colony was picked and incubated in 50 ml of LB broth in a shaking incubator overnight at 37°C. The next morning, the 50-ml cultures were diluted to 1 liter and incubated at 37°C until they were at an optical density (OD) of 0.5. At an OD of 0.5, protein

production was induced with 1 mM isopropyl- β -d-thiogalactopyranoside (final concentration), and the cultures were incubated at 18°C overnight. Bacteria were harvested, resuspended in 30 ml of heparin binding buffer [50 mM Hepes (pH 7.4), 50 mM NaCl, 0.1 mM EDTA, and 5% glycerol] supplemented with lysozyme (10 μ g/ml). Lysates were prepared by lysing the membranes with a sonicator (30% amplitude and 0.5-s on/off cycles for 2 min). The cellular debris was cleared by centrifugation (20,000g for 15 min at 4°C). The clear supernatant was passed through HiTrap heparin column (GE Healthcare) using ÄKTA high-performance LC system. The column was washed (10 column volumes) with heparin binding buffer and eluted over a linear salt gradient ranging from 50 to 1.5 M NaCl over 20 column volumes. Fractions from heparin elution were analyzed by SDS-polyacrylamide gel electrophoresis, followed by Coomassie staining. Appropriate fractions identified from this step were subject to Ni-NTA (nitrilotriacetic acid) (QIAGEN) purification as per the manufacturer's protocol. Briefly, heparin eluate was incubated with Ni-NTA resin for 2 hours on a shaker at 4°C, washed four times, and eluted with 250 mM imidazole. Ni-NTA eluate proteins were resolved by gel filtration using a Superdex 200 10/300 GL column on ÄKTA system (GE Healthcare). Purified proteins were quantified by Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific) and by the enhanced protocol for BCA Assay (Pierce, #23255). These purified recombinant proteins were used for confirming antibody specificity and quantitative immunoblot.

4.10 Quantitative immunoblot

For each cell line, 5e6 cells were plated in two 15cm tissue culture dishes. 24 hours later, 1 dish of cells was counted, and the second dish of cells was washed and scraped with cold PBS into a microfuge tube on ice. Cells were pelleted at 1000g for 5 minutes at 4°C. After discarding

supernatant, cell pellets were snap frozen in liquid nitrogen for 30-60 seconds and stored at -80°C. Subsequently, cell pellets were lysed as described above. Whole cell extracts were loaded at two concentrations for each cell line alongside a standard curve of known amounts of recombinant HSF1 or HSF2. iBlot 2 was used to transfer protein to a nitrocellulose membrane. Membranes were blocked in Odyssey Blocking Buffer (LI-COR). Membranes were incubated in primary antibody (same as in immunoblot) at 4°C overnight in Odyssey Blocking Buffer plus 0.2% Tween20 according to manufacturer protocol. Membranes were washed 3x5 minutes with PBST then incubated in secondary antibody diluted in Odyssey Blocking Buffer plus 0.2% Tween20 for 60 minutes at room temperature with LI-COR IRDye secondary antibodies at 1:5000 dilution: donkey anti-rabbit for detecting HSF1 (#926-32213), goat anti-rat for detecting HSF2 (#926-68076). Membranes were washed with PBST then rinsed in PBS to remove Tween20. Membranes were imaged using the Odyssey CLx imager using auto-exposure settings. Specific bands corresponding to the full length proteins were quantified with Image Studio. Recombinant protein was used to generate a standard curve of known protein amount to LI-COR signal generating a linear model used to interpolate total target protein amount for each cell line. Because total protein was harvested from a known number of cells, we were able to calculate the number of molecules of HSF per cell.

4.11 HSF overexpression constructs

Overexpression constructs of HSF1 or HSF2 were generated by Gateway cloning HSF1 or HSF2 coding sequence in pLenti6.2-ccdB-3xFLAG-V5 that was a gift from M. Taipale (Addgene plasmid #87071) (Taipale et al., 2014). MCF-7 cells were subsequently selected with blasticidin. Expression of HSF1 and HSF2 was confirmed by immunoblot for either HSF or Flag.

4.12 Gene silencing

siGenome or ON-TARGET plus siRNA pools containing 4 gene-specific siRNAs were obtained for HSF1 and HSF2 (Dharmacon, Horizon Discovery), as well as a non-targeting control siRNAs, and transfected using RNAimax (Thermo Fisher) according to the manufacturer protocol. Unless otherwise specified, cells were harvested 72h after siRNA transfection. For serum starvation, serum was removed 48 hours after siRNA treatment and harvested after 48 hours of serum starvation. Knockdown was confirmed for each siRNA experiment by qPCR/RNA-seq or immunoblot. See Table 2 for additional sample and treatment details.

Table 2. siRNA Sample Notes

Cell Line	siRNA Target	siRNA Pool Type	n	Naming Notes
HCC38	NT (non-targeting)	siGenome	2	
HCC38	HSF1	siGenome	2	
HCC38	HSF2	siGenome	2	
MCF7	NT (non-targeting)	siGenome	2	
MCF7	HSF1	siGenome	2	
MCF7	HSF2	siGenome	2	
SKBR3	NT (non-targeting)	siGenome	2	
SKBR3	HSF1	siGenome	2	
SKBR3	HSF2	siGenome	2	
BT20	NT (non-targeting)	siGenome	2	
BT20	HSF1	siGenome	2	
BT20	HSF2	siGenome	2	
BT20	HSF1+HSF2	siGenome	2	
PC3M (population)	NT (non-targeting)	siGenome	2	
PC3M (population)	HSF1	siGenome	2	
PC3M (population)	HSF2	siGenome	2	
PC3M (population)	HSF1+HSF2	siGenome	2	
PC3MX1 (c1 clone)	NT (non-targeting)	siGenome	2	
PC3MX1 (c1 clone)	HSF1	siGenome	2	

PC3MX1 (c1 clone)	HSF2	siGenome	2	
ZR751	NT (non-targeting)	siGenome	2	
ZR751	HSF1	siGenome	2	
ZR751	HSF2	siGenome	2	
NCI-H838	NT (non-targeting)	siGenome	2	
NCI-H838	HSF1	siGenome	2	
NCI-H838	HSF2	siGenome	2	
NCI-H838	HSF1+HSF2	siGenome	2	
DLD1	NT (non-targeting)	OnTargetPlus	3	
DLD1	HSF1	OnTargetPlus	3	
DLD1	HSF2	OnTargetPlus	3	
DLD1	HSF1+HSF2	OnTargetPlus	3	
VCaP	NT (non-targeting)	OnTargetPlus	3	
VCaP	HSF1	OnTargetPlus	3	
VCaP	HSF2	OnTargetPlus	3	
VCaP	HSF1+HSF2	OnTargetPlus	3	
LNCaP	NT (non-targeting)	OnTargetPlus	3	
LNCaP	HSF1	OnTargetPlus	3	
LNCaP	HSF2	OnTargetPlus	3	
LNCaP	HSF1+HSF2	OnTargetPlus	3	
PC3MX1 (c1 clone)	NT (non-targeting)	OnTargetPlus	2	"PC3MX" in Fig. 3
PC3MX1 (c1 clone)	HSF1	OnTargetPlus	2	"PC3MX" in Fig. 3
PC3MX1 (c1 clone)	HSF2	OnTargetPlus	2	"PC3MX" in Fig. 3
MDAMB231	NT (non-targeting)	OnTargetPlus	3	
MDAMB231	HSF1	OnTargetPlus	3	
MDAMB231	HSF2	OnTargetPlus	3	
MDAMB231	HSF1+HSF2	OnTargetPlus	3	

EdgeR analyses were grouped by siRNA pool type

4.13 CRISPR/Cas9 knockout generation

Unless otherwise noted, knockout cells were generated using CRISPR/Cas9 by introducing lentiCRISPRv2 with lentivirus (Sanjana et al., 2014; Shalem et al., 2014). lentiCRISPRv2 was a gift from Feng Zhang (Addgene plasmid #52961). Briefly, 4 sgRNAs sequences targeting HSF1 or HSF2 (see **Table 3**) were selected from the Brunello Human CRISPR Knockout Pooled Library

(Doench et al., 2016) and cloned into lentiCRISPRv2 according to standard protocols. Lentiviral particles were produced using 293T cells transfected with lentiCRISPRv2, envelope plasmid (pMD2.G), and packaging plasmid (psPAX2). pMD2.G and psPAX2 were gifts from D. Trono (Addgene plasmid #12259 and #12260, respectively). Host cells were infected with lentivirus and final concentration of polybrene (8 $\mu\text{g/ml}$; MilliporeSigma, #TR-1003-G) and selected with puromycin (Sigma-Aldrich, P9620) for 2 to 4 days at an empirically determined concentration and duration for each cell line. Knockout populations were confirmed by immunoblot.

For PC3M, we were unable to achieve population knockouts. Instead, clonal knockouts were generated by transducing with virus as described for populations. Single-cell clones were grown after fluorescence-activated cell sorting, and clonal knockouts were confirmed by Sanger sequencing and Western blot. Two clones of wild-type cell lines were also derived as controls. MDA-MB-231 dKO cells were generated through the two vector CRISPR-Cas9 system. Briefly, cells were infected and selected with sgRNA 1 (sgAAVS1 for all single knockouts and sgHSF2 for dKOs) in the lenti-guide-Puro vector (a gift from F. Zhang, plasmid #52963) lacking Cas9 so that cell editing would not yet begin. These cells were subsequently infected and selected with the second sgRNA cloned into the lentiCas9-Blast vector (a gift from F. Zhang, plasmid #52962). Control and single knockout cells received sgAAVS1 as a cutting control, and then the respective gRNAs were used in other cell lines (sgHSF1.1 or sgHSF2.1), with dKO cells receiving both HSF sgRNAs.

