Choosing the target loci:

Heat shock factors HSF1 and HSF2 as regulators of cell stress and development

Anniina Vihervaara

Department of Biosciences, Åbo Akademi University
Turku Centre for Biotechnology, University of Turku and Åbo Akademi University
Turku Doctoral Program of Biomedical Sciences
Finland

2014

Choosing the target loci:

Heat shock factors HSF1 and HSF2 as regulators of cell stress and development

Anniina Vihervaara

Department of Biosciences, Åbo Akademi University
Turku Centre for Biotechnology, University of Turku and Åbo Akademi University
Turku Doctoral Program of Biomedical Sciences
Finland

2014

From the Department of Biosciences, Åbo Akademi University, Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, and Turku Doctoral Program of Biomedical Sciences.

Supervised by

Professor Lea Sistonen, PhD
Department of Biosciences, Åbo Akademi University
and
Turku Centre for Biotechnology, University of Turku and Åbo Akademi University
Turku, Finland

Reviewed by

Docent Maria Vartiainen, PhD
Institute of Biotechnology, University of Helsinki
Helsinki, Finland

Professor Jorma Palvimo, PhD
Institute of Biomedicine, University of Eastern Finland
Kuopio, Finland

Opponent

Professor John T. Lis, PhD
Cornell University, Department of Molecular Biology and Genetics,
Ithaca, New York, U.S.A.

ISBN 978-952-12-3080-6
Painosalama Oy – Turku, Finland 2014

For the curious minds,
who seek for answers
and find questions

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This PhD thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I-III). In addition, unpublished results are included. The original publications have been reproduced with permission of the copyright holders.

I. Åkerfelt M*, Henriksson E*, Laiho A, Vihervaara A, Rautoma K, Kotaja N, Sistonen L. (2008). Promoter ChIP-chip analysis in mouse testis reveals Y chromosome occupancy by HSF2. Proc Natl Acad Sci U S A **105**: 11224-11229. *equal contribution.

II. Åkerfelt M*, Vihervaara A*, Laiho A, Conter A, Christians E, Sistonen L[¥], Henriksson E[¥]. (2010). Heat shock transcription factor 1 localizes to sex chromatin during meiotic repression. J Biol Chem **285**: 34469-34476. *[¥]equal contribution.

III. Vihervaara A, Sergelius C, Vasara J, Blom MAH, Elsing AN, Roos-Mattjus P, Sistonen 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.

ABBREVIATIONS

ACRV1	Acrosomal vesicle protein 1
AD	Activation domain
AHA1/AHSA1	Activator of heat shock 90kDa protein ATPase
ARHGEF	Rho guanine nucleotide exchange factor
BANF1/BAF1	Barrier to autointegration factor 1
bp	Base pair
CCT	Chaperonin containing TCP
CDC37	Cell division cycle 37
CDK	Cyclin-dependent kinase
ChIP	Chromatin immunoprecipitation
ChIP-chip	Chromatin immunoprecipitation on microarray
ChIP-seq	Chromatin immunoprecipitation coupled to high-throughput sequencing
CHRNA	Cholinergic receptor, nicotinic, alpha
CTCF	CCCTC-binding factor
CTD	Carboxyterminal domain of RNA polymerase II
DARS	Aspartyl-tRNA synthetase
DBD	DNA-binding domain
DNAJ	J-domain containing chaperone
DTT	Dithiotreitol
DZ	Dark zone of mouse seminiferous tubule
eEF	Eukaryote elongation factor
ENCODE	Encyclopedia of DNA elements
FTMT	Mitochondrial ferritin
GAF	GAGA-binding factor
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GBA	Glucosidase beta, acid
GEO	Gene expression omnibus
γH2AX	Histone 2A X, phosphorylated at serine (S) 139
HAT	Histone acetyl transferase
HDAC	Histone deacetylase
HOX	Homeobox protein
HR-A/B/C	Heptad repeat A/B/C
HSE	Heat shock element
HSF	Heat shock factor
HSP	Heat shock protein
HSR	Heat shock response
IL	Interleukin
KCNN1	Calcium-activated channel N1
LDLR	Low density lipoprotein receptor
LMNB	Lamin, beta
lncRNA/LINC	Long non-coding RNA
ME1	Malic enzyme 1, NADP(+)-dependent, cytosolic
miRNA	MicroRNA

MLL/KMT2A	Myeloid/lymphoid or mixed-lineage leukemia/Lysine (K)-specific methyl-transferase 2A
mRNA	Messenger RNA
MSCI	Meiotic sex-chromosome inactivation
MRPS	Mitochondrial ribosomal protein
NEAT1	Nuclear paraspeckle assembly transcript 1
NLS	Nuclear localization signal
NUDC	Nuclear distribution C homolog
PAR	Pseudoautosomal region
PcG	Polycomb group protein
PDSM	Phosphorylation-dependent sumoylation motif
PFN3	Profilin 3
PIC	Preinitiation complex
PRKC/PKC	Protein kinase C
PRM	Protamine
PTGES3/p23	Prostaglandin E synthase
PTM	Post-translational modification
PZ	Pale zone of mouse seminiferous tubule
RD	Regulatory domain
RNAi	RNA interference
RNP	RNA polymerase
rRNA	Ribosomal RNA
shRNA	Short hairpin RNA
SLX	Sycp3 like X-linked
SLY	Sycp3 like Y-linked
SNAP29	Synaptonemal-associated protein 29 kDa
SPEER	Spermatogenesis associated glutamate (E)-rich protein
SPT	Suppressor of Ty
SRSY	Serine-rich secreted, Y-linked
SRY	Sex determining region Y
SS	Strong spot of mouse seminiferous tubule
SSTY	Spermiogenesis-specific transcript on the Y
SWI/SNF	Switch/sucrose nonfermenting
TAF	TBP-associated factor
TBP	TATA-box binding protein
TF	General transcription factor
TNP	Transition protein
tRNA	Transfer RNA
trxG	Trithrorax protein
TSS	Transcriptional start site
TUBE	Tandem ubiquitin binding entity
Ub	Ubiquitin
WS	Weak spot of mouse seminiferous tubule
Wt	Wild type
Yp	Short arm of the mouse Y chromosome
Yq	Long arm of the mouse Y chromosome
ZFAND2A/AIRAP	Zinc finger AN1-type domain 2A/Arsenite-induced RNA associated

ABSTRACT

Heat shock factors (HSFs) are an evolutionarily well conserved family of transcription factors that coordinate stress-induced gene expression and direct versatile physiological processes in eukaryote organisms. The essentiality of HSFs for cellular homeostasis has been well demonstrated, mainly through HSF1-induced transcription of heat shock protein (HSP) genes. HSFs are important regulators of many fundamental processes such as gametogenesis, metabolic control and aging, and are involved in pathological conditions including cancer progression and neurodegenerative diseases. In each of the HSF-mediated processes, however, the detailed mechanisms of HSF family members and their complete set of target genes have remained unknown. Recently, rapid advances in chromatin studies have enabled genome-wide characterization of protein binding sites in a high resolution and in an unbiased manner. In this PhD thesis, these novel methods that base on chromatin immunoprecipitation (ChIP) are utilized and the genome-wide target loci for HSF1 and HSF2 are identified in cellular stress responses and in developmental processes. The thesis and its original publications characterize the individual and shared target genes of HSF1 and HSF2, describe HSF1 as a potent transactivator, and discover HSF2 as an epigenetic regulator that coordinates gene expression throughout the cell cycle progression. In male gametogenesis, novel physiological functions for HSF1 and HSF2 are revealed and HSFs are demonstrated to control the expression of X- and Y-chromosomal multicopy genes in a silenced chromatin environment. In stressed human cells, HSF1 and HSF2 are shown to coordinate the expression of a wide variety of genes including genes for chaperone machinery, ubiquitin, regulators of cell cycle progression and signaling. These results highlight the importance of cell type and cell cycle phase in transcriptional responses, reveal the myriad of processes that are adjusted in a stressed cell and describe novel mechanisms that maintain transcriptional memory in mitotic cell division.

1 INTRODUCTION

The cell is the smallest entity of life and the structural and functional foundation of all organisms. To understand nature and the mechanisms of life, a comprehensive appreciation of cellular processes and cellular networks in complex organisms is required. Inside each cell, hereditary information carries the instructions for differentiation, cell type-specific functions and division of labor between tissues. Coordinated execution of cellular processes involves sensing external and internal conditions, organizing signaling cascades and changing the cell's structure or behavior. Signals that reach the genome adjust the expression of genes, the instruction entities for synthesis of cellular components. Consequently, the gene expression programs in individual cells determine the molecular constituents and the possibilities for cellular responses, making the regulation of gene expression one of the most fundamental processes in all living organisms. In this PhD thesis, the coordination of gene expression is investigated in development and in response to acute, protein-damaging stress.

Heat shock factors (HSFs) are an evolutionarily well conserved protein family that coordinates gene expression in a variety of physiological processes. HSFs are best characterized as rapid activators of gene expression upon protein-damaging stress when the overall gene activity in the cell is silenced. Beyond stress, HSFs are crucial regulators of developmental processes and aging, and involved in several pathological conditions such as neurodegenerative diseases and cancer. In this thesis, I have investigated the versatile roles of HSF1 and HSF2 in development and in cellular stress responses, addressing how HSFs interact with the dynamic chromatin environment in different cell types and cellular conditions. As a model system for development, I have used male gametogenesis which consists of strikingly complex and well coordinated changes in the chromatin landscape. The developing gametes undergo clonal expansion via mitosis, reorganization and reduction division of chromosomes during meiosis and a profound morphological differentiation during the haploid phase of spermatogenesis. The rapidly provoked gene expression in response to stress provides a model system where the molecular mechanisms of HSF-mediated transcriptional activation and the cellular processes that maintain homeostasis can be studied. To elucidate how HSF1 and HSF2 interact with chromatin in distinct states, I have investigated the transcriptional programming and the cellular survival mechanisms in freely cycling cells and in cells that undergo mitotic division. In each of these studies, the genome-wide target sites for HSF1 and HSF2 have been characterized using advanced techniques that base on chromatin immunoprecipitation (ChIP). The genome-wide analyses have been elaborated with computational data mining, biochemical characterization of gene-specific regulatory mechanisms for HSF1 and HSF2 and by investigating the biological significance of the transcriptional reprogramming.

The results presented in this thesis and its original publications reveal the importance of the cell type, developmental state and the chromatin environment for transcriptional responses. In male gametogenesis, HSFs are uncovered to control the X- and Y-chromosomal multicopy genes, which are crucial for chromatin compaction in the sperm head and for correct sperm morphology. HSF1 is shown to localize to silenced X- and Y-chromatin in pre- and post-meiotic germ cells whereas HSF2 resides at the dividing genome in meiosis I and II. To date, HSFs remain the only

transcription factors that have been shown to occupy the meiotic sex-chromatin and to regulate the expression of the sex-linked multicopy genes. In cellular stress responses, HSFs are revealed to coordinate the expression of whole chaperone machinery, including protein foldases, disaggregases, inhibitors of aggregate formation and cochaperones. Furthermore, HSFs induce the expression of ubiquitin and regulate the expression of translational components, mediators of cell cycle progression, metabolic processes and signaling cascades. The highly intertwined functions of HSF1 and HSF2 are contrasted with their profoundly distinct mechanisms on the genome, particularly in dividing cells. While HSF1 is efficiently excluded from the mitotic and meiotic chomatin, HSF2 avidly interacts with the condensed genome. HSF2, however, does not compensate for the lack of HSF1 at the stress responsive genes, but instead, is involved in reactivationt of transcription in post-mitotic cells.

2 REVIEW OF THE LITERATURE

2.1 Transcriptional regulation of cellular functions

Faithful implementation and propagation of hereditary information is a prerequisite for the execution of biological processes and development of eukaryotic organisms. How the genome is utilized in individual cells dictates the synthesis of cellular components and coordinates differentiation, cell type-specific functions and cellular responses. In a process called transcription, the genetic information is used as a template to build structural and functional molecules of the cell. The delicate coordination of transcriptional programs is essential for physiological processes and is coordinated by regulatory factors that organize the genetic information and direct the gene expression in response to internal and external stimuli.

2.1.1 Chromatin structure and dynamics

The hereditary information is encoded in the sequence of complementary deoxyribose nucleic acid (DNA) polymers that form a double helix (Figure 1). In eukaryotes, the DNA is organized inside a membrane enclosed nuclear compartment and associated with proteins to form a structure called chromatin (reviewed in Felsenfeld and Groudine, 2003; Schlick *et al*., 2012). The basal constituents of the chromatin are nucleosomes in which 146 base pairs (bps) of DNA encircle a histone octamer. The octamer, in turn, is composed of two sets of histones H2A, H2B, H3 and H4, and the nucleosome structure is stabilized by histone H1 that contacts the DNA at the site where it enters and exits the histone core (Figure 1). The nucleosomes are connected with linker DNA sequences and the chromatin fiber is further organized by scaffold proteins into a higher-order structure (Figure 1; reviewed in Woodcock and Ghosh, 2010).

Organization of genetic material into chromatin enables efficient condensation of the DNA, but also dynamic regulation of the accessibility of distinct genomic regions to transcriptional regulators, replication factors and DNA repair machinery. As a result of different grades of packaging and associating factors, the molecular composition of the chromatin fiber is highly diverse along the length of the chromosomes, making chromosomes among the most complex entities in the cell. The first layer of chromatin condensation is conducted by the histone molecules, which can occur in different variants and undergo post-translational modifications (PTMs) (reviewed in Zentner and Henikoff, 2013). The histone proteins associate with each other via globular core domains, whereas their highly dynamic tails can undergo a wide range of modifications including acetylation, methylation, phosphorylation, ubiquitination, sumoylation and ribosylation. The combination of the histone modifications affects the compactness of the chromatin and profoundly contributes to the accessibility of the underlying DNA (reviewed in Felsenfeld and Groudine, 2003; Zentner and Henikoff, 2013). The histone modifications are catalyzed by enzymes such as acetyl transferases (HATs), deacetylaces (HDACs), methyl transferases, demethylases, ligases and proteases (Brownell and Allis, 1996; Peterson and Laniel, 2004; Shilatifard, 2006). The chromatin is, furthermore, targeted by remodeling factors and architectural proteins that can change the position of histones or organize the genome into a more

or less compacted higher-order structure (reviewed in Woodcock and Ghosh, 2010). In general, chromatin regions that are actively encoded are kept within an open conformation (euchromatin) and are marked by hyperacetylation of histones H3 and H4, as well as by trimethylation of histone H3 at lysine (K) 4 (H3K4me3). Instead, silent chromatin regions reside in a closed state (heterochromatin) and are typically characterized by H3K9me3 and H3K27me3, and heterochromatin-associated proteins (Zentner and Henikoff, 2013). In non-dividing cells, the DNA is relatively loosely packed and display region-specific patterning of condensation. However, at every cell division, the chromatin undergoes profound condensation as the nuclear membrane breaks down and the duplicated metaphase chromosomes are separated into the daughter cells (Figure 1).

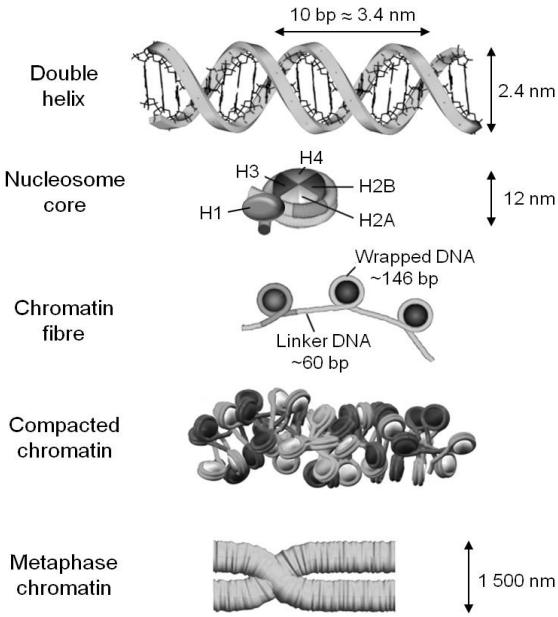


Figure 1. Chromatin structure. DNA double helix is wrapped around a histone octamer and the chromatin fibre is further organized into a higher-order structure. Adapted from Schlick *et al.*, *J. Biol. Chem.* 2012, reprinted with permission from ASBMB.

2.1.2 Organization of the genetic information

Information in the DNA is transcribed into single-stranded sequences of ribonucleic acid (RNA) polymers. Several RNA species exist in the cells, carrying versatile functions. Some of the RNA molecules contain information that is translated into amino acid sequence of proteins, and together, the RNA and protein molecules synthesize other cellular components such as lipids and define the structural and functional constituents of the cell. Traditionally, the sequences of DNA that code for proteins are termed genes, but genome-wide analyses have revealed the multitude of genomic regions that encode RNA as an end product (reviewed by Pennisi, 2012; Qu and Fang, 2013). In this thesis, the term "gene" refers to a DNA sequence that encodes a transcript. When specification is needed, "protein-coding" and "RNA-coding" define the end product of the gene. In human, the 20 687 protein-coding genes identified to date are sparsely distributed along the chromosomes and, due to the many non-coding sequences inside and between the genes, less than 3% of the human genome codes for amino acid sequences (Encyclopedia of DNA Elements, ENCODE, Consortium, 2013). The RNA-coding genes include a variety of small RNA species as well as 18 400 genes for RNA with over 200 bp. Consequently, recent estimates suggest that the majority of the human genome (~76%) is transcribed (The ENCODE Consortium, 2013; reviewed by Pennisi, 2012; Qu and Fang, 2013). A major challenge, however, for interpretation of the current genome-wide analyses is the presence of a number of repetitive sequences in the human genome, which routinely are neglegled from the sequence-based analyses (de Koning *et*

al., 2011; reviewed in Treangen and Salzberg, 2011). It is also worth noting that the genome is complex and certain loci encode several transcripts. Moreover, an RNA molecule might carry a regulatory or structural function besides serving as a template for an amino acid sequence (Salmena *et al*., 2011; Taulli *et al*., 2013).

The genome contains a myriad of non-coding sequences including regulatory elements that contribute to the coordinated expression of the genes (Figure 2; reviewed by Dekker *et al*., 2013; Maston *et al*., 2006; Riethoven, 2010). Core and control promoters locate to the vicinity of transcriptional start site (TSS) and serve as an assembly platform for the protein complexes that synthesize RNA. Enhancers, silencers and insulators are so called distal regulatory elements that can occur either upstream or downstream of the gene and regulate the gene activity also from a great distance (Figure 2; reviewed by Maston *et al*., 2006; Riethoven, 2010). Both enhancers and silencers contain binding sites for activatory and inhibitory factors, and typically organize the spatio-temporal gene expression in different tissues or in response to distinct stimuli. While insulators are able to confine the gene activity and chromatin landscape to a given region, locus control elements (LCRs) contribute to the coordinated activity of an entire locus or a gene cluster (Maston *et al*., 2006; Riethoven, 2010). The genome also contains information for its three dimensional organization (reviewed by Dekker *et al*., 2013; Gibcus and Dekker, 2013). For example, certain regions are targeted by lamins, which are the primary protein constituents of the nuclear envelope and can position genomic regions to the vicinity of the nuclear membrane (reviewed in Andrés and González, 2009). Importantly, the lamins are in direct contact with nuclear actin (Holaska *et al*., 2004; Simon *et al*., 2010) which has been coupled to gene expression, chromatin remodeling and processing of transcripts (reviewed in Akhtar and Gasser, 2007; Grosse and Vartiainen, 2013; de Lanerolle and Serebryannyy, 2011). A line of evidence also points to formation of transcription and replication factories, as several genomic regions come together to utilize the same machinery for a cellular process (Iborra *et al*., 1996; Mitchell and Fraser, 2008; Osborne *et al*., 2004; reviewed in Cook, 1999; Pope *et al*., 2013; Sutherland and Bickmore, 2009). All-in-all, the organization of the genome in space and time is a multi-dimensional task where a number of *cis*-acting DNA sequence elements and *trans*-acting factors interact for coordinated implementation of the hereditary information.

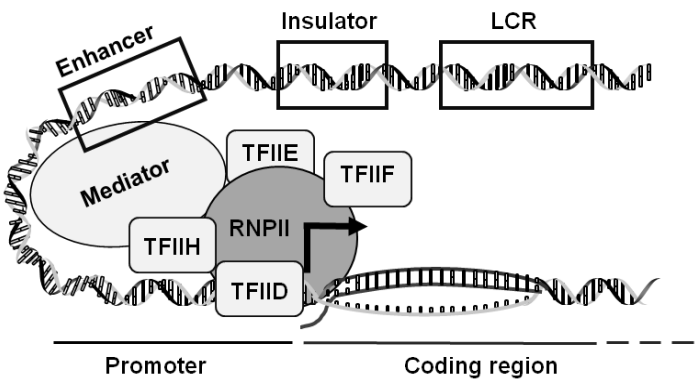


Figure 2. Control regions at an RNA polymerase II (RNPII) transcribed gene. General transcription factors (TFIIs) recognize DNA elements at the promoter and direct the binding and assembly of RNPII at the transcriptional start site (arrow). Transcriptional regulators, including the mediator complex, relay signals from distal control regions such as enhancers, insulators and locus control elements (LCRs) to the transcriptional machinery. Transcription is also controlled by site-specific transcription factors, cofactors and chromatin modifying enzymes (not shown).

2.1.3 Initiation of transcription

Expression of RNA and protein is regulated at multiple steps including transcriptional initiation, elongation and termination, as well as the stability, processing and localization of the produced RNA and protein. Since correct assembly of the transcriptional apparatus at the right genes at the right time is a prerequisite for correct gene expression, the initiation of transcription is a key regulatory step in the coordination of genome-wide transcriptional programming. Gene promoters harbor well conserved DNA elements such as TATA-box, GC-box, BRE-element or initiator that are recognized by general transcription factors (TFs) (reviewed in Kadonaga, 2012). The transcriptional activation is initiated by an ordered assembly of TFs at the promoter to form a preinitiation complex (PIC). PIC, in turn, directs the binding and correct positioning of RNA polymerase (RNP), the enzyme that catalyzes the synthesis of RNA using DNA as a template (Figure 2; reviewed in Cramer *et al*., 2008; Maston *et al*., 2006). In eukaryotes, three RNPs (RNPI-III) exist, each of which associates with a distinct set of TFs (TFI-III, respectively) and encodes a specified set of genes (reviewed in Hamperl *et al*., 2013; Vannini and Cramer, 2012). RNPI catalyzes the synthesis of 45S ribosomal RNA (rRNA) that is processed into 5.8S, 18S and 28S structural elements of protein translating ribosomes. Instead, RNPIII catalyzes S5 rRNA and all the transfer RNA (tRNA) species that interact with ribosomes and recruit amino acids to the growing polypeptide chain. Transcription that is mediated by RNPI and RNPIII localizes to nucleoli, which are subnuclear compartments that contain clustered ribosomal gene copies and, therefore, are prominent transcription-organizing structures in the eukaryote cell (reviewed in Gibcus and Dekker, 2013). The best studied RNP in eukaryotes is, however, RNPII which synthesizes all the protein-coding genes, as well as most microRNA (miRNA) and long non-coding RNA (lncRNA) species.

After the assembly of PIC and recruitment of RNP, several signals confer to the release of the RNP to its elongation mode. At certain gene promoters, the PIC and RNPII are assembled and disassembled in a cyclinc manner (reviewed in Metiviér *et al*., 2006). However, approximately 30% of human gene promoters harbor RNPII that is engaged in transcription but kept paused by negative elongation factors (reviewed in Adelman and Lis, 2012). Indeed, the release of paused RNPII to elongation has emerged as an efficient means to coordinate gene expression programs, particularly during development and in response to activating signals. Efficient transcription also involves melting of the DNA strands and clearance of the gene body from obstructing proteins and DNA coils (reviewed in Feklistov, 2013; Fuda *et al*., 2009; Selth *et al*., 2010). The recruitment and assembly of PIC and RNP, as well as the following steps of promoter escape, elongation, termination, and reinitiation depend on a synergistic action of transcriptional regulators and chromatin modifying enzymes.

2.1.4 Transcription factors

The assembly of PIC and RNP is directed by sequence-specific transcriptional regulators called transcription factors. Transcription factors are characterized by a DNA-binding domain (DBD) that recognizes short, generally in the range of 6-12 bp, DNA sequences (reviewed in Maston *et*

al., 2006; Pabo and Sauer, 1992). Many transcription factors form homo- or hetero-oligomers, which is reflected in the consensus DNA-binding element, often composed of two half-sites. A transcription factor can recognize several variants of its consensus DNA element, but the precise sequence can dictate the regulatory impact, for example, by directing the oligomerization partner or by conferring affinity advantage for certain loci over the others (reviewed in Maston *et al*., 2006). Transcriptional regulators act in a complex chromatin environment where, besides the underlying *cis*-elements, also the composition of other components at the target loci affects the binding ability and the regulatory output. Most transcription factors can directly or indirectly regulate the assembly of PIC, promote the escape of RNP, or recruit chromatin modifying enzymes that either enhance or inhibit the steps of gene expression. Transcriptional activation can also be influenced by cofactors which typically do not bind to DNA, but instead, target transcriptional regulators and modulate their transactivating capacity (reviewed in Roeder, 2005). Thus, the transcriptional control engages synergistic action from several transcriptional regulators and chromatin modifying enzymes that integrate the cellular and physiological signals to a coordinated behavior of cells and organisms.

2.1.5 Transcription elongation, termination and reinitiation

Most of our knowledge on regulation of gene expression originates from RNPII-mediated transcription of protein-coding genes. For RNPII, the promoter escape is mediated by releasing inhibitory proteins, such as negative elongation factor (NELF), and by recruitment of activating factors such as positive transcription elongation factor b (P-TEFb). The P-TEFb complex mediates phosphorylation of serine (S) 2 at the heptad repeat of C-terminal domain (CTD) of RNPII, enabling RNPII to enter elongation. Transcriptional elongation is facilitated by chromatin remodeling factors that clear the gene body from obstructing proteins and by topoisomerases that cut and paste DNA strands to relieve coiling (reviewed in Feklistov, 2013; Fuda *et al*., 2009; Selth *et al*., 2010). At the 3' end of the gene, RNPII meets a polyA site and the transcription terminates. In many cases the components of RNP complex are recycled to the promoter for reinitiation of transcription (reviewed in Gilmour and Fan, 2008; Richard and Manley, 2009).

2.1.6 Post-transcriptional processing of RNA and protein

Protein-coding genes contain exons that code for amino acid sequences and introns that do not code for proteins. Initially, the gene is transcribed as one unit termed pre-messenger RNA (pre-mRNA), but during a process called splicing, the introns are removed and a defined set of exons are united to form a mature mRNA (reviewed by Darnell, 2013; Kornblihtt, 2007). Splicing is an important step in controlling the transcript variant and provides a mechanism for producing several protein isoforms from a single gene. The processing of mRNA includes also methylation of the 5' end, as well as polyadenylation of the 3' end. The 3' polyA tail is bound by proteins that transport the mRNA to the cytosol, whereas the 5' region interacts with ribosomes to initiate translation (reviewed by Darnell, 2013; Kornblihtt, 2007).

The nucleotide sequence of an mRNA is translated into a sequence of amino acids that defines the structure and function of a protein. During translation, ribosomes position tRNA molecules to recognize nucleotide triplets, so called codons, of the mRNA and to transfer a corresponding amino acid to the growing polypeptide chain (reviewed by Schmeing and Ramakrishnan, 2009). Translation is initiated in the cytoplasma but signal sequences emerging at the polypeptide chain can direct the ribosome to a specific cellular location, e.g. to the membrane of endoplasmic reticulum (ER) where all the secreted and membrane-associated proteins are synthesized.

Proteins are versatile molecules whose proper biological functions depend on correct three-dimensional shape and conformational flexibily. While small proteins that comprise only one functional domain fold efficiently also *in vitro*, large proteins in the context of the crowded cellular environment require molecular chaperones for an efficient assembly. Besides assisting in *de novo* folding of newly synthesized proteins, molecular chaperones provide constant surveillance of the proteome and are integrated into the networks that coordinate protein-protein-interactions, localization, stability and degradation (reviewed in Hartl *et al*., 2011).

2.2 Transcriptional regulation of cellular differentiation

Every mitotic cell division leads to symmetric division of the hereditary information. As a consequence, virtually every cell in an organism contains the same genetic information, which incites a question on how the distinct cell types and organs emerge? Despite the symmetrical distribution of the replicated genome, all the material in the cell is not equally divided between the daughter cells. For example at fertilization, the site of sperm entry and the position of polar bodies of the egg define an equatorial line for cleavage and initiate redistribution of maternal RNA and protein (Edwards and Beard, 1997; Kumano, 2012; Piotrowska and Zernicka-Goetz, 2002; Roegiers and Jan, 2004). This cellular polarization determines the dorso-ventral axis of the developing zygote at the very first division. Besides gaining a different set of regulatory molecules, the cells are directed by signals from their surroundings. Particularly during early development, positive and negative regulatory networks culminate on the genome and cause different transcriptional programs that direct the neighboring cells towards distinct lineages. Among the best studied examples of early transcriptional regulators are the homeobox (HOX) proteins, which activate and repress gene groups that coordinate the segmentation of the body (reviewed in Pearson *et al*., 2005). The cell-specific expression patterns of *HOX* genes, in turn, are regulated by trithorax-group (trxG) proteins that maintain genes in an active state and by polycomb-group (PcG) proteins that can repress gene activity over many cell generations (Schuettengruber *et al*., 2007).

2.2.1 Differentiation and commitment

Differentiation towards a committed cell type is accompanied with a genome-wide patterning of the chromatin. This epigenetic chromatin state is maintained over mitotic division and passed on to the daughter cells (reviewed in Delcuve *et al*., 2008; Probst *et al*., 2009). In pluripotent stem

cells, the loosely packed chromatin enables plasticity of gene expression and differentiation towards versatility of cell types (Fussner *et al*., 2011). The lineage-specific genes are kept in a silent but transcriptionally available (poised) state, which is characterized by repressive H3K27me3 and activating H3K4me3 histone modifications. This bivalent chromatin enables rapid transcription when the cell becomes committed for differentiation (Bernstein *et al*., 2006; reviewed in Tollervey and Lunyak, 2012).

Majority of chromatin modifications in differentiating cells, however, is involved in silencing and compacting chromatin regions that are not utilized by the terminally differentiated cells. For example, *β-globin* is in an open state only in cells of erythrocyte lineage where its abundant expression generates components for oxygen transport (Reddy *et al*., 1994). Especially in complex organisms, the billions of nucleotides in each cell create a daunting task for transcriptional regulators to localize a couple of nucleotides long DNA elements. Over the course of evolution, the expansion in the genome size has been accompanied with evolvement of mechanisms for chromatin compaction (reviewed in Mohn and Schübeler, 2009). Histones emerged in early eukaryote development, and through the versatile PTMs, create a specific code that directs transcriptional regulators (reviewed in Zentner and Henikoff, 2013). Moreover, the PcG proteins have undergone high diversification in multicellular organisms (Schuettengruber *et al*., 2007; Whitcomb *et al*., 2007) and, in vertebrates, genome-wide DNA methylation efficiently compacts and represses chromatin regions (Guibert *et al*., 2009; Mohn and Schübeler, 2009). Although the epigenetic status of the chromatin can be modified even in terminally differentiated cells, the distinct features of the genomic regions define the chromatin landscape for transcriptional responses and direct the ability of cells to respond to the guidance cues. As a result, different cell types have profoundly distinct appearances, express different sets of molecules and carry a wide versatility of functions.

During the life cycle of a mammal, a global resetting of the epigenome occurs during gametogenesis and early embryogenesis. This reacquisition of totipotency is crucial for the ability of germ cells to develop into mature gametes, and for the fertilized egg to give rise to all the cell types of the new individual (reviewed in Cantone and Fisher, 2013). In mouse, the global resetting of the epigenome is initiated at E7.5-12.5 when the primordial germ cells migrate to the genital ridges (Cantone and Fisher, 2013). This epigenetic restoration involves incorporation of histone variants, RNA-mediated silencing of repetitive DNA elements, and establishment of germ line-specific pattern of DNA methylation (Hajkova *et al*., 2008; 2010). After reaching the genital ridges, the germ cells undergo a mitotic or meiotic arrest that sustains until sexual maturity. In post-pubertal animals, the gametogenesis is reactivated and produces haploid germ cells that are able to generate individuals with a unique genetic composition. Upon fertilization, a second wave of epigenetic resetting takes place to reassure the capability of sperm and egg DNA to fuse and generate all the cell types of the new individual (Cantone and Fisher, 2013).