Table 3. sgRNA sequence for CRISPR/Cas9 knockout generation.

sgRNA Name	sgRNA Sequence
sgHSF1.1	GCTCCAGCAGATGAGCGCGT
sgHSF2.1	CGGCTTTCCTCAGCAAGCTG
sgHSF1.2	CCGGCGGGAGCATAGACGAG
sgHSF2.2	TATGCACCTGTCATTGAGAG
sgAAVS1	GGGGCCACTAGGGACAGGAT

4.14 RNA harvesting, library prep, sequencing

All reactions involving RNA were completed at an ribonuclease (RNase)-free bench. Cells were seeded in either 12-well or 6-well plates (two to three biological replicates). Cells were harvested with either TRIzol or Buffer RLT (QIAGEN) and then applied to Direct-zol (Zymo Research) or QIAGEN RNeasy Kit with deoxyribonuclease I on-column treatment according to the manufacturer's protocols (QIAGEN). RNA concentration and quality were determined using the Agilent 2100 Bioanalyzer or Agilent 4200 TapeStation according to the manufacturer's protocols. Samples with an RNA integrity number (RIN) greater than 7 were included in subsequent steps. Libraries were prepared from 100 ng of RNA using the Lexogen QuantSeq FWD Kit for Illumina sequencing using the Sciclone G3 NGS Workstation (PerkinElmer) according to the manufacturer's protocol, using 14 cycles of PCR library amplification.

DNA Library quality and fragment size were assessed using Agilent High Sensitivity DNA kits for either the Agilent 2100 Bioanalyzer or Agilent 4200 TapeStation. Library concentrations were determined using the Qubit dsDNA HS assay adapted to a 384-well format, scaling the reaction size down to 20 μ l in triplicate (19 μ l of working reagent and 1 μ l of sample or standard) with a series of 11 Qubit DNA standard dilutions (Invitrogen). Fluorescence was measured using a Tecan Infinite M1000 Pro plate reader (excitation, 480 nm; emission, 530 nm). Sequencing

samples were pooled at equimolar amounts, diluted to a 4 nM final library concentration, denatured with 0.2 M NaOH (final concentration) for 5 min at room temperature, and quenched with 200 mM tris-HCl (pH 7). PhiX spike-in (1%) was used and libraries were run on an Illumina NextSeq or NovaSeq

4.15 RNA-seq data processing and analysis

Raw Bcl files were converted to FASTQ using bcl2fastq software. Sequence quality was assessed with FastQC v0.11.2, FASTQ files were trimmed using BBduk from BBTools v35.92 according to Lexogen QuantSeq manufacturer's parameters, mapped to hg38 using STAR v2.6.0 and gene annotations from Ensembl 78. HTSeq was used to count uniquely mapped reads. Significantly differentially expressed genes were determined using EdgeR v3.26.8. A blocked experimental design analysis strategy was used (edgeR User's Guide, 3.4.2). This strategy limits inclusion of genes that are differentially expressed between different cell lines or cell type-specific siRNA treatment effects, thus emphasizing gene expression changes common to the treatment relative to control (siRNA versus siNT). This analysis was conducted independently for siHSF2 and siHSF1 and for each pool of siRNA hairpins. High confidence differentially expressed genes were defined as the intersection of differentially expressed genes across siRNA pools for HSF2 or HSF1 to minimize inclusion of nonspecific gene expression changes. For the correlation analyses in Fig. 4B, fold change gene expression for the union of differentially expressed genes in any cell line was used.

Enriched GO terms and p-values were determined with the Gene Set Enrichment Analysis (GSEA)(Liberzon et al., 2015; Subramanian et al., 2005) using the Hallmark (H), KEGG and Reactome (C2), GO biological process and molecular function (C5), and transcription factor

targets (TFT) in the Molecular Signatures Database (MSigDb, v7.2, msigdb.org). Analyses used Python v3.7.4, Pandas v0.25.1, Numpy v1.17.2, and data visualized with Seaborn v0.9.0 and Matplotlib 3.3.3.

4.16 ChIP-seq sample preparation and analysis

Cells were cross-linked with 1% formaldehyde for 10 min at room temperature, quenched with 0.125 M glycine (final concentration) for 5 min at room temperature, and washed twice with PBS. Cells were harvested with ice-cold PBS and pelleted at 1000g for 5 min at 4°C, flash-frozen in liquid nitrogen, and stored at -80°C before lysis and micrococcal nuclease (MNase) digestion. Cell pellets were lysed in ice-cold lysis buffer [1% SDS, 10 mM EDTA, 50 mM tris-HCl (pH 8), and 1× protease inhibitor] and gently shook for 10 min at 4°C. Cell suspensions were spun down at 1350g for 5 min at 4°C, supernatant discarded, and resuspended in ice-cold MNase digestion buffer [50 mM tris-HCl (pH 8), 5 mM CaCl₂, and 1× protease inhibitor]. MNase (ChIP grade, 100 U/μl) was diluted (1:10) with MNase digestion buffer, added to cell lysates, and incubated in a 37°C water bath for 30 min, mixing every 5 min. The reactions were quenched with 10 mM EDTA and 20 mM EGTA on ice for 5 min. Cell lysates were spun down at 1350g for 5 min at 4°C to recover nuclei. Cell nuclei were resuspended in 2 ml of ice-cold ChIP buffer [1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM tris-HCl (pH 8), 3 mM CaCl₂, and 1× protease inhibitor] and sonicated with 800 mg of sonication beads for five cycles of 30-s on/off. Sonicated lysates were recovered and spun down at high speed for 10 min at 4°C. Sheared chromatin (10%) was saved as an input, and the rest was incubated with Dynabeads Protein G magnetic beads (Thermo Fisher Scientific, catalog no. 10009D) containing HSF1 antibody validated previously (Mendillo et al.,

2012) (H-311, Santa Cruz Biotechnology) or HSF2 antibody (3E2, Santa Cruz Biotechnology) and rotated at 4°C overnight. Wash steps were conducted as detailed previously (Lee et al., 2006).

HSF-bound beads were recovered using magnetic particle concentrator, washed once with 1 ml of wash buffer B [20 mM tris-HCl (pH 8), 150 mM NaCl, 2 mM EDTA (pH 8), 0.1% SDS, and 1% Triton X-100], washed once with buffer C [20 mM tris-HCl (pH 8), 500 mM NaCl, 2 mM EDTA (pH 8), 0.1% SDS, and 1% Triton X-100], once with buffer D [10 mM tris-HCl (pH 8), 250 mM LiCl, 1 mM EDTA (pH 8), 1% Na-deoxycholate, and 1% IGEPAL CA-360], and once with buffer TE [10 mM tris-HCl (pH 8), 1 mM EDTA (pH 8), and 50 mM NaCl], discarding the supernatant each wash. Beads were spun down at 1000g for 3 min at 4°C to remove any residual TE buffer. Chromatin-protein complexes and 10% inputs were extracted and decross-linked with 300 µl of extraction buffer (20 mM tris-HCl, 10 mM EDTA, 5 mM EGTA, 1% SDS, 300 mM NaCl, and proteinase K) at 65°C overnight. Samples were cooled to room temperature and vortexed, and proteinase K was inactivated with an incubation at 95°C for 10 min. Beads were collected and eluted, and decross-linked chromatin was recovered. Chromatin was incubated with RNase A at final concentration (0.2 mg/ml) at 37°C for 2 hours. Fragmented DNA was purified with Zymo Research ChIP DNA concentrator kit.

Preparation of the ChIP-seq DNA library was performed with KAPA HTP library preparation kit (Roche, KK8234) and sequenced on an Illumina NextSeq as described above. Reads were aligned to hg38 using Bowtie v0.12.9. Peaks were called relative to knockout samples using MACS v1.4.2 and annotated from Ensembl 78. Heatmaps and metaplots were created using NGS plots. ChIP tracks were visualized using UCSC Genome Browser. Relative ChIP binding alongside RNA-seq data represents average peak intensity for MDA-MB-231 and PC3M

experiments. Binding for each gene is plotted as a percentage of the maximum peak for each HSF to account for differences in antibody binding intensity. The scale max is set at 10 to aid visualization of intermediate peak intensities.

4.17 Coessentiality analysis

Gene essentiality data derived from CRISPR-Cas9 genome-scale loss-of-function screening of 739 cancer cell lines (Meyers et al., 2017) using a modified Avana library (Doench et al., 2016) as part of Project Achilles was obtained from the Broad Institute's DepMap portal (20q1 release). Data was downloaded from <https://depmap.org/portal/download/>. Locus-adjusted gene coessentiality was determined as described previously (Amici et al., 2021) and implemented at fireworks.mendillolab.org. Briefly, the dependency score for each gene is corrected using a sliding window approach that minimizes bias at genomic regions which have variable copy number across cell lines. Then, all possible gene-gene correlations are assessed and ranked by Pearson correlation coefficient. Highly ranked correlations indicate similar contextual essentiality and imply functional importance in the same pathway(s).

4.18 Xenograft experiments and analysis

MDA-MB-231 population knockouts or PC3M clonal knockouts and their respective controls were engineered to express pUltra-Chili-Luciferase (Addgene plasmid #48688). NSG mice were obtained from the Jackson Laboratory, female mice were used for MDA-MB-231 breast cancer experiments, and male mice were used for PC3M prostate cancer experiments. Cells were suspended in PBS and Matrigel (50:50), and 5×10^6 cells per mouse were injected into the right mammary fat pad in a volume of 0.05 ml. Eight mice were inoculated for each genetic group (e.g.,

HSF1 knockout, sgNT control) for each cell line. Body weight and tumor volume were measured using calipers twice per week. Mice were euthanized when tumor volume reached 1500 mm³. A repeated-measures analysis of variance (ANOVA) was conducted using GraphPad Prism 9 with Dunnett's multiple comparison test corrections. Significance was considered $\alpha < 0.05$. Survival analyses were conducted with a log-rank Mantel-Cox test in GraphPad Prism 9. No significant differences were observed for body weight in either experiment. Data are plotted as a summary of all eight mice of the same knockout status. All animal experiments were performed according to Institutional Animal Care and Use Committee–approved protocols.

4.19 Xenograft RNA sequencing and analysis

A slice of flash-frozen tumors (no more than 30 mg) was lysed according to the manufacturer's protocol with a rotor-stator homogenizer using the RNeasy Kit (QIAGEN). One section from each of four independent tumors of similar size at time of experiment termination was used. RNA quantity and quality were determined as above. NEBNext Ultra II DNA Library Prep Kit for Illumina was used according to the manufacturer's protocol (NEB #E7103) and sequenced as described above. Where necessary, reads mapping better to mouse than human genomes were removed using Disambiguate software package using the Python implementation (Ahdesmaki et al., 2016). Samples containing too few reads mapped to the human genome were removed from the analysis which included one MDA-MB-231 sgNT, and two PC3Mc2 sgHSF2 tumor samples.