2.2.2 Spermatogenesis

Male gametogenesis is a remarkable process of cellular differentiation that produces millions of sperm cells daily. In testis of an adult organism, the germ cells undergo extensive clonal

expansion through mitosis, reduction division of the genome in meiosis and profound morphological differentiation of haploid gametes (reviewed in Rousseaux *et al*., 2005). These carefully regulated processes of differentiation are spatio-temporarily organized in the seminiferous epithelium of testis and include genome-wide transcriptional programming, chromatin reorganization and transition of the epigenetic state (reviewed in Meikar *et al*., 2012). Spermatogenesis is under an endocrine control via pituitary-hypothalamus-axis and by testosterone that is secreted by Leydig cells in the testicular interstitium (Dohle *et al*., 2003; Rousseaux *et al*., 2005). Inside the seminiferous tubules, Sertoli cells are the only somatic cells and vital for gamete production. The Sertoli cells provide the developing germ cells both physical and nutritional support, and enable coordinated differentiation and migration of germ cells from the basal lamina to the lumen of the tubule (Figure 3). The mature spermatozoa are highly specialized cells that are released to the lumen and transported to the epididymis where they gain the capacity to move and fertilize an egg.

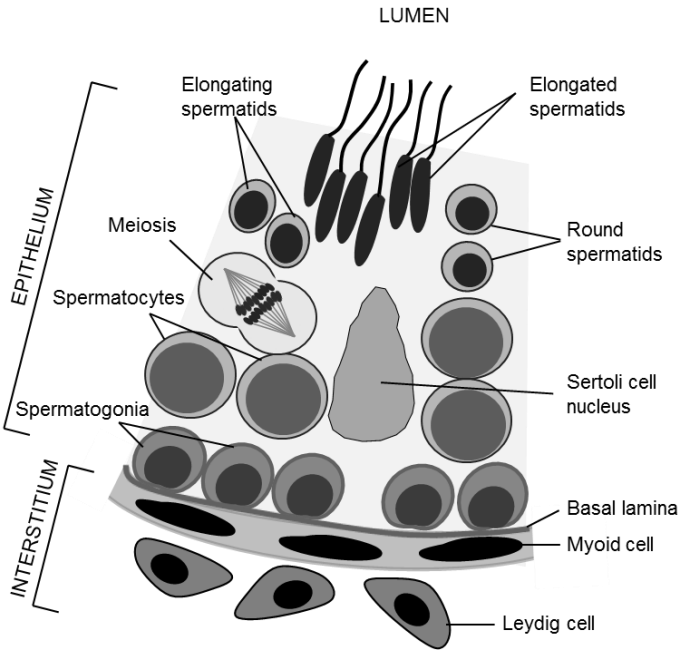


Figure 3. Cells of spermatogenesis. Spermatogenesis takes place in the epithelium of seminiferous tubules of testis. The only somatic cells inside the seminiferous tubules are Sertoli cells which provide the developing germ cells with physical and nutritional support. Spermatogonia reside on the basal lamina and undergo mitotic cell divisions. Primary spermatocytes (leptotene, zygotene, pachytene, diplotene) synthesize DNA and undergo crossing over prior to the meiotic divisions. The haploid spermatids are initially small and round but during spermiogenesis they differentiate into spermatozoa which are released into the lumen of the tubule. Contractions of the Myoid cells beneath the basal lamina flush the immotile spermatozoa to the epididymis. Interstitial cells include Leydig cells which secrete testosterone.

2.2.2.1 Clonal expansion and meiosis of spermatogenic cells

Germ line stem cells and spermatogonia reside attached to the basal lamina and, in a process called spermatocytogenesis, undergo clonal expansion to maintain the stem cell population and to give rise to a large pool of cells that are committed for differentiation. The following meiosis of spermatocytes takes 1.5-2 weeks in mammals, includes synthesis of DNA, active transcription, chromatin remodeling and reduction division of the hereditary material (reviewed in Kimmins *et al*., 2004). The synthesis of DNA occurs in leptotene, and the subsequent pairing of homologous

chromosomes in the zygotene spermatocytes. The following pachytene is characterized by fully formed synaptonemal complexes and exchange of genetic material during crossing over (reviewed in Kimmins and Sassone-Corsi, 2005; Turner, 2007). The pachytene stage is, furthermore, an active phase of RNA and protein synthesis and includes exchange of histones to testis-specific variants (Kimmins and Sassone-Corsi, 2005; Gaucher *et al*., 2010; Soumillon *et al*., 2013). A specific feature of the male gametogenesis is the separation of the X and Y chromosomes, which contain only a small region that is capable for recombination and crossing over. As a consequence of no crossing over, the X and Y chromosomes undergo meiotic sex-chromosome inactivation (MSCI; Turner, 2007), which is mediated by incorporation of histone variants such as macroH2A and H3.3 and specific histone modifications (McKee and Handel, 1993; Khalil *et al*., 2004; Solari, 1974; reviewed in Burgoyne *et al*., 2009). Particularly, phosphorylation of H2AX at serine 139 (γH2AX) has been identified as a hallmark for the transcriptional silencing and sequestration of the X and Y chromosomes into the cellular periphery where they form a structure called sex-body (reviewed in Turner, 2007; Burgoyne *et al*., 2009). In diplotene spermatocytes, the synaptonemal complexes dissolve, and meiosis I and II separate the homologous chromosomes and sister chromatids, respectively. The meiotic divisions are relatively fast processes (completed within 24h) and result in the formation of four haploid spermatids per primary spermatocyte (reviewed in Hess and de Franca, 2008).

2.2.2.2 Morphological differentiation of spermatids

Spermatids are initially small and round, but during a process called spermiogenesis, they develop into specialized spermatozoa with tightly packed DNA in the head, spirally organized mitochondria in the mid piece and a long tail which is composed of microtubules (Rousseaux *et al*., 2005). In mouse, there are 16 steps of spermiogenesis during which spermatid-specific organelles form, the nuclei elongate and the DNA becomes tightly compacted (Oakberg, 1956; reviewed in Hess and de Franca, 2008). A remarkable phenomenon in step 2-6 spermatids is the high rate of transcription and the storage of RNA. The transcripts are stored in specialized organelles called chromatoid bodies and used later when the chromatin is tightly compacted and transcriptionally silent (reviewed in Kotaja and Sassone-Corsi, 2007; Parvinen 2005). Elongation of DNA is initiated in step 9 spermatids and the consequent shift from transcriptional to translational control is reflected in the many RNA species and RNA-binding proteins that are present in elongating spermatids (reviewed in Kleene, 2003; Paronetto and Sette, 2010). During steps 9-11, the DNA-packaging histones are changed to transition proteins (TNPs), which in turn are replaced by protamines (PRMs) in late spermatids. PRMs are small and positively charged proteins, the incorporation of which causes extreme compaction of chromatin and inactivation of transcription (Pogany *et al*., 1981; reviewed in Braun, 2001; Miller *et al*., 2010).

2.2.2.3 Organization of spermatogenic cells in the seminiferous epithelium

During spermatogenesis, every mitotic and later meiotic cell division is followed by an incomplete cytokinesis that results in the formation of cytoplasmic bridges between the germ

cells (Erickson, 1973). These cytoplasmic channels unite germ cells that originate from one progenitor stem cell, and enable exchange of material and information between different developmental states (Ventelä *et al*., 2003). Sharing material is especially crucial for the haploid spermatids that, besides becoming transcriptionally silent, contain only one of the two sex chromosomes. Remarkably, the chromatoid bodies have been shown to translocate via cytoplasmic bridges to other cells in the same cyncytium (reviewed in Parvinen, 2005). Together with Sertoli cells, the cytoplasmic bridges enable synchronized development of germ cells, which gives rise to defined cell associations termed stages. In mouse, there are twelve (I-XII) stages that follow each other as a wave of seminiferous cycle (Figure 4; Oakberg, 1956). Each stage contains a specified set of cells that give characteristic appearance for the tubule region (Figure 4; Kotaja *et al*., 2004; Parvinen and Hecht, 1981). For a given tubule region it takes eight days to undergo all the twelve stages, whereas 4.5 cycles (~36 days) are required for a type A spermatogonia to develop into spermatozoa. Under a dissection microscope, the individual stages are challenging to

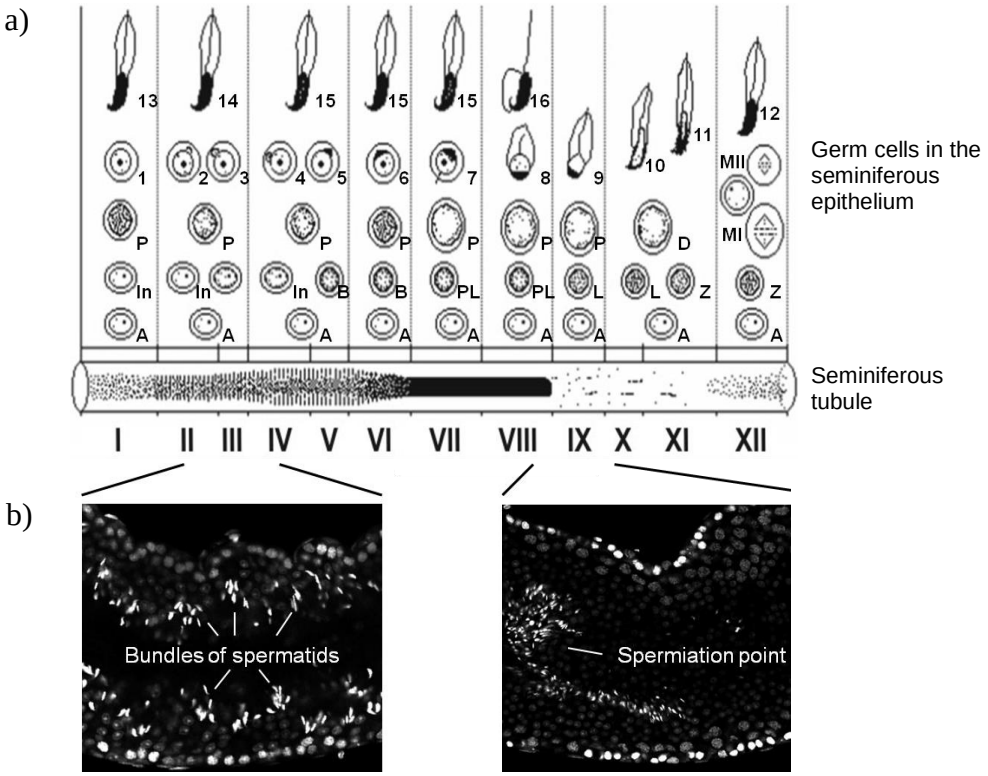


Figure 4. Stages of seminiferous cycle in mouse. a) Each stage (vertical columns, Roman numerals) contains a specific association of germ cells, which gives a specific light-absorbing pattern to the tubule. A: type A spermatogonium; In: intermediate spermatogonium; B: type B spermatogonium; PL: preleptotene spermatocyte; L: leptotene spermatocyte; Z: zygotene spermatocyte; P: pachytene spermatocyte; D: diplotene spermatocyte; MI: meiosis I; MII; meiosis II. The numbers indicate steps of spermiogenesis. 1-8: round spermatids; 9-12 elongating spermatids; 13-16 elongated spermatids. b) Confocal images of seminiferous tubules showing DNA in white.

identify, but four defined regions are easily recognized: Dark zone (DZ) contains stages VII-VIII and gets its light-absorbing appearance from the nearly mature spermatids that pack along the lumen (Figure 4). The abrupt release of the spermatozoa to the lumen is detected as a spermiation point that marks the transition from DZ to pale zone (PZ) (see Figure 4). PZ contains stages IX-XI and its most mature cells are step 9-11 spermatids that are sparsely distributed in the seminiferous epithelium. In the following weak spot (WS) the spermatids form bundles which are detected as small light-absorbing spots in stages XII-I. As the clusters of spermatids grow in size and their DNA condenses, the tubule gains a characteristic brush-like appearance that indicates stages II-VI of the strong spot (SS). Although the stages are identified mainly according to the morphology and associations of elongated spermatids, each stage contains also a defined set of spermatocytes and spermatids (Figure 4; Russell *et al*., 1990). For example, DNA condensation in leptotene spermatocytes occurs in stages IX-X, pachytene phase extends over stages I-IX, meiotic divisions take place in stage XII and round spermatids reside in stages I-VIII (Figure 4).

2.3 The heat shock response

Cells are constantly challenged by external and internal conditions that can cause disrupted homeostasis. Such conditions include elevated temperatures, toxic compounds and cancer progression which all have deleterious effects on cellular organization and affect a broad range of structures from membranes to cytoskeleton and from organelles to the DNA. Mild protein-damaging stress causes microfilaments to reorganize, whereas severe stress leads to collapsed cytoskeletal networks (reviewed in Toivola *et al*., 2010). Moreover, upon stress, the fluidity of plasma membrane increases, nucleoli swell, Golgi and ER become fragmented and mitochondria loose the correct structure (reviewed in Boulon *et al*., 2010; Szalay *et al*., 2007; Vigh *et al*., 2007). As a result, the cellular transport system becomes defected and the production of ATP severely impaired. If not mitigated, the stress eventually causes apoptotic and necrotic death of the cells (reviewed in Richter *et al*., 2010).

The cell's response to protein-damaging conditions is called heat shock response (HSR), a rapid and evolutionarily conserved mechanism that adapts elemental cellular processes to the harmful conditions. The HSR was discovered in 1962 by Ferruzio Ritossa who observed chromosomal puffs in heat-treated salivary gland cells of *D. melanogaster* larvae (Ritossa, 1962). These puffs appeared within minutes of the stress stimuli, coincided repeatedly at specific chromosomal loci and were already at the time known to be hallmarks of active transcription. It, however, took until 1974 before Alfred Tissières and others showed stress-induced production of RNA from these loci and correlated the transcriptional activation to the production of heat shock proteins (HSPs). Importantly, the onset of thermal stress was also detected to halt the production of constitutively expressed proteins owing to global silencing of transcription and translation (Lewis *et al*., 1975; Spradling *et al*., 1975; Tissières *et al*., 1974; reviewed in Lindquist, 1981).

The HSPs are molecular chaperones and proteases that enforce correct folding, assembly, translocation and degradation of nascent and denatured proteins both in the cytosol and organelles (reviewed by Hartl *et al*., 2011). Hence, the first line of defence against protein-

damaging conditions is the rapid production of proteins that maintain homeostasis (reviewed in Morimoto, 1998). HSPs are divided into families based on their molecular size (Kampinga *et al*., 2009), and in human, HSPB1/HSP27, DNAJB1/HSP40, HSPA1A/HSP70.1, and HSPC/HSP90 have been shown to be the main stress-induced HSPs. The members of HSP70 and HSP90 chaperone families are ATP-dependent protein foldases that recognize exposed hydrophobic amino acid residues and provide an environment for an efficient refolding. Small HSPs (sHSPs) and type I and II chaperonins (HSPD1/HSPE1 and CCTs, respectively) work independently of ATP by attaching to misfolded proteins and holding them until cleared by proteasomal degradation or refolded by the HSP40-HSP70 and HSP90 machineries (reviewed in Hartl *et al*., 2011; Richter *et al*., 2010). Besides leading to induced expression of molecular chaperones, heat stress has been detected to cause a stagnation of cell cycle progression and increased expression of metabolic enzymes, membrane modulating proteins, transcription factors and components involved in nucleic acid repair (Gasch *et al*., 2000; Hahn *et al*., 2004; Kühl and Rensing, 2000; Trinklein *et al*., 2004; reviewed in Richter *et al*., 2010).

2.4 The family of heat shock factors

Cloning and deletion mapping of *HSP* genes enabled identification of the promoter element that mediates the heat-induced activation of transcription. This element consisted of inverted nGAAn pentamers and was named the heat shock element (HSE) (Amin *et al*., 1988; Pelham, 1982; Sorger and Pelham, 1988). The identification of HSE, in turn, permitted purification and characterization of the HSE-binding protein which was named the heat shock factor (HSF) (Topol *et al*., 1985; Wiederrecht *et al*., 1987; Wu *et al*., 1987). Since the initial discovery in yeast and fruit fly, the HSFs have been identified as a conserved family of transcription factors that orchestrate gene expression in eukaryote species. Besides providing stressed cells with chaperones that safeguard homeostasis, HSFs act on a plethora of physiological pathways that range from developmental processes to aging and immune responses (reviewed in Åkerfelt *et al*., 2010; Fujimoto and Nakai, 2010). The recently discovered importance of HSFs for a wide range of core physiological processes is highlighted in severe pathologies, e.g. Alzheimer's and other neurodegenerative diseases, cancers, infertility and cataract that have been associated with disrupted activity of HSFs (Christians *et al*., 2000; Dai *et al*., 2007; Fujimoto *et al*., 2004; 2005; Inouye *et al*., 2004; Metchat *et al*., 2009; Santagata *et al*., 2011; 2013; Takaki *et al*., 2006).

A single HSF is expressed in yeasts and invertebrates, whereas two whole-genome duplications during chordate/vertebrate development have led to the existence of four HSFs (HSF1-4) in mammals (reviewed in Fujimoto and Nakai, 2010). HSF1 is a homolog for the single HSF in yeasts and invertebrates, and the major director of stress responses. No other vertebrate HSF can compensate HSF1 for the heat-induced activation of *HSP* gene transcription (McMillan *et al*., 1998; reviewed in Vihervaara and Sistonen, 2014). Also HSF2 is ubiquitously expressed in vertebrate species and coupled to responses to acute and chronic stress. However, the role of HSF2 as a regulator of stress responses has remained enigmatic and it has mainly been considered as a developmental factor and a modulator of HSF1-driven transcription (Chang *et al*., 2006; Kallio *et al*., 2002; Östling *et al*., 2007; Shinkawa *et al*., 2011; Sistonen *et al*., 1992).

Similarly to HSF1, HSF3 is a stress-responsive factor, but it induces the expression of *HSP*s only in avian species (Fujimoto *et al*., 2010). Also mice harbor a functional HSF3, but it is involved in the expression of non-classical heat shock genes (Fujimoto *et al*., 2010). In humans, HSF3 is likely a pseudogene since no HSF3 transcript has been identified. HSF2 and HSF4 are mainly associated with developmental processes and, intriguingly, no stress-related function for HSF4 has been detected, indicating functional diversity of the HSF family members (Fujimoto *et al*., 2004, reviewed in Fujimoto and Nakai, 2010). Both HSF2 and HSF4 have been shown to interplay with HSF1 at the target gene promoters, albeit at distinct tissues, physiological pathways and through distinct mechanisms. While HSF1 and HSF2 form heterotrimers and collaborate during stress responses and gametogenesis (Loison *et al*., 2006; Östling *et al*., 2007; Sandqvist *et al*., 2009; Wang *et al*., 2004), HSF4 competes with HSF1 for the same promoters in sensory plaque (Fujimoto *et al*., 2004; Takaki *et al*., 2006). Consequently, appropriate cooperation of HSF1 and HSF4 coordinates the spatio-temporal gene transcription and drives development and maintenance of lens and olfactory epithelium. The HSF family also includes HSFX and HSFY, which are of sex chromosomal origin. Both HSFX and HSFY are highly expressed in testis and coupled to defective spermatogenesis (Bhowmick *et al*., 2007; Kinoshita *et al*., 2006; Tessari *et al*., 2004). However, the cellular functions and molecular mechanisms of these non-classical HSFs are still undetermined.

2.4.1 The functional domains of HSFs

HSFs are characterized by a conserved N-terminal DBD that contains a looped helix-turn-helix structure (Figure 5; Littlefield and Nelson, 1999; Sorger and Pelham, 1988). Unlike most

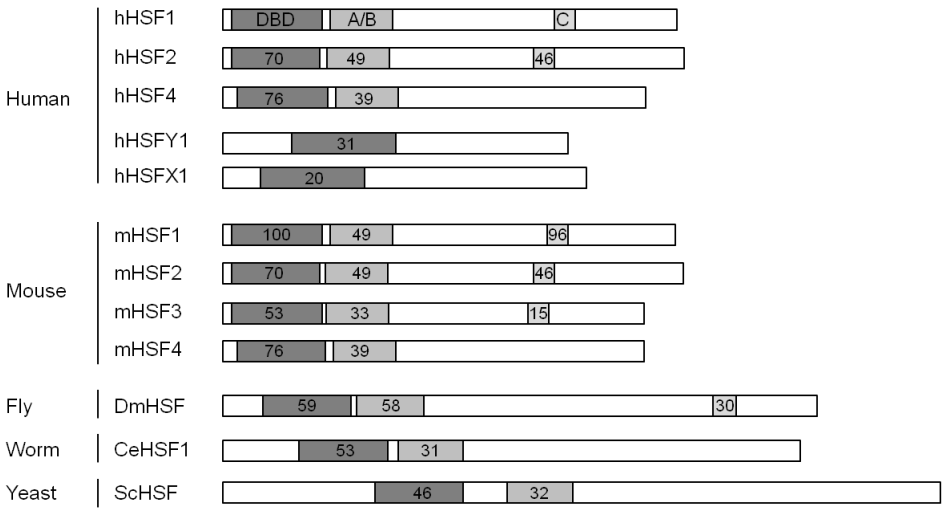


Figure 5. Functional domains and evolutional conservation of HSFs. Schematic presentation of HSF family members. The numbers in specified functional domains indicate percentage of amino acid similarity to human HSF1 and are according to Fujimoto and Nakai (2010). DBD: DNA-binding domain; A/B: oligomerization domain; C: hydrophobic heptad repeat C; h: *Homo sapiens*; m: *Mus musculus*; Dm*: Drosophila melanogaster*; Ce: *Caenorhabditis elegans*; Sc: *Saccharomyces cerevisiae*.

transcription factors, HSFs bind to the DNA as trimers and the oligomerization is mediated by a conserved HR-A/B domain that is constituted of heptad repeats (Peteranderl *et al*., 1992; Rabindran *et al*., 1993; Sorger and Nelson, 1989). Hydrophobic amino acid residues of the HR-A/B enable formation of a coiled-coil structure between HSFs and allow for the trimerization. Spontaneous HSF activation is thought to be inhibited by a C-terminal heptad repeat (HR-C) that forms intramolecular contacts with the HR-A/B and, thereby, allosterically blocks the oligomerization (Rabindran *et al*., 1993). This hypothesis is supported by deletion studies where lack of HR-C rendered HSF1 constitutively active (Rabindran *et al*., 1993; Zuo *et al*., 1994, Nakai *et al*., 2000), and by yeast HSF and mammalian HSF4, both of which lack the HR-C (Figure 5) and are constitutively bound to the DNA (Jakobsen and Pelham, 1988; Nakai *et al*., 1997). HSFs also contain regulatory domains (RDs), nuclear localization signals (NLSs) and transactivation domains (ADs), which are targeted by several protein complexes that coordinate the steps and the magnitude of HSF activation (reviewed in Anckar and Sistonen, 2011). However, the number, location and amino acid sequences of these regulatory regions vary between distinct HSFs.

All HSFs recognize *cis*-acting HSEs, but show preferences over the precise sequence and architecture of the nGAAn repeats (Kroeger and Morimoto, 1994; Yamamoto *et al*., 2009). Intriguingly, increasing number of nGAAn pentamers have been shown to mediate cooperative binding of HSFs, and the exact composition of the HSE to affect the transcriptional activation of target genes (Hashikawa *et al*., 2007; Perisic *et al*., 1991; Sakurai *et al*., 2010; Xiao *et al*., 1991).

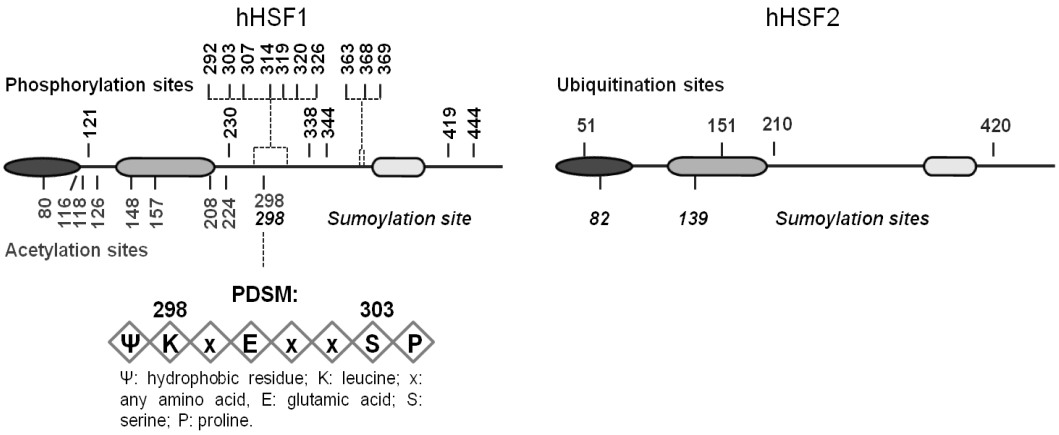


Figure 6. Post-translational modifications of human HSF1 and HSF2. HSF1 (left) is phosphorylated at multiple amino acids (indicated by numbers above the schematic HSF1) and subjected to acetylation (gray, below HSF1) as well as sumoylation (italicized). Only a few amino acids of HSF2 (right) are ubiquitinated (above HSF2) or sumoylated (italicized). The augmented regions in the HSFs denote DBD, HR-A/B and HR-C and are indicated as in Figure 5. The RD of HSF1 locates between the HR-A/B and HR-C domains (not shown) PDSM: phosphorylation-dependent sumoylation motif.

2.4.2 Regulation of HSF1 and HSF2

The molecular structure and stress-sensitivity of HSF1 is highly conserved in eukaryote species. The need for rapid and delicate control of HSF1 is highlighted by the multitude of protein-protein interactions and PTMs that coordinate its activity (Figures 6 and 7; reviewed in Anckar and Sistonen, 2011; Vihervaara and Sistonen, 2014). According to the current model, HSPs can directly bind to and inhibit HSF1, which creates an autoregulatory mechanism that senses protein folding in the cell and fine-tunes the intensity of the stress responses (Baler *et al*., 1992; Shi *et al*., 1998). While HSP90 has been shown to restrain HSF1 in a monomeric state, HSP40 and HSP70 interact with trimeric HSF1 and dampen its transctivating potential (Abravaya *et al*., 1992; Guo *et al*., 2001; Zou *et al*., 1998). Upon heat stress, HSF1 dissociates from the HSPs, trimerizes and binds to the DNA (Figure 7). The stress-induced activation of HSF1 is accompanied with hyperphosphorylation of several amino acid residues at the RD that is localized between HR-A/B and HR-C domains (Figure 6; reviewed in Anckar and Sistonen, 2011). Although the hyperphosphorylation of HSF1 is associated with its transactivating capacity, several amino acids are phosphorylated also in the absence of stress and repress the activation (Chu *et al*., 1996; Kline and Morimoto, 1997). Intriguingly, of the multitude of PTM events that occur on HSF1, only phosphorylations of S230 and S326 have been coupled to increased transactivation (Holmberg *et al*., 2001; Guettouche *et al*., 2005). Instead, several PTMs contribute to defining the extent and duration of HSF1-mediated transcription. At so called phosphorylation-dependent sumoylation motif (PDSM), phosphorylation of S303 primes K298 for sumoylation, which in turn, restrains HSF1-mediated transcription (Hietakangas *et al*., 2003; 2006). The PDSM was first characterized on HSF1 but it is now recognized as a conserved mechanism that controls a range of transcription factors (Anckar and Sistonen, 2007). During prolonged stress, HSF1 activity attenuates and the *HSP70* transcription ceases. The removal of HSF1 from the chromatin is facilitated by acetylation of K80, a site that directly contacts the DNA (Figures 6 and 7, Westerheide *et al*., 2009). The HSF1 occupancy on DNA, and thus the transcription of *HSP70*, can be prolonged by Sirtuin 1 (SIRT1), which is a deacetylase, nutrient sensor and longevity factor that has been suggested to link HSF1 to the metabolic state of the organism (Westerheide *et al*., 2009).

In striking contrast to the multitude of PTMs on HSF1, HSF2 is mainly regulated at the level of its expression (Ahlskog *et al*., 2010; Björk *et al*., 2010; reviewed in Björk and Sistonen, 2010). To date, no phosphorylation of HSF2 has been reported and the only PTMs detected on HSF2 are ubiquitination of S51, K151, K210 and K420, as well as sumoylation of K82 and K139 (Figure 6; Xu, *et al*., 2012). Also the activation of HSF1 and HSF2 are controlled by distinct mechanisms: Althoug HSF2 binds to the DNA as a trimer, in unstressed conditions it exists mainly as a dimer (Figure 7; Sistonen *et al*., 1994). Moreover, the mere increase in HSF2 concentration has been shown to cause its translocation to the nucleus and binding to the DNA (Sandqvist *et al*., 2009). HSF1 and HSF2 co-occupy *HSP* promoters, but since HSF2 is a poor activator of *HSP* genes (Kroeger *et al*., 1993; Sarge *et al*., 1993; Yoshima *et al*., 1998), these factors are likely to display diverse mechanisms at the target genes. The factor-specific control and different transcriptional contributions are partially explained by diverging amino acid sequences of HSF1 and HSF2. While the DBD, HR-A/B and HR-C are well conserved between HSF1 and HSF2 (Figure 5),

their overall amino acid similarity is only 35% (reviewed in Pirkkala *et al*., 2001). The distinct regulation of HSF1 and HSF2 are highlighted by the selective removal of HSF2 from the *HSP70* promoter during heat stress (see Figure 7; Ahlskog *et al*., 2010). The detailed mechanisms of regulation and transcriptional contributions of HSFs remain to be elucidated, but the highly specific coordination of HSF1 *versus* HSF2 demonstrates that cells have both the need and the means to delicately control the interplay of HSFs (Figure 7).

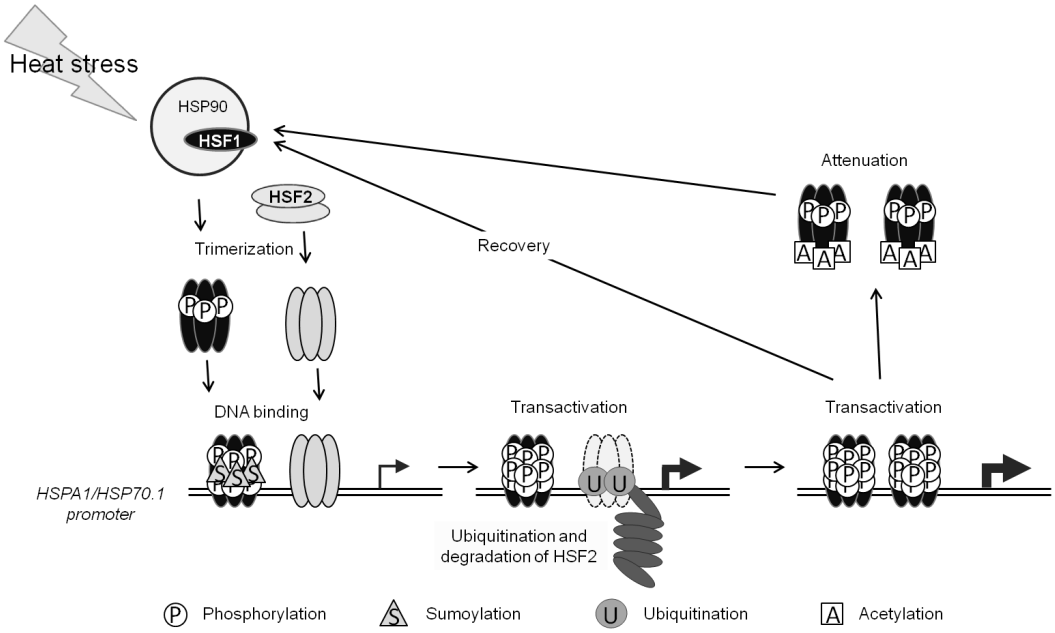


Figure 7. Activation-attenuation cycle of human HSF1 and HSF2 upon heat stress. Heat stress induces trimerization of HSF1 and HSF2 and leads to their binding to *HSP70.1*/*HSPA1A* promoter. The activity of HSF1 is controlled by HSPs and a multitude of PTMs, whereas HSF2 is mainly controlled at the level of expression. From the HSF1-HSF2 complex at the *HSP70.1* promoter, HSF2 is selectively degraded during the heat stress.