4.20 Data and materials availability

Data generated in this study can be found at www.ncbi.nlm.nih.gov/geo/ under accession code GSE194098. All data needed to evaluate the conclusions are present in the paper and/or the Supplementary Materials available at: <https://www.science.org/doi/10.1126/sciadv.abj6526>.

REFERENCES

Ahdesmaki, M.J., Gray, S.R., Johnson, J.H., and Lai, Z. (2016). Disambiguate: An open-source application for disambiguating two species in next generation sequencing data from grafted samples. *F1000Res* 5, 2741. 10.12688/f1000research.10082.2.

Akerfelt, M., Henriksson, E., Laiho, A., Vihervaara, A., Rautoma, K., Kotaja, N., and Sistonen, L. (2008). Promoter CHIP-chip analysis in mouse testis reveals Y chromosome occupancy by HSF2. *P Natl Acad Sci USA* 105, 11224-11229. 10.1073/pnas.0800620105.

Akerfelt, M., Morimoto, R.I., and Sistonen, L. (2010a). Heat shock factors: integrators of cell stress, development and lifespan. *Nat Rev Mol Cell Biol* 11, 545-555. 10.1038/nrm2938.

Akerfelt, M., Vihervaara, A., Laiho, A., Conter, A., Christians, E.S., Sistonen, L., and Henriksson, E. (2010b). Heat shock transcription factor 1 localizes to sex chromatin during meiotic repression. *J Biol Chem* 285, 34469-34476. 10.1074/jbc.M110.157552.

Alasady, M.J., and Mendillo, M.L. (2020). The Multifaceted Role of HSF1 in Tumorigenesis. *Adv Exp Med Biol* 1243, 69-85. 10.1007/978-3-030-40204-4_5.

Amici, D.R., Jackson, J.M., Truica, M.I., Smith, R.S., Abdulkadir, S.A., and Mendillo, M.L. (2021). FIREWORKS: a bottom-up approach to integrative coessentiality network analysis. *Life Sci Alliance* 4. 10.26508/lsa.202000882.

Anckar, J., and Sistonen, L. (2011). Regulation of HSF1 Function in the Heat Stress Response: Implications in Aging and Disease. *Annual Review of Biochemistry*, Vol 80 80, 1089-1115. 10.1146/annurev-biochem-060809-095203.

Aury, J.M., Jaillon, O., Duret, L., Noel, B., Jubin, C., Porcel, B.M., Segurens, B., Daubin, V., Anthouard, V., Aiach, N., et al. (2006). Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature* 444, 171-178. 10.1038/nature05230.

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Peña-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., et al. (2008). Variation in Homeodomain DNA Binding Revealed by High-Resolution Analysis of Sequence Preferences. *Cell* 133, 1266-1276. 10.1016/J.CELL.2008.05.024/ATTACHMENT/E35B0839-A901-4917-97B7-7AAAF4578037/MMC1.PDF.

Bierkamp, C., Luxey, M., Metchat, A., Audouard, C., Dumollard, R., and Christians, E. (2010). Lack of maternal Heat Shock Factor 1 results in multiple cellular and developmental defects, including mitochondrial damage and altered redox homeostasis, and leads to reduced survival of mammalian oocytes and embryos. *Dev Biol* 339, 338-353. 10.1016/j.ydbio.2009.12.037.

Bjork, J.K., Ahonen, I., Mirtti, T., Erickson, A., Rannikko, A., Butzow, A., Nordling, S., Lundin, J., Lundin, M., Sistonen, L., et al. (2018). Increased HSF1 expression predicts shorter disease-specific survival of prostate cancer patients following radical prostatectomy. *Oncotarget* 9, 31200-31213. 10.18632/oncotarget.25756.

Bjork, J.K., Akerfelt, M., Joutsen, J., Puustinen, M.C., Cheng, F., Sistonen, L., and Nees, M. (2016). Heat-shock factor 2 is a suppressor of prostate cancer invasion. *Oncogene* 35, 1770-1784. 10.1038/onc.2015.241.

Blanc, G., and Wolfe, K.H. (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *Plant Cell* 16, 1679-1691. 10.1105/tpc.021410.

Boone, C., Bussey, H., and Andrews, B.J. (2007). Exploring genetic interactions and networks with yeast. *Nat Rev Genet* 8, 437-449. 10.1038/nrg2085.

Bridges, C.B. (1935). Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila melanogaster*. *J Hered* 26, 60-64. DOI 10.1093/oxfordjournals.jhered.a104022.

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Liechti, A., Ascencao, K., Rummel, C., Ovchinnikova, S., et al. (2019). Gene expression across mammalian organ development. *Nature* 571, 505-509. 10.1038/s41586-019-1338-5.

Carithers, L.J., Ardlie, K., Barcus, M., Branton, P.A., Britton, A., Buia, S.A., Compton, C.C., DeLuca, D.S., Peter-Demchok, J., Gelfand, E.T., et al. (2015). A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank* 13, 311-319. 10.1089/bio.2015.0032.

Chang, Y.H., Ostling, P., Akerfelt, M., Trouillet, D., Rallu, M., Gitton, Y., El Fatilmy, R., Fardeau, V., Le Crom, S., Morange, M., et al. (2006). Role of heat-shock factor 2 in cerebral cortex formation and as a regulator of p35 expression. *Gene Dev* 20, 836-847. 10.1101/gad.366906.

Chen, F., Fan, Y.M., Liu, X.P., Zhang, J.H., Shang, Y.A., Zhang, B., Liu, B., Hou, J.J., Cao, P.X., and Tan, K. (2022). Pan-Cancer Integrated Analysis of HSF2 Expression, Prognostic Value and Potential Implications for Cancer Immunity. *Front Mol Biosci* 8. 10.3389/fmolb.2021.789703.

Chen, H.J., Mitchell, J.C., Novoselov, S., Miller, J., Nishimura, A.L., Scotter, E.L., Vance, C.A., Cheetham, M.E., and Shaw, C.E. (2016). The heat shock response plays an important role in TDP-43 clearance: evidence for dysfunction in amyotrophic lateral sclerosis. *Brain* 139, 1417-1432. 10.1093/brain/aww028.

Chen, R., Liliental, J.E., Kowalski, P.E., Lu, Q., and Cohen, S.N. (2011). Regulation of transcription of hypoxia-inducible factor-1 alpha (HIF-1 alpha) by heat shock factors HSF2 and HSF4. *Oncogene* 30, 2570-2580. 10.1038/onc.2010.623.

Conant, G.C., and Wolfe, K.H. (2006). Functional partitioning of yeast co-expression networks after genome duplication. *Plos Biology* 4, 545-554. 10.1371/journal.pbio.0040109.

Conant, G.C., and Wolfe, K.H. (2008). Turning a hobby into a job: How duplicated genes find new functions. *Nature Reviews Genetics* 9, 938-950. 10.1038/nrg2482.

Cunningham, F., Allen, J.E., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Austine-Orimoloye, O., Azov, A.G., Barnes, I., Bennett, R., et al. (2022). Ensembl 2022. *Nucleic Acids Res* 50, D988-D995. 10.1093/nar/gkab1049.

Dai, C. (2018). The heat-shock, or HSF1-mediated proteotoxic stress, response in cancer: from proteomic stability to oncogenesis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 373. 10.1098/RSTB.2016.0525.

Dai, C., and Sampson, S.B. (2016). HSF1: Guardian of Proteostasis in Cancer. *Trends Cell Biol* 26, 17-28. 10.1016/j.tcb.2015.10.011.

Dai, C., Santagata, S., Tang, Z., Shi, J., Cao, J., Kwon, H., Bronson, R.T., Whitesell, L., and Lindquist, S. (2012). Loss of tumor suppressor NF1 activates HSF1 to promote carcinogenesis. *J Clin Invest* 122, 3742-3754. 10.1172/JCI62727.

Dai, C., Whitesell, L., Rogers, A.B., and Lindquist, S. (2007). Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. *Cell* 130, 1005-1018. 10.1016/j.cell.2007.07.020.

Darwin, C. (1859). *On The Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life* (London: John Murray).

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34, 184-191. 10.1038/nbt.3437.

Dong, B.S., Jaeger, A.M., Hughes, P.F., Loiselle, D.R., Hauck, J.S., Fu, Y., Haystead, T.A., Huang, J.T., and Thiele, D.J. (2020). Targeting therapy-resistant prostate cancer via a direct inhibitor of the human heat shock transcription factor 1. *Science Translational Medicine* 12. 10.1126/scitranslmed.abb5647.

Duchateau, A., de Thonel, A., El Fatimy, R., Dubreuil, V., and Mezger, V. (2020). The "HSF connection": Pleiotropic regulation and activities of Heat Shock Factors shape pathophysiological brain development. *Neurosci Lett* 725, 134895. 10.1016/j.neulet.2020.134895.

El Fatimy, R., Miozzo, F., Le Mouel, A., Abane, R., Schwendimann, L., Saberan-Djoneidi, D., de Thonel, A., Massaoudi, I., Paslaru, L., Hashimoto-Torii, K., et al. (2014). Heat shock factor 2 is a stress-responsive mediator of neuronal migration defects in models of fetal alcohol syndrome. *EMBO Mol Med* 6, 1043-1061. 10.15252/emmm.201303311.

Engelman, J.A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J.O., Lindeman, N., Gale, C.M., Zhao, X., Christensen, J., et al. (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 316, 1039-1043. 10.1126/science.1141478.

Farkas, T., Kutskova, Y.A., and Zimarino, V. (1998). Intramolecular repression of mouse heat shock factor 1. *Mol Cell Biol* 18, 906-918. 10.1128/MCB.18.2.906.

Feng, N., Feng, H., Wang, S., Punekar, A.S., Ladenstein, R., Wang, D.C., Zhang, Q., Ding, J., and Liu, W. (2021). Structures of heat shock factor trimers bound to DNA. *iScience* 24, 102951. 10.1016/j.isci.2021.102951.

Fujimoto, M., Izu, H., Seki, K., Fukuda, K., Nishida, T., Yamada, S., Kato, K., Yonemura, S., Inouye, S., and Nakai, A. (2004). HSF4 is required for normal cell growth and differentiation during mouse lens development. *Embo Journal* 23, 4297-4306. 10.1038/sj.emboj.7600435.