2.4.3 HSF1-mediated transcriptional initiation

The rapid HSF-mediated increase in *HSP* expression during stress has provided a robust model for studying transcriptional responses (reviewed in Guertin *et al*., 2010). Particularly the polytene chromosomes in fruit fly *D. melanogaster* have been exploited for detailed investigations on the dynamic recruitment of transcription factors to the *Hsp70* loci (Yao *et al*., 2006; Zobeck *et al*., 2010). The *Hsp70* promoter is primed for activation by GAGA binding factor (GAF) and transcriptionally engaged RNPII that is paused by DRB sensitivity-inducing factor (DSIF) and NELF (Rougvie and Lis, 1988; Rasmussen and Lis, 1993; Wu *et al*; 2003). Pausing of transcriptionally engaged RNPII was discovered at the *Hsp70* loci and is today acknowledged as a common mechanism that confers rapid gene activation and temporal coordination of transcriptional programs (Rasmussen and Lis., 1993; Rougvie and Lis, 1988; reviewed in Adelman and Lis, 2012). Heat stress induces binding of HSF to the *Hsp70* promoter and leads to recruitment of the mediator complex as well as P-TEFb that phosphorylates the C-terminal region

of RNPII (Brès *et al*., 2008; Lis *et al*., 2000; Marshal *et al*., 1996; Park *et al*,. 2001). The transcriptional elongation is coupled to chromatin remodeling via removal of histones by facilitates transcription (FACT) and suppressors of Ty (SPT5, SPT6) as well as relieving of DNA coils by topoisomerase I. Moreover, poly-ADP ribose polymerase (PARP) generates a transcriptional compartment in which transcriptional components are efficiently recycled (Andrulis *et al*., 2000; Gilmour *et al*., 1986; Kaplan *et al*., 2000; Petesch and Lis, 2008; Zobeck *et al*., 2010).

Also in mammals the recruitment of HSF1 leads to the release of paused RNPII and removal of nucleosomes along the *HSP70* (Brown *et al*., 1996; Brown and Kinston, 2007). However, the highly conserved function of HSF1 as inducer of *HSP* expression has species-specific mechanisms. In mammals, HSF1 collaborates with replication factor A (RPA) to maintain the *HSP70* promoter in an accessible state and, upon stress, the nucleosomes are removed along the gene via switch/sucrose nonfermenting (SWI/SNF) chromatin remodeling complex (Brown *et al*., 1996; Fujimoto *et al*., 2012; Sullivan *et al*., 2001). In mammals, HSF1 also interacts with HSF2 which has been suggested, promoter-specifically, to modulate the HSF1-driven gene expression (Östling *et al*., 2007; Sandqvist *et al*., 2009).

2.4.4 HSF1 and HSF2 in developmental processes

The significance of HSFs for developmental processes was revealed by inactivation studies. In yeast, HSF is indispensable for growth and viability in non-stressed conditions, and in fruit fly, it is required for oogenesis and larval development (Gallo *et al*., 1993; Jedlicka *et al*., 1997; Sorger and Pelham, 1988). In mammals, HSFs are involved in a variety of developmental processes from embryogenesis to gametogenesis. Although *Hsf1*$^{-/-}$ mice are viable, they show growth retardation and female sterility due to placental insufficiencies (Christians *et al*., 2000; Xiao *et al*., 1999). Furthermore, HSF1 is required for correct oogenesis, IgG production and formation and maintenance of sensory epithelium (Fujimoto *et al*., 2004; Inoye *et al*., 2004; Le Masson *et al*., 2011; Metchat *et al*., 2009; Takaki *et al*., 2006). Abundant HSF2 expression has been detected during embryogenesis, particularly in the developing nervous system (Min *et al*., 2000; Rallu *et al*., 1997), and consequently, HSF2 has been shown to be essential for correct cortical lamination (Chang *et al*., 2006). Male gametogenesis has been reported to be intact (Izu *et al*., 2004) or only slightly defected (Salmand *et al*., 2008) in *Hsf1*$^{-/-}$ mice, whereas the deficiency of HSF2 causes impaired spermatogenesis due to increased apoptosis of meiotic spermatocytes (Kallio *et al*., 2002; Wang *et al*., 2003). The fine cooperation between HSF1 and HSF2 is manifested by *Hsf1*$^{-/-}$*Hsf2*$^{-/-}$ double knockout mice where a total block in meiosis inhibits formation of post-meiotic germ cells (Wang *et al*., 2004). Thus, knocking out both HSF1 and HSF2 intensifies the meiotic defects in *Hsf2*$^{-/-}$ mouse to sterility, although the lack of HSF1 or HSF2 alone does not seem to remarkably impair male germ cell production. Intriguingly, HSFs have not been coupled to HSP expression during development and are suggested to be activated by distinct mechanisms in stress and development (Jedlicka *et al*., 1997, reviewed in Abane and Mezger, 2010).

2.4.5 HSFs in cancer progression

The importance of HSFs for core cellular processes is highlighted in the central role they display in cancer progression and metastasis (Dai *et al*., 2007; Mendillo *et al*., 2012; Santagata *et al*., 2011; 2013). A ground breaking study by Dai and coworkers revealed that mice devoid of HSF1 are protected from carcinogen-induced skin tumors (Dai *et al*., 2007; reviewed in Solimini *et al*., 2007). Later, large patient studies coupled high HSF1 activity to cancer metastasis and identified HSF1 as a major marker for poor prognosis (Mendillo *et al*., 2012; Santagata *et al*., 2011; 2013). HSF1 was, furthermore, shown to drive a complex transcriptional program in human breast cancer cell lines (Mendillo *et al*., 2012). Since cancer cells have defective cell cycle control, are highly proliferative, mutation prone, and live in crowded and oxygen-deficient environments (Hanahan and Weinberg, 2011), a commonly held view is that HSF1 enables survival and metastasis by allowing cancer cells to adapt to the hostile conditions. To date, the role of HSF2 in cancer development remains uncharacterized. Consequently, elucidating the detailed mechanisms, the interplay and the complete set of target genes for HSF1 and HSF2 is required for understanding the physiological functions of HSFs in distinct conditions.

3 AIMS OF THE STUDY

Prior to the work described in this thesis, HSFs were known as the master regulators of HSP expression upon stress, and their activity had been coupled to developmental processes. However, the genome-wide target genes for HSF1 and HSF2 and their mechanisms on targets beyond the *HSP*s remained unknown. In the focus of my PhD thesis was to identify the genome-wide target sites for HSF1 and HSF2 in male germ cell development and in cellular stress responses. Through characterization of the transcriptional programs that HSF1 and HSF2 mediate, the cellular and physiological functions of HSFs were studied, their collaboration in distinct physiological conditions analyzed and their importance in male gametogenesis and cellular stress responses assessed. From a transcription factors' point of view, the chromatin appears remarkably different depending on the cell type and the phase of the cell cycle. By utilizing the complex developmental process of male germ cells in mouse testis and analyzing HSF1 and HSF2 target loci in interphase and mitotic human cells, I addressed how the chromatin compaction affects the genomic distribution of HSF1 and HSF2 and the ability of cells to coordinate transcription.

The specific aims of this research were:

(1) To identify the genome-wide target genes for HSF1 and HSF2 in mouse spermatogenesis.

(2) To uncover the mechanisms by which HSF1 and HSF2 regulate male gametogenesis.

(3) To analyze HSF1 and HSF2 binding sites in the whole human genome by ChIP-sequencing.

(4) To characterize the HSF1- and HSF2-mediated transcriptional reprogramming in acute stress.

(5) To investigate the ability of HSF1 and HSF2 to interact with mitotic chromatin, and the capacity of mitotic cells to mount transcriptional responses when challenged by protein toxicity.

4 EXPERIMENTAL PROCEDURES

4.1 Cell culture, mitotic arrest, and heat shock treatments (III)

Human K562 erythroleukemia cells were maintained at 37°C in a humidified 5% CO_2 atmosphere and cultured in RPMI medium (Sigma) containing 10% fetal calf serum, 2 mM L-glutamate, 100 µg/ml streptomycin, and 100 U/ml penicillin. Cells were arrested in mitosis using double thymidine block that collects cells in early S-phase, and by subsequent nocodazole treatment that stalls the cells in G2/M (modified from Whitfield *et al*., 2000): The first thymidine treatment (2 mM, Sigma) for 16 hours was followed by wash with PBS, and progression of the cell cycle for 8 hours. The cells were collected in S-phase by the second thymidine treatment (2 mM) for 24 hours. After progression of cell cycle for 5 hours, the cells were arrested in mitosis with a 12-hour nocodazole (100 ng/ml, Fluka) treatment. The nocodazole was removed, and the cells were either harvested, heat treated in a 42°C water bath or cultured at 37°C to allow entry into G1 and S phases.

4.2 Determination of the cell cycle phase (III)

The cells were washed with PBS, fixed in 90% EtOH at -20°C for 24 hours, and the DNA was stained with propidium iodide (40 µg/ml, Sigma) in the presence of RNase A (10 µg/ml, Sigma) and Tween20. The DNA content of the cells was determined with fluorescence-mediated counting by FACSCalibur (BD Biosciences) and the histograms depicting proportions of cells with a given DNA content were generated with CellQuest Pro 6.0 (BD Biosciences) or Flowing Software 2.5 (Turku Centre for Biotechnology). Statistical analyses were conducted using unpaired student's t-test.

4.3 Depletion of HSF1 and HSF2 by small hairpin RNA (shRNA) (III)

HSF1, HSF2, or they both together, were depleted from the cells using shRNA constructs in pSUPER vectors (Oligoengine) as described (Östling *et al*., 2007). The oligonucleotides were specific for HSF1 (GCT CAT TCA GTT CCT GAT C) or HSF2 (CAG GCG AGT ACA ACA GCA T) and a scrambled sequence (GCG CGC TTT GTA GGA TTC G) was used as a control. The shRNA constructs were transfected into cells by electroporation (970 µF, 220 mV), after which the cells were incubated for 72 hours prior harvesting. Synchronization of cells to mitosis was initiated after a 7-hour recovery from the transfection, allowing for simultaneous sample preparation of the cycling and mitotic cells.

4.4 Mouse strains and tissue preparations (I-II)

The mice were kept in pathogen free facilities under a 12-hour light:12-hour dark cycle, and provided with complete pellet chow and tap water. The mice were handled according to

institutional animal care policies of the Åbo Akademi University (Turku, Finland) and Departmental Veterinary Office (Haute-Garonne, France). *Hsf1*$^{-/-}$ mice (McMillan *et al*., 1998) were bred form C57BL/6J strain, maintained in a mixed genetic background, and intercrossed with BALB/c strain. *Hsf2*$^{-/-}$ mice (Kallio *et al*., 2002) were obtained by heterozygous mating, and maintained in C57Bl/6N background. Wild type mice (wt) of the respective strains were used, except in genome-wide target gene assays in which testes from male hybrid mice of the B6129SF2/J strain were used. The mice were killed by CO_2 asphyxiation or cervical dislocation. Testes or epididymes were isolated form adult (60-80 days old) mice.

4.4.1 Comet assay and analyses of sperm head morphology (I-II)

Fragmentation of sperm DNA was analyzed with comet assay (modified from Sakkas *et al*., 2002). Sperm was released from the tail of epididymis (*cauda epididymis*), mixed with 0.8% low-melting point agarose in PBS (at 37°C), and laid on a 1.5% agarose layer that was pre-casted on Superfrost Plus microscope slides (Menzel). After solidification, a third layer of 0.5% low-melting point agarose was mounted on the slide. The slides were incubated for 1 hour in lysis buffer (2.5 M NaCl, 100 mM EDTA, 10 mM Tris, 1% Triton X, 10 mM DTT, pH10) at 4°C, and 1 hour in proteinase K (10 µg/ml) at 37°C. The sperm was subjected to electrophoresis (25V, 20 min) in Tris-Acetate-EDTA, pH7.3, and stained with SybrGreen (10 µg/ml, S7563, Invitrogen) for 15 minutes. Cover glass was mounted, and the proportions of intact *versus* fragmented sperm heads were blind counted from five wt and five knockout mice using fluorescent microscopy (Nikon 800i).

The sperm head morphologies were analyzed as described in Touré *et al*., 2004. Briefly, sperm from *cauda epididymis* was spread onto Superfrost Plus microscope slides (Menzel), fixed with 4% PFA, and stained with hematoxylin. Sperm from five wt and five knockout mice were analyzed under a light microscope and grouped into three categories: normal, slightly abnormal, and grossly abnormal.

4.4.2 Histochemistry and immunofluorescence of mouse germ cells (I-III)

Testes or epididymes were isolated, fixed in 4% PFA, embedded in paraffin, and cut to 4 µm thick sections. For histological analyses, the sections were stained with SybrGreen (10 µg/ml, S7563, Invitrogen), or hematoxylin and eosin, and investigated with light microscope. For immunofluorescence analyses, the sections were stained over night with primary antibodies against HSF1 (Sarge *et al*., 1993; Ab-4, Thermoscience), HSF2 (Sarge *et al*., 1993; Sistonen lab ab4506) or γH2AX (Millipore). IgG (sc-2027) was used as a negative antibody control. The secondary antibodies were conjugated to Alexa 488 or Alexa 568 (Invitrogen). For colocalization analyses, dye-swap of secondary antibodies was controlled. DNA was stained with Hoechst 33342 (H-1399, Molecular Probes). Images for all channels were sequentially captured from a single confocal section using a Zeiss Meta510 confocal microscope. The channels were merged and the colocalizations analyzed with ImageJ (Schneider *et al*., 2012).

4.4.3 Isolation of stages of seminiferous tubules (I-II)

Semininiferous tubules were released from testes of adult mice, and the stages identified under a phase-contrast microscope (Kotaja *et al*., 2004). For morphological analyses the tubules were examined and imaged using a phase-contrast microscope (Nikon 800i). For quantification of mRNA levels during male gametogenesis, the stages VII-VIII (dark zone; DZ), IX-XI (pale zone; PZ), XII-I (weak spot; WS) and II-VI (strong spot; SS) were collected.

4.5 Chromatin immunoprecipation (ChIP) (I-III)

ChIP was conducted as previously described (Östling *et al*., 2007) using ChIP-verified antibodies against HSF1 (Spa-901, Enzo) and HSF2 (Östling *et al*., 2007; Sarge *et al*., 1993). Normal rabbit IgG (Santa Cruz, sc-2027) was used as a negative, and AcH4 (Upstate, 06-866) as a positive antibody control. Decapsulated mouse testis, or $16x10^6$ cycling and $8x10^6$ mitotic human K562 cells were cross-linked in 1% formaldehyde, quenched in 2.5 mM glycine, and lysed in 1% SDS, 10 mM EDTA, 50 mM Tris pH8, and 1x protease inhibitors (Roche). The chromatin was sheared to 100-500 bp fragments with Bioruptor (Diagenode), and the lysate pre-cleared with 1:1 protein G sepharose beads in TE-buffer pH8. Immunoprecipitation was performed at 4°C over night, and the immunocomplexes were washed three times with wash buffer 1 (0.1% SDS, 1% Triton X, 2 mM EDTA, 150 mM NaCl, 20 mM Tris pH8.0), twice with wash buffer 2 (0.1% SDS, 1% Triton X, 2 mM EDTA, 500 mM NaCl, 20 mM Tris pH8.0), and three times with wash buffer 3 (20 mM Tris pH8.0, 1 mM EDTA, 10% glycerol). Cross-links were reversed at 65°C, and the DNA purified twice with phenol:chloroform:isoamyl alcohol (25:24:1), and once with chloroform:isoamylalcohol (24:1). After ethanol-precipitation, the DNA was air dried and dissolved in TE-buffer pH8. The PCR primers for the studied genomic loci are listed in the original publications (I, II and III).