Fujimoto, M., and Nakai, A. (2010). The heat shock factor family and adaptation to proteotoxic stress. *FEBS J* 277, 4112-4125. 10.1111/j.1742-4658.2010.07827.x.

Fujimoto, M., Oshima, K., Shinkawa, T., Wang, B.B., Inouye, S., Hayashida, N., Takii, R., and Nakai, A. (2008). Analysis of HSF4 binding regions reveals its necessity for gene regulation during development and heat shock response in mouse lenses. *J Biol Chem* 283, 29961-29970. 10.1074/jbc.M804629200.

Fujimoto, M., Takaki, E., Hayashi, T., Kitaura, Y., Tanaka, Y., Inouye, S., and Nakai, A. (2005). Active HSF1 significantly suppresses polyglutamine aggregate formation in cellular and mouse models. *J Biol Chem* 280, 34908-34916. 10.1074/jbc.M506288200.

Gabai, V.L., Meng, L., Kim, G., Mills, T.A., Benjamin, I.J., and Sherman, M.Y. (2012). Heat Shock Transcription Factor Hsf1 Is Involved in Tumor Progression via Regulation of Hypoxia-Inducible Factor 1 and RNA-Binding Protein HuR. *Molecular and Cellular Biology* 32, 929-940. 10.1128/Mcb.05921-11.

Gomez-Pastor, R., Burchfiel, E.T., Neef, D.W., Jaeger, A.M., Cabisco, E., McKinstry, S.U., Doss, A., Aballay, A., Lo, D.C., Akimov, S.S., et al. (2017). Abnormal degradation of the neuronal stress-protective transcription factor HSF1 in Huntington's disease. *Nat Commun* 8, 14405. 10.1038/ncomms14405.

Gomez-Pastor, R., Burchfiel, E.T., and Thiele, D.J. (2018). Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat Rev Mol Cell Biol* 19, 4-19. 10.1038/nrm.2017.73.

Green, M., Schuetz, T.J., Sullivan, E.K., and Kingston, R.E. (1995). A heat shock-responsive domain of human HSF1 that regulates transcription activation domain function. *Mol Cell Biol* 15, 3354-3362. 10.1128/MCB.15.6.3354.

Guertin, M.J., and Lis, J.T. (2010). Chromatin landscape dictates HSF binding to target DNA elements. *PLoS Genet* 6, e1001114. 10.1371/journal.pgen.1001114.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature* 431, 99-104. 10.1038/nature02800.

Hayashida, N., Fujimoto, M., Tan, K., Prakasam, R., Shinkawa, T., Li, L., Ichikawa, H., Takii, R., and Nakai, A. (2010). Heat shock factor 1 ameliorates proteotoxicity in cooperation with the transcription factor NFAT. *EMBO J* 29, 3459-3469. 10.1038/emboj.2010.225.

Hentze, N., Le Breton, L., Wiesner, J., Kempf, G., and Mayer, M.P. (2016). Molecular mechanism of thermosensory function of human heat shock transcription factor Hsf1. *Elife* 5, 1-24. 10.7554/eLife.11576.

Hickox, A.E., Wong, A.C., Pak, K., Strojny, C., Ramirez, M., Yates, J.R., 3rd, Ryan, A.F., and Savas, J.N. (2017). Global Analysis of Protein Expression of Inner Ear Hair Cells. *J Neurosci* 37, 1320-1339. 10.1523/JNEUROSCI.2267-16.2016.

Homma, S., Jin, X., Wang, G., Tu, N., Min, J., Yanasak, N., and Mivechi, N.F. (2007). Demyelination, astrogliosis, and accumulation of ubiquitinated proteins, hallmarks of CNS disease in hsf1-deficient mice. *J Neurosci* 27, 7974-7986. 10.1523/JNEUROSCI.0006-07.2007.

Jaeger, A.M., Pemble, C.W.t., Sistonen, L., and Thiele, D.J. (2016). Structures of HSF2 reveal mechanisms for differential regulation of human heat-shock factors. *Nat Struct Mol Biol* 23, 147-154. 10.1038/nsmb.3150.

Jedlicka, P., Mortin, M.A., and Wu, C. (1997). Multiple functions of *Drosophila* heat shock transcription factor in vivo. *EMBO J* 16, 2452-2462. 10.1093/emboj/16.9.2452.

Jiang, Y.Q., Wang, X.L., Cao, X.H., Ye, Z.Y., Li, L., and Cai, W.Q. (2013). Increased heat shock transcription factor 1 in the cerebellum reverses the deficiency of Purkinje cells in Alzheimer's disease. *Brain Res* 1519, 105-111. 10.1016/j.brainres.2013.04.059.

Jin, X., Moskophidis, D., and Mivechi, N.F. (2011). Heat shock transcription factor 1 is a key determinant of HCC development by regulating hepatic steatosis and metabolic syndrome. *Cell Metab* 14, 91-103. 10.1016/j.cmet.2011.03.025.

Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-Binding Specificities of Human Transcription Factors. *Cell* 152, 327-339. 10.1016/J.CELL.2012.12.009.

Joutsen, J., Da Silva, A.J., Luoto, J.C., Budzynski, M.A., Nylund, A.S., de Thonel, A., Concordet, J.P., Mezger, V., Saberan-Djoneidi, D., Henriksson, E., and Sistonen, L. (2020). Heat Shock Factor 2 Protects against Proteotoxicity by Maintaining Cell-Cell Adhesion. *Cell Rep* 30, 583-597 e586. 10.1016/j.celrep.2019.12.037.

Kallio, M., Chang, Y.H., Manuel, M., Alastalo, T.P., Rallu, M., Gitton, Y., Pirkkala, L., Loones, M.T., Paslaru, L., Larney, S., et al. (2002). Brain abnormalities, defective meiotic chromosome synapsis and female subfertility in HSF2 null mice. *Embo Journal* 21, 2591-2601. DOI 10.1093/emboj/21.11.2591.

Kim, E., Wang, B., Sastry, N., Maslah, E., Nelson, P.T., Cai, H., and Liao, F.F. (2016). NEDD4-mediated HSF1 degradation underlies alpha-synucleinopathy. *Hum Mol Genet* 25, 211-222. 10.1093/hmg/ddv445.

Koonin, E.V. (2005). Orthologs, Paralogs, and Evolutionary Genomics 1. 10.1146/annurev.genet.39.073003.114725.

Korfanty, J., Stokowy, T., Widlak, P., Gogler-Piglowska, A., Handschuh, L., Podkowinski, J., Vydra, N., Naumowicz, A., Toma-Jonik, A., and Widlak, W. (2014). Crosstalk between HSF1 and HSF2 during the heat shock response in mouse testes. *Int J Biochem Cell Biol* 57, 76-83. 10.1016/j.biocel.2014.10.006.

Kourtis, N., Lazaris, C., Hockemeyer, K., Balandran, J.C., Jimenez, A.R., Mullenders, J., Gong, Y., Trimarchi, T., Bhatt, K., Hu, H., et al. (2018). Oncogenic hijacking of the stress response machinery in T cell acute lymphoblastic leukemia. *Nat Med* 24, 1157-1166. 10.1038/s41591-018-0105-8.

Lan, X., and Pritchard, J.K. (2016). Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals. *Science* 352, 1009-1013. 10.1126/science.aad8411.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc* 1, 729-748. 10.1038/nprot.2006.98.

Li, J., Chauve, L., Phelps, G., Brielmann, R.M., and Morimoto, R.I. (2016). E2F coregulates an essential HSF developmental program that is distinct from the heat-shock response. *Genes Dev* 30, 2062-2075. 10.1101/gad.283317.116.

- Li, J., Labbadia, J., and Morimoto, R.I. (2017). Rethinking HSF1 in Stress, Development, and Organismal Health. *Trends Cell Biol* 27, 895-905. 10.1016/j.tcb.2017.08.002.
- Li, P., Sheng, C., Huang, L.L., Zhang, H., Huang, L.H., Cheng, Z.N., and Zhu, Q.B. (2014). MiR-183/-96/-182 cluster is up-regulated in most breast cancers and increases cell proliferation and migration. *Breast Cancer Research* 16. 10.1186/s13058-014-0473-z.
- Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1, 417-425. 10.1016/j.cels.2015.12.004.
- Lindquist, S. (1986). The heat-shock response. *Annu Rev Biochem* 55, 1151-1191. 10.1146/annurev.bi.55.070186.005443.
- Lindquist, S., and Craig, E.A. (1988). The heat-shock proteins. *Annu Rev Genet* 22, 631-677. 10.1146/annurev.ge.22.120188.003215.
- Littlefield, O., and Nelson, H.C. (1999). A new use for the 'wing' of the 'winged' helix-turn-helix motif in the HSF-DNA cocystal. *Nat Struct Biol* 6, 464-470. 10.1038/8269.
- Liu, X.D., Liu, P.C., Santoro, N., and Thiele, D.J. (1997). Conservation of a stress response: human heat shock transcription factors functionally substitute for yeast HSF. *EMBO J* 16, 6466-6477. 10.1093/emboj/16.21.6466.
- Luo, J., Solimini, N.L., and Elledge, S.J. (2009). Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* 136, 823-837. 10.1016/j.cell.2009.02.024.
- Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151-1155. 10.1126/SCIENCE.290.5494.1151/ASSET/47A1732E-F865-418C-AA6A-63332D48FC0F/ASSETS/GRAPHIC/SE4408976003.JPEG.
- Lynch, M., and Katju, V. (2004). The altered evolutionary trajectories of gene duplicates. *Trends in Genetics* 20, 544-549. 10.1016/j.tig.2004.09.001.
- Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 10.1093/nar/gkac240.

Mahat, D.B., Salamanca, H.H., Duarte, F.M., Danko, C.G., and Lis, J.T. (2016). Mammalian Heat Shock Response and Mechanisms Underlying Its Genome-wide Transcriptional Regulation. *Mol Cell* 62, 63-78. 10.1016/j.molcel.2016.02.025.

Mathew, A., Mathur, S.K., Jolly, C., Fox, S.G., Kim, S., and Morimoto, R.I. (2001). Stress-specific activation and repression of heat shock factors 1 and 2. *Mol Cell Biol* 21, 7163-7171. 10.1128/MCB.21.21.7163-7171.2001.

Mathew, A., Mathur, S.K., and Morimoto, R.I. (1998). Heat shock response and protein degradation: regulation of HSF2 by the ubiquitin-proteasome pathway. *Mol Cell Biol* 18, 5091-5098. 10.1128/mcb.18.9.5091.