4.6 Chromatin immunoprecipitation on promoter microarray (ChIP-chip) (I-II)

HSF1- or HSF2-bound chromatin regions were isolated in testis of three individual wt mice using antibodies against HSF1 (Spa-901, Stressgen) and HSF2 (Sarge *et al*., 1993). As controls, 20 µg of input sample, containing whole fragmented genome of the same tissue material, were used. Purified DNA of the ChIP and input samples were amplified using ligation-mediated PCR and the amplicons blunted with T4 DNA polymerase (New England Biolabs) in the presence of dNTP mix (Promega). The blunted DNA fragments were annealed to linker oligos (5'-GCG GTG ACC GGG AGA TCT GAA TTC-'3 and 5'-GAA TTC AGA TC-'3), and amplified with Taq and Pfu polymerases (New England Biolabs and Stratagen, respectively). An aliquote of the final DNA was verified for the fragment sizes on an agarose gel. The experimental HSF1 or HSF2 amplicons were labeled with Cy5, and the input control amplicons with Cy3, including one dye-swap for control. The labeled DNA fragments were hybridized to NimbleGen high-density oligonucleotide tiling arrays, containing 388 000 probes that cover 26 129 promoters of the mouse genome (genome build MM5, NCBI, May 2004, NibleGen Systems Inc.). On each

1.5 kbp promoter (spanning from -1200 bp to +300 bp), 15 probes with a length of 50 bp were designed with approximately 100 bp spacing, depending on the sequence composition of the corresponding regions (NimbleGen Systems Inc.).

The signal from HSF1- and HSF2-immunoprecipitated chromatin was compared to signal from control input chromatin according to standard operating procedures by NimbleGen Systems Inc. The intensity ratio was plotted against genomic position to identify promoter regions with HSF1 and/or HSF2 occupancy. The raw data is available at Gene Expression Omnibus (GEO) database (http:www.ncbi.nlm.nih.gov.geo) under accession GSE22492 for HSF1, and GSE9289 for HSF2.

The dye-specific variation in the two-channel raw data was corrected with Lowess normalization (Cleveland and Devlin, 1988), and ChIP-to-input log2-ratios were independently produced for all replicates. The target promoters were ranked according the average log2-ratio produced by all the probes on the promoter. The log2-ratios in the three replicates showed positive correlation yielding Pearson's correlation between 0.34-0.41. The reliability of the bound promoters was determined with RankProd package of R/Bioconductor (Breitling *et al*., 2004). The data was filtered with p-value < 0.005, and promoters occurring in at least two of the three replicates were included, resulting in a list of 742 target promoters for HSF1, and 546 target promoters for HSF2 in mouse testis. The intensities of HSF1 and HSF2 occupancies on the target promoters were visualized with SignalMap (NimbleGen Inc.).

4.7 ChIP-sequencing (ChIP-seq) (III)

HSF1- and HSF2-bound chromatin fragments were isolated in untreated and heat-treated (30 minutes, 42°C), cycling and mitotic K562 cells. Negative antibody controls (IgG) were included for each condition examined, and input samples of cycling and mitotic cells were used as indicators of genomic background. For sequencing, ten ChIP-replicates per sample were collected (see III: Figure S1, for a schematic presentation of sample preparation and data analysis) and purified using QIAquick DNA purification columns (Qiagen). Sequencing libraries were generated using New England BioLabs NEBNext sample preparation kits. Adapters, PCR and Index primers were from Bioo Scientific AIR DNA Barcodes kit. Template amplification and cluster generation were performed using cBot and Truseq SR Cluster kit. From each fragment, 36 nucleotides were sequenced with Illumina Genome Analyzer IIx using v5 TruSeq SBS sequencing kits. After quality trim and removal of duplicates, the sequenced reads were mapped to human genome (GRCh37/hg19) with Bowtie (Langmead *et al*., 2009). The peaks were called with the MACS 1.4 program (Feng *et al*., 2012) using input as control. The complete raw data is available at GEO database (http://www.ncbi.nlm.nih.gov/geo) under accession: GSE43579. To identify HSF1 and HSF2 target sites, a minimum fold enrichment of five times over input was set as a cut-off criterion, and any site that exceeded the cut-off in the negative IgG control sample was discarded.

4.8 Data analyses of HSF1 and HSF2 target sites (I-III)

4.8.1 Identification of targeted DNA sequences and genomic regions (III)

The consensus DNA sequences of HSF1 and HSF2 binding sites were identified in cycling and mitotic cells using MEME-ChIP (Machanic and Bailey, 2011). The consensus sequences were searched from 120 bp regions, centred on the summit points of HSF targets. HSF1 and HSF2 target loci were annotated to genomic regions using exon and intron coordinates provided by RefSeq, and by defining a core promoter to span from -1200 to +300 bp from TSS. From each identified peak, 50% of peak length was centered on the summit point. Peaks that fell on exon-intron boundaries were counted as exons. For visualization of HSF1 and HSF2 target sites in cycling and mitotic cells, Integrative Genomics Viewer (Robinson *et al*., 2011) was used.

4.8.2 Gene ontology analyses (I-III)

Biological processes associated with HSF1 or HSF2 target genes in mouse testis (p-value $\leq$ 0.001), and in cycling and mitotic human K562 cells (fold enrichment $\geq$ 5) were identified with DAVID (Dennis *et al*., 2003), which uses Fisher's exact test for calculation of the p-value for enriched gene ontology terms.

4.8.3 Characterization of chromatin state at HSF target sites (III)

The HSF1 and HSF2 target sites were searched for the presence or absence of RNPII or DNaseI hypersensentivity, using data provided by the ENCODE consortium (RNPII ChIP-seq: wgEncodeEH000529, Iyer Laboratory, University of Texas at Austin; DNaseI hypersensitive regions in G2/M: wgEncodeEH003472, Crawford Laboratory, Duke University). The occurrence of promoter-proximally paused RNPII was analyzed on selected HSF1 and HSF2 target genes utilizing density signal of RNPII (wgEncodeEH000616, Snyder Laboratory, Yale University). Promoters where RNPII-enrichment at least five times exceeded the overall RNPII density on the gene were designed as paused (pausing index $\geq$ 5, as in Adelman and Lis, 2012).

4.8.4 Defining the chromosomal distribution of HSF target promoters (I-II)

HSF1 and HSF2 target promoters in testis were searched for possible chromosomal accumulation using GeneMerge tool (Castillo-Davis and Hartl, 2003).

4.9 Analyses of RNA using quantitative real-time PCR (Q-RT-PCR) (I-III)

RNA was isolated using the RNeasy kit (Qiagen). 1 µg of total RNA was DNaseI treated (Promega) and reverse transcribed with Moloney Murine Leukemia Virus Reverse Transcriptase RNase H(–) (Promega). Real-time RT-PCR reactions were prepared and run as described earlier

(Östling *et al*., 2007) using ABI Prism 7900 (Applied Biosystems). The primers were purchased from Oligomer (Helsinki) and the probes from Roche Applied Science. Primer and probe sequences are listed in the original publications (I-III) and in Table 1. Relative quantity of the target gene RNA was normalized to a reference gene, glyceraldehyde-3-phosphate dehydrogenase (Gapdh) in K562 cells, and β-tubulin, acrosomal vesicle protein 1 (Acrv1) and profilin 3 (Pfn3) in the mouse testis. Fold inductions were calculated against the respective RNA level in non-treated cells, or against the respective RNA level in the wt testis. All reactions were made in triplicate for samples derived from at least three biological replicates. Standard deviations were calculated and are shown in the graphs. Independent student's t-test was used to determine the p-value when comparing RNA levels between scrambled-transfected and shRNA-transfected cells, or in between wt and knockout mice.

4.10 Protein analyses with Western blotting (I-III)

Cells were lysed with buffer C (25% glycerol, 20 mM HEPES pH7.4, 1.5 mM $MgCl_2$, 0.42 M NaCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT) and the protein concentration in the soluble fraction was measured using Bradford analysis. 20 µg of total protein was boiled in Laemmli sample buffer and subjected to SDS-PAGE in 8-10% acryl amide gels. The tissues were lysed and boiled in Leammli, and the proteins separated by SDS-PAGE using 8% or 14% acryl amide gels depending on the protein size. After transfer to nitrocellulose membrane (Protran nitrocellulose, Schleicher & Schuell), the proteins were analyzed with primary antibodies against HSF1 (Spa-901, Enzo; Sarge *et al*., 1993; Ab-4, Thermoscience), HSF2 (3E2, Upstate), β-actin (AC-40, Sigma) α-tubulin (ab57062, Abcam), β-tubulin (ab6046, Abcam), TNP1 (M-88), PRM1 (M-51), and PRM2 (M-107). The secondary antibodies were horseradish peroxidase-conjugated (GE Healthcare) and the blots were developed using an enhanced chemiluminescence method (ECL kit, GE Healthcare).

4.11 Isolation of polyubiquitinated proteins

Non-treated or heat-treated cells (0.5, 1, 2, or 6 hours at 42°C) were lysed (50 mM Tris-HCl pH7.5, 0.15 M NaCl, 1 mM EDTA, 1% NP-40, 10% glycerol) and the soluble fraction was collected by centrifugation. 2 mg total protein was diluted in 300 µl lysis buffer and incubated with agarose-conjugated Tandem Ubiquitin Binding Entities (agarose-TUBEs; Lifesensors) which bind to polyubiquitinated proteins with high specificity (Hjerpe *et al*., 2009). After centrifugation, the supernatant (unbound fraction) was collected and the agarose (bound fraction) was washed twice with Tris pH8. The samples were boiled in Laemmli sample buffer and the proteins were separated on SDS-PAGE using 4-20% gradient acrylamide gels (Bio-Rad). The gels containing the unbound fraction were stained with Coomassie Brilliant blue to detect the protein pool that was depleted form the polyubiquitinated proteins. The samples containing the bound fraction were analyzed by Western blotting to detect ubiquitinated proteins. The primary antibody was against mono- and polyubiquitinated proteins (Adi-Spa-200). Input samples were

taken prior to agarose-TUBE selection and analyzed for HSF1 (Adi-Spa-901) and β-tubulin (ab6046, Abcam) expression.

4.12 Oligonucleotide-mediated pull-down assay (III)

The oligonucleotide-mediated pull-down assay was performed as described previously (Anckar *et al*., 2006). The double-stranded biotinylated oligonucleotides (Oligomer) either contained an HSE (5′ -biotin - TCG ACT AGA AGC TTC TAG AAG CTT CTA G - 3′) or lacked an HSF-binding element, serving as a scrambled (Scr) control (5′- biotin - AAC GAC GGT CGC TCC GCC TGG CT - 3′). Buffer C extracts of 100-400 µg total protein were annealed to oligonucleotides (0.5 µM) in binding buffer (20 mM Tris pH7.5, 100 mM NaCl, 2 mM EDTA, 10% glycerol) containing salmon sperm DNA (0.5 µg/µl). The samples were pre-cleared, and the oligonucleotides precipitated with UltraLink streptavidin beads (Pierce) in binding buffer (1:1). Bound fractions were washed three times with binding buffer containing 0.1% Triton X. DNA-bound proteins were eluted with Laemmli sample buffer and detected by SDS-PAGE and Western blotting.

Transcript		Primers and probes (5' to 3')		Amplicon		
				start¥	end¥	lenght
NM_001349.2	DARS	Forward	GTTTTGTGGAAATCCAAACTCC	+825	+933	109nt
		Reverse	GGGGACTGAGCCAGGTATG			
		Probe	# 63			
NM_153201.1	HSPA8	Forward	TTTTTGTGGCTTCCTTCGTT	+30	+90	61nt
		Reverse	TCCCTTGGACATGGTTGC			
		Probe	# 36			
NM_002157.2	HSPE1	Forward	CAGTAGTCGCTGTTGGATCG	+574	+688	115nt
		Reverse	AGAACTACTTTGGTGCCTCCAT			
		Probe	# 7			
NM_001195802.1	LDLR	Forward	GGCTACAAGTGCCAGTGTGA	+944	+1032	89nt
		Reverse	AAGAAGAGGTAGGCGATGGAG			
		Probe	#11			
NR_028272.1	NEAT1	Forward	ATTGATGCCTGCAGATTGAA	+1270	+1338	69nt
		Reverse	GCCTGGAAACAGAACATTGG			
		Probe	# 63			
NM_002738.6	PRKCB	Forward	AGCGGTGCCATGAATTTG	+424	+519	96nt
		Reverse	TGTGGATCTTAAACTTGTGTTTGC			
		Probe	# 37			

Table 1. Primer and probe sequences. Probes are from universal probe library. ¥: indicates position from the transcription start site.

5 RESULTS AND DISCUSSION

5.1 HSF1 and HSF2 coordinate male germ cell development (I-II)

5.1.1 Disrupted organization of seminiferous epithelium in Hsf1$^{-/-}$ and Hsf2$^{-/-}$ mice

The role of HSF1 in male gametogenesis has remained controversial. Mice deficient of HSF1 have been reported to display normal spermatogenesis (Izu *et al*., 2004), and to contain seminiferous tubules with missing layers of germ cells (Salmand *et al*., 2008). To uncover the importance of HSF1 in male germ cell development, and to understand why simultaneous absence of HSF1 and HSF2 causes sterility (Wang *et al*., 2004), we investigated the cytological and morphological integrity of the seminiferous epithelium in HSF1-deficient mice. In stark contrast to the faithfully executed seminiferous cycle in wt testis, the detailed characterization of seminiferous tubules in *Hsf1*$^{-/-}$ mice revealed long stretches of abnormally pale regions that altered with dark zone-resembling regions (II: Figure 2B; Figure 8). From histological analyses we uncovered that the abnormal pale zones were devoid of both meiotic and post-meiotic germ cells and contained only spermatogonia (II: Figure 2A). Instead, the regions resembling dark zones also lacked spermatids but gained the light-reflecting appearance from apoptotic spermatocytes (II: Figure 2A-B). The absence of spermatocytes and spermatids in a subset of tubule regions suggested that *Hsf1*$^{-/-}$ mice display defects in meiotic and post-meiotic phases of gametogenesis.

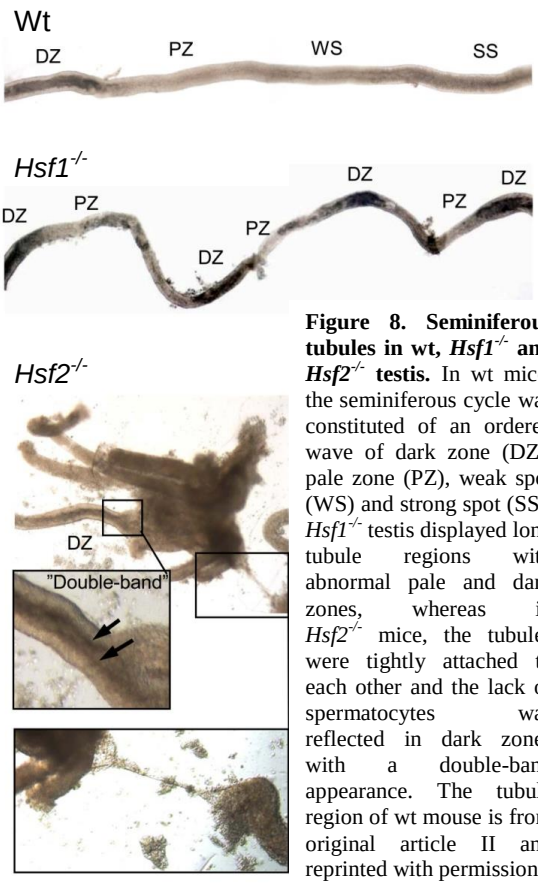


Figure 8. Seminiferous tubules in wt, *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ testis. In wt mice, the seminiferous cycle was constituted of an ordered wave of dark zone (DZ), pale zone (PZ), weak spot (WS) and strong spot (SS). *Hsf1*$^{-/-}$ testis displayed long tubule regions with abnormal pale and dark zones, whereas in *Hsf2*$^{-/-}$ mice, the tubules were tightly attached to each other and the lack of spermatocytes was reflected in dark zones with a double-band appearance. The tubule region of wt mouse is from original article II and reprinted with permission.