Mavroudis, I.A., Fotiou, D.F., Adipepe, L.F., Manani, M.G., Njau, S.D., Psaroulis, D., Costa, V.G., and Baloyannis, S.J. (2010). Morphological Changes of the Human Purkinje Cells and Deposition of Neuritic Plaques and Neurofibrillary Tangles on the Cerebellar Cortex of Alzheimer's Disease. *Am J Alzheimers Dis* 25, 585-591. 10.1177/1533317510382892.

McMillan, D.R., Xiao, X., Shao, L., Graves, K., and Benjamin, I.J. (1998). Targeted disruption of heat shock transcription factor 1 abolishes thermotolerance and protection against heat-inducible apoptosis. *J Biol Chem* 273, 7523-7528. 10.1074/jbc.273.13.7523.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardi, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* 10, 730-736. 10.1038/nmeth.2557.

Mendillo, M.L., Santagata, S., Koeva, M., Bell, G.W., Hu, R., Tamimi, R.M., Fraenkel, E., Ince, T.A., Whitesell, L., and Lindquist, S. (2012). HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. *Cell* 150, 549-562. 10.1016/j.cell.2012.06.031.

Meng, L., Gabai, V.L., and Sherman, M.Y. (2010). Heat-shock transcription factor HSF1 has a critical role in human epidermal growth factor receptor-2-induced cellular transformation and tumorigenesis. *Oncogene* 29, 5204-5213. 10.1038/onc.2010.277.

Meng, X., Chen, X., Lu, P., Ma, W., Yue, D., Song, L., and Fan, Q. (2017). miR-202 Promotes Cell Apoptosis in Esophageal Squamous Cell Carcinoma by Targeting HSF2. *Oncol Res* 25, 215-223. 10.3727/096504016X14732772150541.

Metchat, A., Akerfelt, M., Bierkamp, C., Delsinne, V., Sistonen, L., Alexandre, H., and Christians, E.S. (2009). Mammalian Heat Shock Factor 1 Is Essential for Oocyte Meiosis and Directly Regulates Hsp90 alpha Expression. *Journal of Biological Chemistry* 284, 9521-9528. 10.1074/jbc.M808819200.

Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H., Dharia, N.V., Montgomery, P.G., Cowley, G.S., Pantel, S., et al. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* 49, 1779-1784. 10.1038/ng.3984.

Min, J.N., Huang, L., Zimonjic, D.B., Moskophidis, D., and Mivechi, N.F. (2007). Selective suppression of lymphomas by functional loss of Hsf1 in a p53-deficient mouse model for spontaneous tumors. *Oncogene* 26, 5086-5097. 10.1038/sj.onc.1210317.

Mosser, D.D., Theodorakis, N.G., and Morimoto, R.I. (1988). Coordinate changes in heat shock element-binding activity and HSP70 gene transcription rates in human cells. *Mol Cell Biol* 8, 4736-4744. 10.1128/mcb.8.11.4736.

Munoz-Sanchez, J., and Chanez-Cardenas, M.E. (2019). The use of cobalt chloride as a chemical hypoxia model. *J Appl Toxicol* 39, 556-570. 10.1002/jat.3749.

Nakai, A., Tanabe, M., Kawazoe, Y., Inazawa, J., Morimoto, R.I., and Nagata, K. (1997). HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. *Molecular and Cellular Biology* 17, 469-481. Doi 10.1128/Mcb.17.1.469.

Neef, D.W., Jaeger, A.M., Gomez-Pastor, R., Willmund, F., Frydman, J., and Thiele, D.J. (2014). A direct regulatory interaction between chaperonin TRiC and stress-responsive transcription factor HSF1. *Cell Rep* 9, 955-966. 10.1016/j.celrep.2014.09.056.

Neef, D.W., Jaeger, A.M., and Thiele, D.J. (2011). Heat shock transcription factor 1 as a therapeutic target in neurodegenerative diseases. *Nat Rev Drug Discov* 10, 930-944. 10.1038/nrd3453.

Neef, D.W., Turski, M.L., and Thiele, D.J. (2010). Modulation of Heat Shock Transcription Factor 1 as a Therapeutic Target for Small Molecule Intervention in Neurodegenerative Disease. *Plos Biology* 8. 10.1371/journal.pbio.1000291.

Neudegger, T., Verghese, J., Hayer-Hartl, M., Hartl, F.U., and Bracher, A. (2016). Structure of human heat-shock transcription factor 1 in complex with DNA. *Nat Struct Mol Biol* 23, 140-146. 10.1038/nsmb.3149.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.P., Tong, F., et al. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10, 515-527. 10.1016/j.ccr.2006.10.008.

Ohno, S. (1970). *Evolution by gene duplication* (Springer). <https://doi.org/10.1007/978-3-642-86659-3>.

Ostling, P., Bjork, J.K., Roos-Mattjus, P., Mezger, V., and Sistonen, L. (2007). Heat shock factor 2 (HSF2) contributes to inducible expression of hsp genes through interplay with HSF1. *J Biol Chem* 282, 7077-7086. 10.1074/jbc.M607556200.

Pan, J., Meyers, R.M., Michel, B.C., Mashtalir, N., Sizemore, A.E., Wells, J.N., Cassel, S.H., Vazquez, F., Weir, B.A., Hahn, W.C., et al. (2018). Interrogation of Mammalian Protein Complex Structure, Function, and Membership Using Genome-Scale Fitness Screens. *Cell Syst* 6, 555-568 e557. 10.1016/j.cels.2018.04.011.

Parker, C.S., and Topol, J. (1984). A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an hsp 70 gene. *Cell* 37, 273-283. 10.1016/0092-8674(84)90323-4.

Peterfy, M., Phan, J., Xu, P., and Reue, K. (2001). Lipodystrophy in the fld mouse results from mutation of a new gene encoding a nuclear protein, lipin. *Nat Genet* 27, 121-124. 10.1038/83685.

Pincus, D. (2020). Regulation of Hsf1 and the Heat Shock Response. *Adv Exp Med Biol* 1243, 41-50. 10.1007/978-3-030-40204-4_3.

Puustinen, M.C., and Sistonen, L. (2020). Molecular Mechanisms of Heat Shock Factors in Cancer. *Cells* 9. 10.3390/cells9051202.

Qiao, A., Jin, X., Pang, J., Moskophidis, D., and Mivechi, N.F. (2017). The transcriptional regulator of the chaperone response HSF1 controls hepatic bioenergetics and protein homeostasis. *J Cell Biol* 216, 723-741. 10.1083/jcb.201607091.

Rabindran, S.K., Haroun, R.I., Clos, J., Wisniewski, J., and Wu, C. (1993). Regulation of Heat-Shock Factor Trimer Formation - Role of a Conserved Leucine Zipper. *Science* 259, 230-234. DOI 10.1126/science.8421783.

Ritossa, F. (1962). New Puffing Pattern Induced by Temperature Shock and Dnp in *Drosophila*. *Experientia* 18, 571-&. Doi 10.1007/Bf02172188.

Riva, L., Koeva, M., Yildirim, F., Pirhaji, L., Dinesh, D., Mazor, T., Duennwald, M.L., and Fraenkel, E. (2012). Poly-glutamine expanded huntingtin dramatically alters the genome wide binding of HSF1. *J Huntingtons Dis* 1, 33-45.

Rogers, J.M., and Bulyk, M.L. (2018). Diversification of transcription factor-DNA interactions and the evolution of gene regulatory networks. *Wiley Interdiscip Rev Syst Biol Med*, e1423. 10.1002/wsbm.1423.

Roos-Mattjus, P., and Sistonen, L. (2021). Interplay between mammalian heat shock factors 1 and 2 in physiology and pathology. *FEBS J.* 10.1111/febs.16178.

Sandqvist, A., Bjork, J.K., Akerfelt, M., Chitikova, Z., Grichine, A., Vourc'h, C., Jolly, C., Salminen, T.A., Nymalm, Y., and Sistonen, L. (2009). Heterotrimerization of heat-shock factors 1 and 2 provides a transcriptional switch in response to distinct stimuli. *Mol Biol Cell* 20, 1340-1347. 10.1091/mbc.E08-08-0864.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11, 783-784. 10.1038/nmeth.3047.

Santagata, S., Hu, R., Lin, N.U., Mendillo, M.L., Collins, L.C., Hankinson, S.E., Schnitt, S.J., Whitesell, L., Tamimi, R.M., Lindquist, S., and Ince, T.A. (2011). High levels of nuclear heat-shock factor 1 (HSF1) are associated with poor prognosis in breast cancer. *Proc Natl Acad Sci U S A* 108, 18378-18383. 10.1073/pnas.1115031108.

Santagata, S., Mendillo, M.L., Tang, Y.C., Subramanian, A., Perley, C.C., Roche, S.P., Wong, B., Narayan, R., Kwon, H., Koeva, M., et al. (2013). Tight coordination of protein translation and HSF1 activation supports the anabolic malignant state. *Science* 341, 1238303. 10.1126/science.1238303.

- Santopolo, S., Riccio, A., Rossi, A., and Santoro, M.G. (2021). The proteostasis guardian HSF1 directs the transcription of its paralog and interactor HSF2 during proteasome dysfunction. *Cell Mol Life Sci* 78, 1113-1129. 10.1007/s00018-020-03568-x.
- Santos, S.D., and Saraiva, M.J. (2004). Enlarged ventricles, astrogliosis and neurodegeneration in heat shock factor 1 null mouse brain. *Neuroscience* 126, 657-663. 10.1016/j.neuroscience.2004.03.023.
- Sarge, K.D., Parksarge, O.K., Kirby, J.D., Mayo, K.E., and Morimoto, R.I. (1994). Expression of Heat-Shock Factor-2 in Mouse Testis - Potential Role as a Regulator of Heat-Shock Protein Gene-Expression during Spermatogenesis. *Biol Reprod* 50, 1334-1343. DOI 10.1095/biolreprod50.6.1334.
- Scherz-Shouval, R., Santagata, S., Mendillo, M.L., Sholl, L.M., Ben-Aharon, I., Beck, A.H., Dias-Santagata, D., Koeva, M., Stemmer, S.M., Whitesell, L., and Lindquist, S. (2014). The reprogramming of tumor stroma by HSF1 is a potent enabler of malignancy. *Cell* 158, 564-578. 10.1016/j.cell.2014.05.045.
- Schuetz, T.J., Gallo, G.J., Sheldon, L., Tempst, P., and Kingston, R.E. (1991). Isolation of a cDNA for HSF2: Evidence for two Heat-Shock Factor Genes in Humans. *P Natl Acad Sci USA* 88, 6911-6915. DOI 10.1073/pnas.88.16.6911.
- Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84-87. 10.1126/science.1247005.
- Shi, Y., Mosser, D.D., and Morimoto, R.I. (1998). Molecular chaperones as HSF1-specific transcriptional repressors. *Genes Dev* 12, 654-666. 10.1101/gad.12.5.654.
- Shinkawa, T., Tan, K., Fujimoto, M., Hayashida, N., Yamamoto, K., Takaki, E., Takii, R., Prakasam, R., Inouye, S., Mezger, V., and Nakai, A. (2011). Heat shock factor 2 is required for maintaining proteostasis against febrile-range thermal stress and polyglutamine aggregation. *Mol Biol Cell* 22, 3571-3583. 10.1091/mbc.E11-04-0330.
- Silveira, M.A., Tav, C., Berube-Simard, F.A., Cuppens, T., Leclercq, M., Fournier, E., Cote, M.C., Droit, A., and Bilodeau, S. (2021). Modulating HSF1 levels impacts expression of the estrogen receptor alpha and antiestrogen response. *Life Sci Alliance* 4. 10.26508/lsa.202000811.

Sistonen, L., Sarge, K.D., and Morimoto, R.I. (1994). Human heat shock factors 1 and 2 are differentially activated and can synergistically induce hsp70 gene transcription. *Mol Cell Biol* *14*, 2087-2099. 10.1128/mcb.14.3.2087.

Sistonen, L., Sarge, K.D., Phillips, B., Abravaya, K., and Morimoto, R.I. (1992). Activation of heat shock factor 2 during hemin-induced differentiation of human erythroleukemia cells. *Mol Cell Biol* *12*, 4104-4111. 10.1128/mcb.12.9.4104.

Skloot, R. (2010). *The immortal life of Henrietta Lacks* (Crown Publishers).

Smith, R.S., Takagishi, S.R., Amici, D.R., Metz, K., Gayatri, S., Alasady, M.J., Wu, Y., Brockway, S., Taiberg, S.L., Khalatyan, N., et al. (2022). HSF2 cooperates with HSF1 to drive a transcriptional program critical for the malignant state. *Sci Adv* *8*, eabj6526. 10.1126/sciadv.abj6526.

Solimini, N.L., Luo, J., and Elledge, S.J. (2007). Non-oncogene addiction and the stress phenotype of cancer cells. *Cell* *130*, 986-988. 10.1016/j.cell.2007.09.007.

Solis, E.J., Pandey, J.P., Zheng, X., Jin, D.X., Gupta, P.B., Airoidi, E.M., Pincus, D., and Denic, V. (2016). Defining the Essential Function of Yeast Hsf1 Reveals a Compact Transcriptional Program for Maintaining Eukaryotic Proteostasis. *Mol Cell* *63*, 60-71. 10.1016/j.molcel.2016.05.014.

Sorger, P.K., and Nelson, H.C. (1989). Trimerization of a yeast transcriptional activator via a coiled-coil motif. *Cell* *59*, 807-813. 10.1016/0092-8674(89)90604-1.

Su, K.H., Dai, S., Tang, Z., Xu, M., and Dai, C. (2019). Heat Shock Factor 1 Is a Direct Antagonist of AMP-Activated Protein Kinase. *Mol Cell* *76*, 546-561 e548. 10.1016/j.molcel.2019.08.021.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* *102*, 15545-15550. 10.1073/pnas.0506580102.

Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., and Lindquist, S. (2012). Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. *Cell* *150*, 987-1001. 10.1016/j.cell.2012.06.047.

Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z.Y., Larsen, B., Choi, H., Berger, B., Gingras, A.C., and Lindquist, S. (2014). A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* *158*, 434-448. 10.1016/j.cell.2014.05.039.

Takaki, E., Fujimoto, M., Sugahara, K., Nakahari, T., Yonemura, S., Tanaka, Y., Hayashida, N., Inouye, S., Takemoto, T., Yamashita, H., and Nakai, A. (2006). Maintenance of olfactory neurogenesis requires HSF1, a major heat shock transcription factor in mice. *J Biol Chem* *281*, 4931-4937. 10.1074/jbc.M506911200.

Tanabe, M., Sasai, N., Nagata, K., Liu, X.D., Liu, P.C., Thiele, D.J., and Nakai, A. (1999). The mammalian HSF4 gene generates both an activator and a repressor of heat shock genes by alternative splicing. *J Biol Chem* *274*, 27845-27856. 10.1074/jbc.274.39.27845.

Taylor, J.S., and Raes, J. (2004). Duplication and divergence: the evolution of new genes and old ideas. *Annu Rev Genet* *38*, 615-643. 10.1146/annurev.genet.38.072902.092831.

Teichmann, S.A., and Babu, M.M. (2004). Gene regulatory network growth by duplication. *Nature Genetics* *36*, 492-496. 10.1038/ng1340.

Teves, S.S., and Henikoff, S. (2013). The heat shock response: A case study of chromatin dynamics in gene regulation. *Biochem Cell Biol* *91*, 42-48. 10.1139/bcb-2012-0075.

Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a Cancer Dependency Map. *Cell* *170*, 564-576 e516. 10.1016/j.cell.2017.06.010.

Vihervaara, A., Duarte, F.M., and Lis, J.T. (2018). Molecular mechanisms driving transcriptional stress responses. *Nat Rev Genet* *19*, 385-397. 10.1038/s41576-018-0001-6.

Vihervaara, A., Mahat, D.B., Guertin, M.J., Chu, T., Danko, C.G., Lis, J.T., and Sistonon, L. (2017). Transcriptional response to stress is pre-wired by promoter and enhancer architecture. *Nat Commun* *8*, 255. 10.1038/s41467-017-00151-0.

Vihervaara, A., Sergelius, C., Vasara, J., Blom, M.A., Elsing, A.N., Roos-Mattjus, P., and Sistonon, L. (2013). Transcriptional response to stress in the dynamic chromatin environment of cycling and mitotic cells. *Proc Natl Acad Sci U S A* *110*, E3388-3397. 10.1073/pnas.1305275110.

- Vydra, N., Janus, P., Toma-Jonik, A., Stokowy, T., Mrowiec, K., Korfanty, J., Dlugajczyk, A., Wojtas, B., Gielniewski, B., and Widlak, W. (2019). 17 beta-Estradiol Activates HSF1 via MAPK Signaling in ER alpha-Positive Breast Cancer Cells. *Cancers* *11*. 10.3390/cancers11101533.
- Wang, G., Ying, Z., Jin, X., Tu, N., Zhang, Y., Phillips, M., Moskophidis, D., and Mivechi, N.F. (2004). Essential requirement for both hsf1 and hsf2 transcriptional activity in spermatogenesis and male fertility. *Genesis* *38*, 66-80. 10.1002/gene.20005.
- Wang, G.H., Zhang, J., Moskophidis, D., and Mivechi, N.F. (2003). Targeted disruption of the heat shock transcription factor (hsf)-2 gene results in increased embryonic lethality, neuronal defects, and reduced spermatogenesis. *Genesis* *36*, 48-61. 10.1002/gene.10200.
- Wang, T., Yu, H., Hughes, N.W., Liu, B., Kendirli, A., Klein, K., Chen, W.W., Lander, E.S., and Sabatini, D.M. (2017). Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras. *Cell* *168*, 890-903 e815. 10.1016/j.cell.2017.01.013.
- Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. *Nature* *449*, 54-U36. 10.1038/nature06107.
- Wessel, D., and Flugge, U.I. (1984). A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* *138*, 141-143. 10.1016/0003-2697(84)90782-6.
- Westerheide, S.D., Bosman, J.D., Mbadugha, B.N., Kawahara, T.L., Matsumoto, G., Kim, S., Gu, W., Devlin, J.P., Silverman, R.B., and Morimoto, R.I. (2004). Celastrols as inducers of the heat shock response and cytoprotection. *J Biol Chem* *279*, 56053-56060. 10.1074/jbc.M409267200.
- Whitesell, L., Santagata, S., Mendillo, M.L., Lin, N.U., Proia, D.A., and Lindquist, S. (2014). HSP90 empowers evolution of resistance to hormonal therapy in human breast cancer models. *Proc Natl Acad Sci U S A* *111*, 18297-18302. 10.1073/pnas.1421323111.
- Widlak, W., and Vydra, N. (2017). The Role of Heat Shock Factors in Mammalian Spermatogenesis. *Adv Anat Embryol Cell Biol* *222*, 45-65. 10.1007/978-3-319-51409-3_3.
- Wu, C. (1984). Activating protein factor binds in vitro to upstream control sequences in heat shock gene chromatin. *Nature* *311*, 81-84. 10.1038/311081a0.

Wu, C. (1995). Heat shock transcription factors: Structure and regulation. *Annu Rev Cell Dev Bi* 11, 441-469. DOI 10.1146/annurev.cb.11.110195.002301.

Xi, C.X., Hu, Y.Z., Buckhaults, P., Moskophidis, D., and Mivechi, N.F. (2012). Heat Shock Factor Hsf1 Cooperates with ErbB2 (Her2/Neu) Protein to Promote Mammary Tumorigenesis and Metastasis. *Journal of Biological Chemistry* 287, 35646-35657. 10.1074/jbc.M112.377481.

Xiao, X.Z., Zuo, X.X., Davis, A.A., McMillan, D.R., Curry, B.B., Richardson, J.A., and Benjamin, I.J. (1999). HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. *Embo Journal* 18, 5943-5952. DOI 10.1093/emboj/18.21.5943.

Xing, H., Wilkerson, D.C., Mayhew, C.N., Lubert, E.J., Skaggs, H.S., Goodson, M.L., Hong, Y., Park-Sarge, O.K., and Sarge, K.D. (2005). Mechanism of hsp70i gene bookmarking. *Science* 307, 421-423. 10.1126/science.1106478.

Yang, Y., Zhou, Y., Xiong, X., Huang, M., Ying, X., and Wang, M. (2018). ALG3 Is Activated by Heat Shock Factor 2 and Promotes Breast Cancer Growth. *Med Sci Monit* 24, 3479-3487. 10.12659/MSM.907461.