The gametogenic defects in *Hsf1*$^{-/-}$ mice profoundly differed from the phenotype of *Hsf2*$^{-/-}$ mice, in which the increased apoptosis in meiosis leads to tubule regions that lack the middle layer of spermatocytes (Kallio *et al*., 2002). Our studies confirmed the missing layer of spermatocytes in *Hsf2*$^{-/-}$ mice, but also identified tubule regions that were strongly attached to each other (Figure 8). The tight attachment of tubules argued for altered composition of the interstitium. HSF2 has been shown to regulate immunological genes in mouse testis and the *Hsf2*$^{-/-}$ testis to express reduced levels of interleukin-10 (IL-10) and increased levels of IL-6 (Vihervaara, 2007). Since male germ cells

enter meiosis after puberty, the highly specialized structures of spermatids are generated after the immune tolerance has been established. Several mechanisms contribute to generation of immune privilege in testis, including blood-testis-barrier, specific leukocyte composition, cytokine environment and reduced major histocombatibility complex I expression on spermatids (reviewed by Fijak and Meinhardt, 2006). The altered levels of immunological mediators in HSF2-deficient mice imply that HSF2 directs the expression of immune related genes in mouse testis, but could also reflect a systemic effect where the absence of HSF2 alters the composition of immunological cells in the organism. All in all, the distinct spermatogenic defects in HSF1- and HSF2-deficient mice emphasize the diverse functions that HSFs have in directing male germ cell development and argue for cumulative effects that cause sterility in the *Hsf1-Hsf2* double knockout mice.

5.2 HSF1 and HSF2 target genes in mouse testis (I-II)

To elucidate the mechanisms of *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ in male germ cell development, we characterized the target genes for HSF1 and HSF2 in whole mouse testis using promoter ChIP-chip assay. ChIP enables isolation of genomic regions that are bound by a given protein (Gilmour and Lis, 1984; 1985; Solomon *et al*., 1988) and identification of the isolated DNA regions by microarray is a powerful method for characterizing the protein binding sites in a genome-wide scale (Iyer *et al*., 2001; Ren *et al*., 2000; reviewed in Farnham, 2009; Gottardo, 2009). We were among the first laboratories to utilize ChIP-chip in tissue material and to characterize transcriptional regulation during germ cell development. Genomic regions that were bound by HSF1 or HSF2 were isolated from testis of three wt mice, and input material from the same tissues was used as indicators of the genomic background. The purified DNA regions were labeled and hybridized to NimbleGen's promoter microarray containing 15 probes per 26 129 promoters of the mouse genome. By comparing the HSF signal to the input background, and filtering the data with p-value 0.005, 742 putative target genes for HSF1, and 546 for HSF2 were identified (I: Table S1; II: Table S1). Selected HSF1 and HSF2 target genes were verified with ChIP-PCR using *Hsf1*$^{-/-}$ mice (II: Figure 3A) and an alternative HSF2 antibody (I: Figure 1B) as controls. Genes from different parts of the lists were selected for verification and included the members of the spermatogenesis associated glutamate (E)-rich protein (*Speer*) family, mitochondrial ferritin (*Ftmt*), and interleukins. Curiously, no classical *Hsp* genes were detected among the putative HSF1 or HSF2 target genes (I: Table S1; II: Table S1), implying that the transcriptional program that HSF1 and HSF2 regulate is highly dependent on the cell type, stimulus and developmental state. On the contrary, *Hsp70*, which is the classical HSF1-responsive gene in heat-treated cells, was used as a negative control for the HSF-ChIP in mouse testis (II: Figure 3A).

Comparison of the putative HSF1 and HSF2 target sites uncovered that promoters of only 55 genes were co-occupied by HSF1 and HSF2, corresponding to 10% of HSF1 and 13% of HSF2 targets (Figure 9; II: Table S3). HSF1 has been shown to facilitate the localization of HSF2 to target loci during stress (Östling *et al*., 2007; Sandqvist *et al*., 2009), but the radically different sets of target promoters in testis suggest that HSF1 and HSF2 carry out individual transcriptional functions in developing male germ cells. These distinct transcriptional programs that are

mediated by HSF1 and HSF2 could partially explain the different testicular phenotypes of *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice.

Figure 9. HSF1 and HSF2 target promoters in mouse spermatogenesis. Number of the putative HSF1 (dark circle) and HSF2 (light circle) target genes in mouse testis, revealing that only 55 promoters are co-occupied by both HSF1 and HSF2 in mouse testis. Note that any multicopy gene is counted as a single gene.

5.2.1 HSF1 and HSF2 occupy promoters of sex-chromosomal multicopy genes

Analyzing the chromosomal distribution of HSF1 and HSF2 target sites uncovered a striking occupancy of HSF1 and HSF2 on the Y chromosome (I: Table 1; II: Table S2). Of the 105 investigated promoters on the Y chromosome, over a third was bound by HSF1 (42/105) or HSF2 (35/105). The mouse Y chromosome consists of 95 Mbp of DNA that is mainly heterochromatin and contains long and short palindromic sequences that undergo intrachromosomal recombination during meiosis (David Page, Whitehead Institute, Massachusetts Institute of Technology Dept. of Biology, USA, personal communications; PhD thesis of Jessica Alföldi, Massachusetts Institute of Technology, USA, 2008). The short arm of the mouse Y chromosome (Yp) is 3 Mbp in size and contains eight genes in a single copy, including the sex-determinant *Sry*, and one duplicated gene. However, none of the genes on Yp were occupied by HSF1 or HSF2 (I: Table S1; II: Table S1). Instead, the long arm of Y chromosome (Yq), which is highly repetitive and palindromic, harbors multiple copies of Sycp3 like Y-linked (*Sly*), spermiogenesis-specific transcript on the Y (*Ssty1* and *Ssty2*) and serine-rich, secreted, Y-linked (*Srsy*) genes. Remarkably, all the HSF1 or HSF2 target sites on the Y chromosome were comprised of the distinct copies of the *Sly*, *Ssty1* and *Ssty2* promoters, or of hypothetical open reading frames which appeared to be included in the repetitive elements (I: Table S1; II: Table S1; see also Figure 10). *Srsy* was not included in the NimbleGen array, being discovered later, but by ChIP-PCR also *Srsy* was identified as a target gene for HSFs (II: Figure 3A).

Over the course of 300 million years, the X and Y chromosomes have diverged from a regular pair of autosomes to male-specific Y and to the X chromosome that occurs in both genders (reviewed in Bachtrog, 2013). Although genomic regions have been transferred between the sex chromosomes, only a 0.7-Mbp pseudoautosomal region (PAR) in mouse can homologously recombine between the sex chromosomes (Ohno *et al*., 1959; Perry *et al*., 2001; Tres, 1977). The X chromosome, however, also contains multicopy gene families that are higly expressed in testis and reside in repetitive and palindromic sequences that resemble the sequences of Yq. One of these X-chromosomal regions contains an estimate of 25-100 copies of Sycp3 like X-linked (*Slx*), which is a paralog for *Sly* (Mueller *et al*., 2008). *Slx* and *Sly* have arisen by a genomic transfer between the X and Y chromosomes and they both contain a Cor1 domain that can interact with chromatin (Touré *et al*., 2005). Intriguingly, several copies of *Slx* were identified as putative HSF1 and HSF2 target genes by ChIP-chip (Figure 10; I: Figure 1A and Table S1; II: Table S1) and verified as true targets by ChIP-PCR (I: Figure 1B; II: Figure 3A). Accumulation of HSFs on

the Yq, and on an 8-Mbp region containing multiple copies of *Slx*, was demonstrated with SignalMap (I: Figure S3), confirming the presence of HSFs on repetitive and palindromic regions of the sex chromosomes.

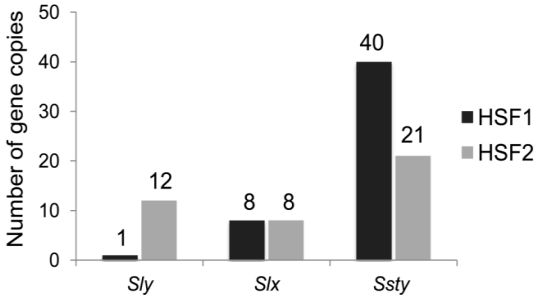


Figure 10. HSF1 and HSF2 occupancy at distinct promoters of X- and Y-chromosomal multicopy genes. The number of *Sly*, *Slx* and *Ssty* gene copies bound by HSF1 (black) and HSF2 (gray), indicating a preference for HSF1 to bind to *Ssty*, and HSF2 to bind to *Sly*, promoters.

The transcripts of *Slx* and *Sly* have been suggested to be participants of so called genomic conflict between the X and Y chromosomes, owing to their antagonizing effects on the MSCI and capacity to direct sex-ratio of offspring to favor the chromosome of the transcript's origin (Cocquet *et al*., 2012; Ellis *et al*., 2005). Given the antagonizing nature of the X- and Y-chromosomal multicopy transcripts, the binding of HSF1 and HSF2 to distinct copies of *Slx*, *Sly* and *Ssty1/2* promoters was investigated. Surprisingly, HSF1 prioritized the promoters of *Ssty*, whereas HSF2 accumulated on the promoters of *Sly* (Figure 10). The number of occupied *Slx* promoters was equal for HSF1 and HSF2 (Figure 10). The tendency of HSF1 and HSF2 to localize to distinct set of the multicopy gene promoters highlights their individual target genes in the mouse testis and rises up the possibility that the composition of HSF1 and HSF2 at the promoters directs the outcome of gene expression.

5.2.2 HSF1 and HSF2 collaborate on activating the expression of Sly and Ssty, but carry out different effects on the promoter of Slx

The X- and Y-chromosomal multicopy genes are expressed in round spermatids and crucial for the sperm head formation and fertility (Touré *et al*., 2005). To address whether HSF1 or HSF2 affect the expression of the multicopy genes, we quantified the mRNA levels of Ssty, Sly, and Slx in the testis of wt, *Hsf1*$^{-/-}$, and *Hsf2*$^{-/-}$ mice. Since both *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice have defects in meiosis and, subsequently, lowered spermatid count, the mRNA levels were normalized against spermatid-specifically expressed Acrv1 and Pfn3, allowing for quantitative comparison of the transcripts in the round spermatids. Lack of HSF1 or HSF2 caused decreased expression of Sly and Ssty mRNA (I: Figure 2B; II: Figure 3B), indicating cooperation of HSF1 and HSF2 in activation of the Y-chromosomal multicopy genes. Also the mRNA levels of Slx were decreased in the absence of HSF1 but, surprisingly, lack of HSF2 increased the amount of Slx mRNA to fourfold (I: Figure 2B; II: Figure 3B). The opposing effects of HSF1 and HSF2 on *Slx* uncovered diversity in the HSF1- *versus* HSF2-mediated transcriptional regulation and highlighted the need for a delicate determination of the HSF1-HSF2-composition at the target promoters.

5.3 HSF1 and HSF2 are expressed in a cell type-specific manner in mouse testis (I-II)

During homologous recombination, the X- and Y-chromosomes are silenced, covered with γH2AX and retained in the periphery of the cell (reviewed in Handel, 2004; Burgoyne *et al*., 2009). The sex-chromosomal silencing is partially relieved after meiotic divisions, allowing for multicopy gene expression in round spermatids (Ellis *et al*., 2005; Reynard *et al*., 2009). To characterize the spatio-temporal relationship of HSF1 and HSF2 with the multicopy genes, we isolated stages IX-XI, XII-I, II-VI, and VII-VIII of the seminiferous tubules in wt mice, and quantified the respective mRNA levels during the development of male germ cells (I: Figure 2A; Vihervaara, 2007). The mRNA of Slx, Sly, and Ssty showed similar expression patterns, since they all were abundant in stages II-VI that contain the round spermatids (I: Figure 2A). Expression of Hsf2 mRNA correlated well with that of multicopy genes since high levels of Hsf2 were detected during stages XII-I and II-VI, preceding and coinciding with the induction of the sex-linked multicopy genes (I: Figure 2A; Vihervaara, 2007). Unlike Hsf2, the mRNA levels of Hsf1 remained relatively unchanged throughout the stages of the seminiferous cycle (Vihervaara, 2007), suggesting constant HSF1 expression during male gametogenesis.

5.3.1 The subcellular localizations of HSF1 and HSF2 are carefully controlled during male germ cell development

The stages of seminiferous cycle contain multiple cell types that develop in a cyncytium (Oakberg, 1956; Eriksson, 1973). Thus, deciphering mRNA levels in isolated stages cannot determine the cell types that express the transcripts, or characterize whether the RNAs are distributed to neighboring cells via cytoplasmic bridges. I, therefore, analyzed the expression patterns or HSF1 and HSF2 proteins in mouse testis using immunohistochemistry and confocal microscopy.

The localization of HSF1 was carefully determined in every cell type of the male gametogenesis. The specificity of the HSF1 antibody (Sarge *et al*., 1993) was confirmed with the *Hsf1*$^{-/-}$ testis, where no signal was detected in the seminiferous epithelium (II: Figure 1A-B). Furthermore, the HSF1 expression in mouse germ cells was verified using another HSF1-specific antibody (Ab-4, Thermoscience) which reproduced the observed HSF1 localization with strikingly high accuracy (Figure 11a). HSF1 was not detected in spermatogonia that undergo mitosis, but it appeared in stage II-III spermatocytes at the zygotene-to-pachytene transition (II: Figures 1A and S1A). Throughout the pachytene, HSF1 was detected in distinct sub-nuclear regions (II: Figures 1A and S1B-F). At the transition from pachytene to diplotene, the levels of HSF1 rapidly increased and HSF1 localized throughout the nuclei of stage X-XI diplotene cells (II: Figures 1A and S1F). The expression of HSF1 remained high in the following meiotic divisions in stage XII, but it did not localize to the dividing chromatin (II: Figures 1A and S1A). In haploid cells, HSF1 was abundantly expressed in round spermatids troughout the stages I-VIII (II: Figures 1A and S1A-D), but its levels abruptly declined upon elongation of the spermatid nuclei (II: Figures 1A and S1E). In step 10-16 spermatids, no HSF1 expression was detected (II: Figures 1A and S1A-F).

The transient HSF1 expression from pachytene spermatocytes to early elongating spermatids is schematically summarized in II: Figure 5.

HSF1 has been described as a stable protein that is ubiquitously expressed in different cell types and tissues (Fiorenza *et al*., 1995; reviewed in Fujimoto and Nakai 2010). However, in testis, we uncovered striking variation in both the level and localization of HSF1, indicating the capacity and the need of the germ cells for spatio-temporal control of HSF1 expression. Intriguingly, high HSF1 expression in germ cells coincided with the phases of active transcription, chromatin modifications and exchange of histones to their testis-specific variants.