Yuan, C.X., Czarnecka-Verner, E., and Gurley, W.B. (1997). Expression of human heat shock transcription factors 1 and 2 in HeLa cells and yeast. *Cell Stress Chaperones* 2, 263-275. 10.1379/1466-1268(1997)002<0263:eohhst>2.3.co;2.

Zhang, Y., Huang, L., Zhang, J., Moskophidis, D., and Mivechi, N.F. (2002). Targeted disruption of hsf1 leads to lack of thermotolerance and defines tissue-specific regulation for stress-inducible Hsp molecular chaperones. *Journal of Cellular Biochemistry* 86, 376-393. 10.1002/jcb.10232.

Zhong, Y.H., Cheng, H.Z., Peng, H., Tang, S.C., and Wang, P. (2016). Heat shock factor 2 is associated with the occurrence of lung cancer by enhancing the expression of heat shock proteins. *Oncol Lett* 12, 5106-5112. 10.3892/ol.2016.5368.

Zhu, Z., and Mivechi, N.F. (1999). Regulatory domain of human heat shock transcription Factor-2 is not regulated by hemin or heat shock. *Journal of Cellular Biochemistry* 73, 56-69. 10.1002/(sici)1097-4644(19990401)73:1<56::Aid-jcb7>3.0.Co;2-7.

Zou, Y., Zhu, W., Sakamoto, M., Qin, Y., Akazawa, H., Toko, H., Mizukami, M., Takeda, N., Minamino, T., Takano, H., et al. (2003). Heat shock transcription factor 1 protects cardiomyocytes from ischemia/reperfusion injury. *Circulation* *108*, 3024-3030.
10.1161/01.CIR.0000101923.54751.77.

Zuo, J., Baler, R., Dahl, G., and 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.
10.1128/mcb.14.11.7557-7568.1994.

VITA

EDUCATION

Jun 2016 – Present	Northwestern University, Feinberg School of Medicine , Chicago, IL MD-PhD Candidate , Medical Scientist Training Program, 2024 USMLE Step 1: 257, ~94 th percentile; Driskill Graduate Program GPA: 4.0
Sep 2010 – May 2014	University of Notre Dame , Notre Dame, IN Bachelor of Science, Biological Sciences Cumulative GPA: 3.853
Spring 2013	Notre Dame International Study Abroad: Fundación Ortega y Gasset , Toledo, Spain. GPA: 4.0

RESEARCH EXPERIENCE

Jun 2018 – Jun 2022	Graduate Student, Doctor of Philosophy, Driskill Graduate Program in Life Sciences Dr. Marc Mendillo Lab, Feinberg School of Medicine, Northwestern University, Chicago, IL. We are investigating the role of heat shock factors in cancer, specifically how their role in maintaining protein homeostasis is hijacked by cancer cells to support rapid growth and resistance to therapies.
Jun 2014 – May 2016 Summer 2013	Research Technician; Summer Undergraduate Researcher Dr. Marc Ladanyi Lab, Memorial Sloan Kettering Cancer Center, New York, NY. I developed and characterized models of rare, oncogenic fusion-driven lung cancers and pediatric sarcomas using primary patient samples and CRISPR/Cas9 gene editing. These models were instrumental for subsequent studies identifying biological vulnerabilities of these tumors and validating targets for therapy.
Jan 2011 – May 2014	Undergraduate Research Student Dr. Michael T. Ferdig Lab, University of Notre Dame, Notre Dame, IN. Characterized <i>Plasmodium falciparum</i> transcriptional response to diverse small molecules to interpret mechanism of action and identify novel therapeutic opportunities. Honors Thesis: “Validating small molecule responses in the <i>Plasmodium falciparum</i> transcriptome.”

Programming Languages: Python | R | Unix/Linux

LEADERSHIP

Mar 2020 – Jun 2021	Co-Founder, GetMePPE Chicago. The GetMePPEChicago team has been assessing PPE needs and coordinating donations to healthcare and community organizations in Chicago. We have donated over 800,000 pieces of PPE. Our efforts also included regional and national donations coordinated with key industry and manufacturing partners.
Aug 2018 – Aug 2019	President, Northwestern Medical Scientist Training Program Student Council.
Aug 2017 – Present	Chair, Feinberg School of Medicine Student IT Advisory Committee.
Oct 2018 – Present	Leadership Fellow, Center for Leadership, Northwestern University
Jul 2018 – Present	Student Member, Northwestern MSTP Admissions Committee
Jul 2017 – Present	Student Member, Feinberg School of Medicine Curriculum Review Committee
Aug 2017 – Jun 2018	Student Review Panel, Phase I Curriculum, Feinberg School of Medicine
Jul 2014 – May 2016	Member, MSKCC Green Team Committee
Aug 2013 – May 2014	Co-Chair, Senior Leadership Committee, Department of Biological Sciences
Jan 2012 – May 2014	Co-President, Spanish Club

HONORS and AWARDS

Jun 2022	The Graduate School Conference Travel Grant, for attending the 3 rd Cell Death, Cell Stress and Metabolism Conference, Cancún, Mexico
Jul 2020 - 2022	ARCS Scholar, Achievement Rewards for College Scientists, Chicago, IL
Jun 2019	Heller MSTP Travel Award, Gordon Research Conference Stress Proteins, 2020, Lucca, Italy
Nov 2018	Travel Award, LINCS CMap Workshop Series 2018, Boston, MA
May 2014	Magna cum laude, University of Notre Dame, Notre Dame, IN
May 2014	Outstanding Biological Scientist Award, Dept. of Biological Sciences, Univ. of Notre Dame, Notre Dame, IN
Summer 2013	Rubin and Sarah Shaps Scholar – for exceptional performance in the Gerstner Sloan Kettering Summer Undergraduate Research Program, New York, NY
Summer 2011, 2012	Summer Undergraduate Research Fellowship, College of Science, University of Notre Dame, Notre Dame, IN
2010 – 2014	Dean's Honor List, 7 of 8 semesters, University of Notre Dame, Notre Dame, IN

TEACHING EXPERIENCE

Fall 2019	<i>Teaching Assistant, Data Science for Biomedical Researchers, The Driskill Graduate Program, Feinberg School of Medicine</i>
Jul - Aug 2019	<i>Invited Teaching Assistant, Cold Spring Harbor Labs Gene Expression and Epigenetics Course, 2019</i>

INVITED TALKS

Aug 2020	<i>Student-faculty panel on COVID-19 activities, Feinberg School of Medicine, Panelist</i>
Aug 2018	<i>Modes of Learning, Feinberg School of Medicine, Panelist</i>

PROFESSIONAL MEMBERSHIPS

2016 - Present	Member , American Medical Association
2017 - Present	Member , American Physician Scientists Association
2017 – Present	Member , American Society of Clinical Oncology

PUBLICATIONS

1. Amici DR, Ansel D, Metz K, **Smith RS**, Phoumyvong C, Gayatri S, Srivastava S, Chamera T, Edwards S, Brockway S, Takagishi S, O'Hara B, Cho B, Ah Goo Y, Kelleher NL, Ben-Sahra I, Foltz DR, Li J, Mendillo M. C16orf72/HAPSTR1 is a molecular rheostat in an integrated network of stress response pathways. PNAS. 2022, In press.
2. **Smith RS***, Takagishi SR*, Amici DR, Metz K, Gayatri S, Alasady MJ, Wu Y, Brockway S, Taiberg SL, Khalatyan N, Taipale M, Santagata S, Whitesell L, Lindquist S, Savas JN, Mendillo ML. HSF2 cooperates with HSF1 to drive a transcriptional program critical for the malignant state. Sci Adv. 2022;8(11):eabj6526. doi: 10.1126/sciadv.abj6526. *co-first authors.
3. **Smith RS***, Odintsov I*, Zebing L, Lui A, Hayashi T, Vojnic M, Suehara Y, Delasos L, Mattar MS, Hmeljak J, Ramirez HA, Shaw M, Bui G, Hartono AB, Gladstone E, Kunte S, Magnan H, Khodos I, De Stanchina E, La Quaglia MP, Yao J, Laé M, Lee SB, Spraggon L, Pratilas CA, Ladanyi M, Somwar R. Novel patient-derived models of desmoplastic small round cell tumor confirm a targetable dependency on ERBB signaling. Disease Models & Mechanisms. 2021 Nov 29; 15(1):dmm047621. doi: 10.1242/dmm.047621. PMID: 34841430. *co-first authors.
4. Merchant TS, Hormozian S, **Smith RS**, Pendergrast T, Siddiqui A, Wen Z, Sheldon M. Ethical principles in PPE inventory management decisions and partnerships across state lines. Public Health Reports. 2021 Dec 30. doi:10.1177/00333549211058734. PMID: 34969322.

5. Clos ALT, Cohen AP, Edwards LE, Martin AHH, Merchant TS, Pendergrast TR, Siegel MA, **Smith RS**. Ethical Considerations in PPE Allocation During COVID-19: A Case Study. *Harvard Medical Student Review*. 2021 Mar 19. <https://www.hmsreview.org/covid/ethical-considerations-in-ppe-allocation>.
6. Amici DR, Jackson JM, Truica M, **Smith RS**, Abdulkadir SA, Mendillo ML. FIREWORKS: a bottom-up approach to integrative coessentiality network analysis. *Life Science Alliance*. 2020 Dec 16;4(2):e202000882. doi: 10.26508/lsa.202000882. PMID: 33328249.
7. Hayashi T*, Odintsov I*, **Smith RS***, Ishizawa K, Liu AJW, Delasos L, Kurzatkowski C, Tai H, Gladstone E, Vojnic M, Kohsaka S, Suzawa K, Liu Z, Kunte S, Mattar MS, Khodos I, Davare MA, Drilon A, Cheng E, De Stanchina E, Ladanyi M, Somwar R. RET inhibition in novel patient-derived models of RET fusion-positive lung adenocarcinoma reveals a role for MYC upregulation. *Disease Models & Mechanisms*. 2020 Dec 14;14(2):dmm047779. doi: 10.1242/dmm.047779. PMID: 33318047. *co-first authors.
8. Suzawa K, Offin M, Lu D, Kurzatkowski C, Vojnic M, **Smith RS**, Sabari JK, Tai H, Mattar M, Khodos I, De Stanchina E, Rudin CM, Kris MG, Arcila ME, Lockwood WW, Drilon A, Ladanyi M, Somwar R. Activation of KRAS mediates resistance to targeted therapy in MET exon 14 mutant non-small cell lung cancer. *Clinical Cancer Research*. 2019 Feb 15. doi: 10.1158/1078-0432.CCR-18-1640. PubMed PMID: 30352902.
9. Drilon A, Somwar R, Mangatt BP, Edgren H, Desmeules P, Ruusulehto A, **Smith RS**, Delasos L, Vojnic M, Plodkowski AJ, Sabari J, Ng K, Montecalvo J, Chang J, Tai H, Lockwood WW, Martinez V, Riely GJ, Rudin CM, Kris MG, Arcila ME, Matheny C, Benayed R, Rekhtman N, Ladanyi M, Ganji G. Response to ERBB3-Directed Targeted Therapy in *NRG1*-Rearranged Cancers. *Cancer Discovery*. 2018 Apr 2. doi: 10.1158/2159-8290.CD-17-1004. PubMed PMID: 29610121.
10. Hayashi T, Desmeules P, **Smith RS**, Drilon A, Somwar R, Ladanyi M. *RASAI* and *NFI* are Preferentially Co-Mutated and Define A Distinct Genetic Subset of Smoking-Associated Non-Small Cell Lung Carcinomas Sensitive to MEK Inhibition. *Clinical Cancer Research*. 2018 Mar 15;24(6):1436-1447. doi: 10.1158/1078-0432.CCR-17-2343. Epub 2017 Nov 10. PubMed PMID: 29127119.
11. Li GG, Somwar R, Joseph J, **Smith RS**, Hayashi T, Martin L, Franovic A, Schairer A, Martin ES, Riely GJ, Harris J, Yan S, Wei G, Oliver J, Patel R, Multani P, Ladanyi M, Drilon A. Antitumor Activity of RXDX-105 in Multiple Cancer Types with RET Rearrangements or Mutations. *Clin Cancer Res*. 2016 Dec 23. pii: clincanres.1887.2016. doi: 10.1158/1078-0432.CCR-16-1887. PubMed PMID: 28011461.
12. Drilon A, Rekhtman N, Arcila M, Wang L, Ni A, Albano M, Van Voorthuysen M, Somwar R, **Smith RS**, Montecalvo J, Plodkowski A, Ginsberg MS, Riely GJ, Rudin CM, Ladanyi M,

- Kris MG. Cabozantinib in patients with advanced RET-rearranged non-small-cell lung cancer: an open-label, single-centre, phase 2, single-arm trial. *Lancet Oncol.* 2016 Dec;17(12):1653-1660. doi: 10.1016/S1470-2045(16)30562-9. PubMed PMID: 27825636; PubMed Central PMCID: PMC5143197.
13. Shukla N*, Somwar R*, **Smith RS***, Ambati S, Munoz S, Merchant M, D'Arcy P, Wang X, Kobos R, Antczak C, Bhinder B, Shum D, Radu C, Yang G, Taylor BS, Ng CK, Weigelt B, Khodos I, de Stanchina E, Reis-Filho JS, Ouerfelli O, Linder S, Djaballah H, Ladanyi M. Proteasome Addiction Defined in Ewing Sarcoma Is Effectively Targeted by a Novel Class of 19S Proteasome Inhibitors. *Cancer Res.* 2016 Aug 1;76(15):4525-34. doi: 10.1158/0008-5472.CAN-16-1040. PubMed PMID: 27256563. *co-first authors
 14. Drilon A, Bergagnini I, Delasos L, Sabari J, Woo KM, Plodkowski A, Wang L, Hellmann MD, Joubert P, Sima CS, **Smith R**, Somwar R, Rekhman N, Ladanyi M, Riely GJ, Kris MG. Clinical outcomes with pemetrexed-based systemic therapies in RET-rearranged lung cancers. *Annals of Oncology.* 2016 Jul;27(7):1286-91. doi: 10.1093/annonc/mdw163. PubMed PMID: 27056998; PubMed Central PMCID: PMC4922319.
 15. Wang D, Narula N, Azzopardi S, **Smith RS**, Nasar A, Altorki NK, Mittal V, Somwar R, Stiles BM, Du YN. Expression of the receptor for hyaluronic acid mediated motility (RHAMM) is associated with poor prognosis and metastasis in non-small cell lung carcinoma. *Oncotarget.* 2016 Jun 28;7(26):39957-39969. doi: 10.18632/oncotarget.9554. PubMed PMID: 27220886; PubMed Central PMCID: PMC5129984.
 16. Drilon A, Somwar R, Wagner JP, Vellore NA, Eide CA, Zabriskie MS, Arcila ME, Hechtman JF, Wang L, **Smith RS**, Kris MG, Riely GJ, Druker BJ, O'Hare T, Ladanyi M, Davare MA. A Novel Crizotinib-Resistant Solvent-Front Mutation Responsive to Cabozantinib Therapy in a Patient with ROS1-Rearranged Lung Cancer. *Clinical Cancer Research.* 2016 May 15;22(10):2351-8. doi: 10.1158/1078-0432.CCR-15-2013. PubMed PMID: 26673800; PubMed Central PMCID: PMC4867287.
 17. Siwo GH, **Smith RS**, Tan A, Button-Simons KA, Checkley LA, Ferdig MT. An integrative analysis of small molecule transcriptional responses in the human malaria parasite *Plasmodium falciparum*. *BMC Genomics.* 2015 Dec 4;16:1030. doi: 10.1186/s12864-015-2165-1. PubMed PMID: 26637195; PubMed Central PMCID: PMC4670519.

ABSTRACTS/ PRESENTATIONS/ CONFERENCE PROCEEDINGS

1. **HSF2 cooperates with HSF1 to drive a transcriptional program critical for the malignant state.** **Smith RS***, Takagishi SR*, Amici DR, Metz K, Gayatri S, Alasady MJ, Wu Y, Brockway S, Taiberg SL, Khalatyan N, Taipale M, Santagata S, Whitesell L,

Lindquist S, Savas JN, Mendillo, ML. 3rd Cell Death, Cell Stress and Metabolism Conference. May 7 – 10, 2022. Cancún, Mexico. Oral presentation

2. **HSF2 cooperates with HSF1 to drive a transcriptional program critical for the malignant state.** Smith RS*, Takagishi SR*, Amici DR, Metz K, Gayatri S, Alasady MJ, Wu Y, Brockway S, Taiberg SL, Khalatyan N, Taipale M, Santagata S, Whitesell L, Lindquist S, Savas JN, Mendillo, ML. American Physician Scientists Association Annual Meeting. April 8 - 10, 2022. Chicago, IL. Poster presentation.
3. **The Role of Heat Shock Factor 2 in Human Malignancies.** Smith RS, Takagishi S, Wu Y, Metz K, Alasady M, Gayatri S, Mendillo ML. Gordon Research Conference: Stress Proteins in Growth, Development and Disease. June 23 – 28, 2019. Lucca, Italy. Poster presentation.
4. **The role of heat shock factor 2 in human malignancies.** Seesha Takagishi, R.S. Smith (presenter), Yaqi Wu, Kyle Metz, Milad Alasady, Marc Mendillo. 24th Annual Midwest Stress Response and Molecular Chaperone Meeting January 19, 2019. Evanston, IL. Oral presentation.
5. **Role of ERBB signaling in RET-rearranged lung cancer and contribution of EGFR amplification to cabozantinib resistance.** R.S. Smith, A.E. Drilon, S. Kunte, K. Suzawa, T. Hayashi, L. Delasos, H. Tai, T. Hitchman, I. Khodos, M. Mattar, S. Kohsaka, E. de Stanchina, W. Lockwood, M. Ladanyi, R. Somwar. ASCO June 3, 2017. Chicago, IL. Poster presentation.
6. **EGFR is a Potential Therapeutic Target for Desmoplastic Small Round Cell Tumor.** R.S. Smith, R. Somwar, M. Shaw, G. Bui, L. Spraggon, H. Ramirez, I. Khodos, M. Mattar, E. DeStanchina, M. LaQuaglia, H. Magnan, C. Pratilas and M. Ladanyi. *Anderson Cancer Symposium*; September 18, 2015; Rockefeller University, New York, NY. Poster presentation.
7. **Validating Small Molecule Responses in *Plasmodium falciparum*.** R.S. Smith, G.H. Siwo, K.A. Button-Simons, A. Tan, L. Checkley, M.T. Ferdig. Oral presentation, *College of Science Joint Annual Meeting*; May 2, 2014; University of Notre Dame, Notre Dame, IN. Poster presentation.
8. **A Genomics Platform for Anti-Malarial Drug Discovery.** R.S. Smith, G.H. Siwo, A. Tan, L. Checkley, M.T. Ferdig. *Fall Undergraduate Research Fair*; October 31, 2013; University of Notre Dame, IN. Poster presentation.
9. **Preclinical Characterization of a New Potential Targeted Therapy for Ewing's Sarcoma.** R.S. Smith, R. Somwar, N. Shukla, M. Ladanyi. *Gerstner Sloan Kettering SURP Poster Session*; August 8, 2013; Memorial Sloan Kettering Cancer Center, NY, NY. Poster presentation.

10. **Network Biology: A Tool for Understanding Drug Mechanism of Action in the Malaria Parasite.** R.S. Smith, G.H. Siwo, L. Checkley, A. Tan, M.T. Ferdig. *American Society of Tropical Medicine and Hygiene 61st Annual Meeting*; November 11-15, 2012; Atlanta, GA.
11. **Predicting Antimalarial Mode of Action from Gene Expression Signatures.** R.S. Smith, G.H. Siwo, A. Tan, M.T. Ferdig. Poster presentation.
 - a. Notre Dame alumni-outreach fundraising event; April 26, 2012; Lake Forest, IL.
 - b. *National Collegiate Research Conference* presented by the Harvard College Undergraduate Research Association; January 19-21, 2012; Harvard University, Cambridge, MA.
 - c. *Fall Undergraduate Research Fair*; October 27, 2011; University of Notre Dame, Notre Dame, IN.
 - d. *Transforming the Next Generation of Notre Dame Medical Doctors Through Research and Discovery*; October 12, 2011; Santa Fe Building, Chicago, IL.