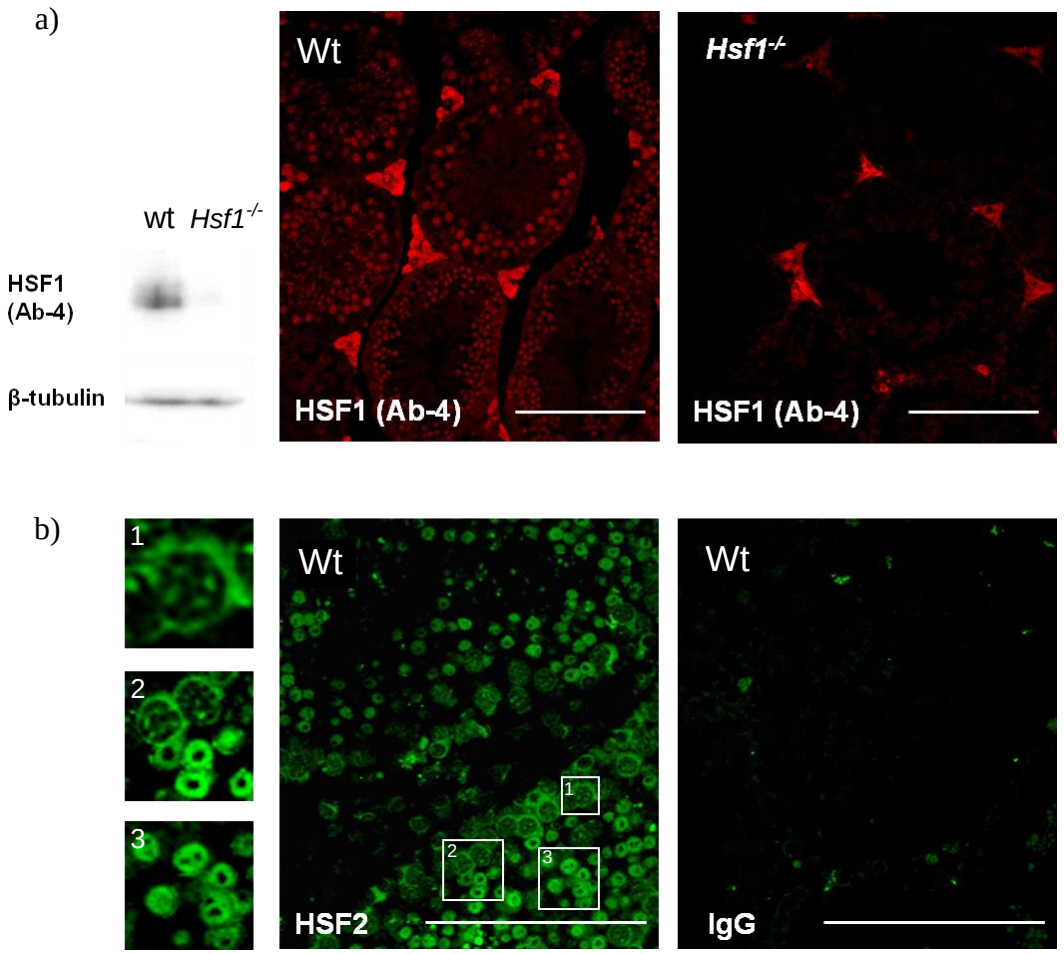


Figure 11. Localization of HSF1 and HSF2 in the seminiferous epithelium of mouse testis. a) Western blotting and confocal microscopy of wt and *Hsf1*$^{-/-}$ testes using Ab-4 antibody against HSF1. The detailed localization of HSF1 in mouse testis is illustrated in II: Figures 1 and S1. b) Cross sectioned wt testis stained with anti-HSF2 (left) and IgG (right) antibodies. The insets show HSF2 localization in pachytene spermatocytes (1 and 2) and round spermatids (2 and 3). Scale bars: 100 µm.

The expression of HSF2 has been characterized in rat and mouse testes (Alastalo *et al*., 1998; Björk *et al*., 2010), and our results were in accordance with the high HSF2 levels reported in pachytene spermatocytes and round spermatids (Figure 11b). By fluorescent confocal microscopy, we discovered localization of HSF2 mainly in the cytosol of pachytene spermatocytes but also in specific nuclear regions that, like rims, extended from the cytosol to the nucleus (Figures 11 and 12). Remarkably, in diplotene spermatocytes, HSF2 localized to the dividing chromatin (Figure 12). Although both HSF1 and HSF2 were present in germ cells that undergo active phases of transcription, their specific subcellular localizations indicate distinct spatio-temporal regulation and suggest factor-specific mechanisms for coordination of germ cell development. Particularly the diverging expressions of HSF1 and HSF2 in meiotic germ cells are in accordance with their distinct target genes (Figure 9; I: Table S1; II: Table S1) and with the different meiotic defects in *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice (Figure 8; II: Figure 2A-B). The high HSF1 and HSF2 expression in round spermatids correspond, in turn, with the shared spermatid-specifically expressed target genes. The different preferences of HSF1 and HSF2 for Y-chromosomal multicopy gene promoters (Figure 10) and opposing effects on Slx expression (I: Figure 2B; II: Figure 3B) suggest, however, factor-specific mechanisms for controlling the transcription of the multicopy genes.

5.3.2 HSF1 localizes to transcriptionally repressed sex-chromatin

HSF1 and HSF2 displayed both activating and inhibiting effects on the X- and Y-chromosomal multicopy genes (I: Figure 2B; II: Figure 3B), suggesting promoter-specific interaction with the transcriptional machinery or modification of the sex-chromatin environment. To understand the role of HSF1 on multicopy gene promoters, I investigated the HSF1 expression in respect to meiotic and post-meiotic sex-chromatin (II: Figures 4 and S1). During the homologous recombination in pachytene, HSF1 localized to the silenced sex-chromatin, as indicated by colocalization of HSF1 and γH2AX (II: Figures 4 and S1). During diplotene, HSF1 disappeared from the sex-chromatin and was not detected at the chromatin during meiotic divisions (II: Figure S1). In haploid spermatids, where the silencing of the sex-chromatin is partially relieved, the X or Y chromosome can be detected as a cloud-like structure next to the DNA-dense chromocenter (Namekawa *et al*., 2006). We did not detect HSF1 at the step 1 spermatids' sex-chromatin (II: Figure S1A), but observed HSF1 localization to the sex-chromatin in step 2-9 spermatids (II: Figure S1B-E). To the best of our knowledge, HSF1 is the first transcriptional regulator that has been shown to localize to the silenced sex-chromatin in male gametogenesis. Moreover, HSF1 and HSF2 are the first two transcription factors that have been shown to control the expression of X- and Y-chromosomal multicopy genes.

The localization of HSF2 in respect to the sex-chromatin is still inconclusively determined. However, we have detected HSF2 at the outer edges of the sex-body (Figure 12) using two antibodies (Sarge *et al*., 1993; Sistonen lab ab4506). This radically different pattern of HSF1 and HSF2 localization to the pre-meiotic sex-chromatin suggests distinct mechanisms for HSF1 and HSF2 during MSCI, which could modify the sex-chromatin for the post-meiotic expression of the sex-linked genes. Furthermore, our preliminary results show that also HSF2 can localize to the

post-meiotic sex-chromatin (data not shown), which argues for interplay between HSF1 and HSF2 at the X- and Y-chromosomes during the transcription of the sex-linked multicopy genes.

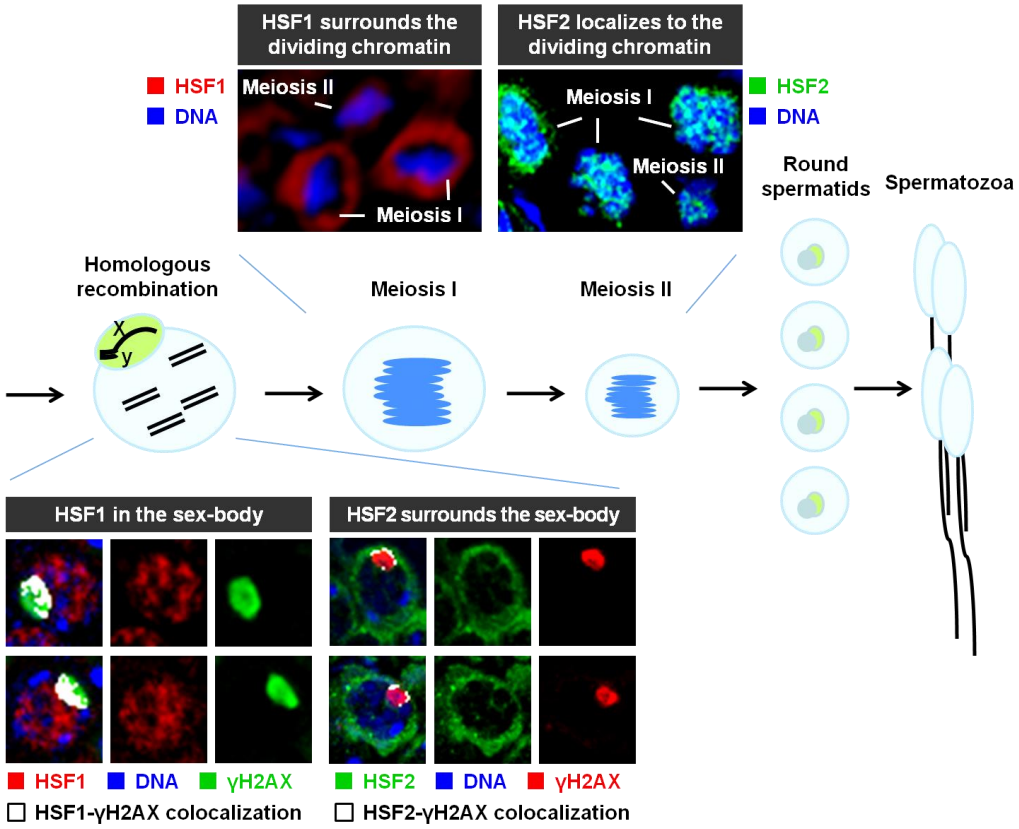


Figure 12. Subcellular localization of HSF1 and HSF2 in pachytene spermatocytes and in meiosis I and II. HSF1 is highly expressed in the nucleus of pachytene spermatocytes and localizes to the silenced pre-meiotic sex-chromatin. In contrast, HSF2 is abundantly expressed in the cytoplasma of the pachytene spermatocytes but detected also at the surroundings of the sex-chromatin. HSF1 is abundantly expressed in dividing cells but localizes outside the dividing chromatin. Instead, HSF2 colocalizes with the condensed chromatin at meiosis I and II.

5.4 HSF1 and HSF2 are required for correct DNA packing and sperm quality (I-II)

The haploid genome is tightly compacted and organized in the sperm head (reviewed in Braun, 2001; Miller *et al.*, 2010). The process of DNA organization is initiated already during embryogenesis and early spermatogenesis when somatic histones are changed to testis-specific histone variants. Later, the histones are replaced by TNPs in early elongating spermatids, and the TNPs are changed to PRMs in elongated spermatids (reviewed in Kimmins and Sassone-Corsi, 2005; Gaucher *et al.*, 2010). Since the X- and Y-chromosomal multicopy genes are involved in the DNA compaction (Ellis *et al.*, 2005; Riel *et al.*, 2013), we investigated the sperm head

morphology and the chromatin integrity in *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice (I: Figures 3 and 4B-C; II Figure 2C). Intriguingly, both HSFs were required for correct sperm head morphology as indicated by the increased proportion of sperm with deformed head structure in the absence of HSF1 or HSF2 (I: Figure 3; II: Figure 2C). Moreover, the lack of HSF2 caused increased fragmentation of DNA as detected by comet assay (I: Figure 4B-C). We also uncovered a profound retention of TNPs and incomplete incorporation of PRMs in *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice (I: Figure 4A; II: Figure 2D). Since genes encoding TNPs or PRMs were not among the putative HSF1 or HSF2 targets (I: Table S1; II: Table S1), HSFs are likely to carry indirect effects on TNP and PRM expression.

5.5 Conclusions and future perspectives on HSF1 and HSF2 in male gametogenesis (I-II)

The identification of HSF-bound promoters in mouse testis revealed a large number of putative target genes that HSF1 and HSF2 control in male germ cell development (I: Table S1; II: Table S1). As a result, HSF1 and HSF2 were shown to bind to a profoundly different set of target genes and display factor-specific mechanisms in male gametogenesis (Figures 8-10; II: Figure 2). Spermatogenesis provides a fascinating model of development during which cells and cellular processes are delicately coodinated in space and time. Although investigations in intact tissues give invaluable information about the mechanisms of life, the associations of a number of cell types also presents a major challenge: Testis contains a collection of cells including germ cells in distinct developmental states, Sertoli cells, Leydig cells, myoid cells and immune cells (Rouseaux *et al*., 2005). Therefore, identification of HSF1 and HSF2 target genes in a whole testis cannot determine the cell type(s) where HSFs bind to the target promoters. The presense of several cell types also imply that the strength of the binding signal becomes diluted and the binding intensities cannot be compared between target promoters or coupled to spatio-temporal organization of cellular processes. To overcome the restrictions of using whole testis, we utilized confocal microscopy for detailed characterizations of HSF1 and HSF2 in the cells of seminiferous epithelium (Figures 11-12; II: Figures 1 and S1). Although HSF1 and HSF2 were both expressed in pachytene spermatocytes and round spermatids, they displayed strikingly distinct subcellular localizations, implying different regulation of HSF1 and HSF2 and emphasizing their individual functions in male gametogenesis. The most surprising result was, however, the profound accumulation of HSF1 and HSF2 on the promoters of X- and Y-chromosomal multicopy genes (Figure 10; I: Figures 1 and S3; Table S1; II: Figure 3 and Table S1). These genes are expressed in round spermatids, regulate their own silencing and are coupled to defects in the sperm head morphology (Cocquet *et al*., 2012; Ellis *et al*., 2005; Reynard *et al*., 2009; Riel *et al*., 2013; Touré *et al*., 2004; 2005). Direct HSF-mediated regulation of the multicopy genes was indicated by the binding of HSFs to their promoters, localization of HSFs to the sex-chromatin and by the altered levels of *Stty*, *Sly* and *Slx* transcripts in *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice (Figures 9-12; I: Figures 1-2; II: Figures 1, 3-4, S1). Physiological consequence for the transcriptional regulation was suggested by the impaired incorporation of DNA packing protamines in HSF1- and HSF2-deficient mice (I: Figure 4A; II: Figure 2D), as well as by the

sperm head anomalies (I: Figure 3; II: Figure 2C), the severity of which correlated well in *Hsf1*$^{-/-}$ mice, *Hsf2*$^{-/-}$ mice and in mice that lack a 2/3 proportion of the multicopy gene containing Yp (Touré *et al*., 2004; 2005).

The long arm of Y chromosome has undergone high diversification during evolution and, consequently, no homologs of *Stty*, *Sly* or *Srsy* are detected in human. However, the highly repetitive and palindromic nature of the Yq is conserved in human and mouse (David Page, MIT, personal communications) and the palindromes of the Yq have been shown to undergo intrachromosomal crossing-over (Skaletsky *et al*., 2003). To date, the complete sequence of the mouse Y chromosome remains unpublished. However, a detailed characterization is provided in the PhD thesis of Jessica Alföldi (Alföldi, 2008), showing that the multicopy genes, targeted by HSF1 and HSF2, reside in the palindromes. Given that yH2AX marks chromatin regions that do not cross-over in meiosis (Turner 2007), the localization of HSF1 and HSF2 in respect to the yH2AX provides an intriguing point for speculation: If a subset of the palindromic sequences cross-over and are not marked by yH2AX, where do these regions localize. Could they be found in the interior of the sex-body where HSF1 is detected or would they perhaps locate to the outer edges that contain HSF2? In the future, the localization of multicopy genes, their promoters as well as transcripts in respect to HSFs should be studied using *in situ* hybridization.

Localization of HSF1 to the silenced sex-chromatin in pre- and post-meiotic germ cells reveals an unexpected aspect of HSF1 biology: So far HSF1 has been considered mainly as a potent transactivator that is rapidly recruited to open chromatin regions in stressed cells. Here, by identifying HSF1 in pre-meiotic sex-chromatin, where no transcription has been detected, HSF1 is suggested to carry out a function that diverges from its classical role as a transactivator. Intriguingly, HSF1 seems to be removed from the chromatin before meiotic divisions and to return to the sex-chromatin in post-meiotic cells (II: Figure S1) where the MSCI is partially releaved and the multicopy genes are expressed. The altered levels of multicopy transcripts in *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice suggest transcriptional regulation by HSFs. However, the action mechanisms of HSFs at the multicopy gene promoters, particularly in the context of pre- versus post-meiotic sex-chromatin remain uncharacterized. As a future challenge, elucidating the chromatin environment at the multicopy genes and visualizing the production of multicopy gene transcripts in wt, *Hsf1*$^{-/-}$ and *Hsf2*$^{-/-}$ mice, would give valuable information about the HSF1- and HSF2-mediated transcriptional regulation in germ cell development.

Taken together, this thesis identified several novel developmentally crucial target genes for HSF1 and HSF2 in mouse testis, characterized their distinct mechanisms to regulate spermatogenesis, and revealed their requirement for correct chromatin packing and male fertility. The analyses of HSF1 and HSF2 target genes in mouse testis demonstrated that the HSF-driven transcriptional programming can highly depend on the cell type and the developmental state. Furthermore, these studies uncovered that HSF2 localizes to the dividing meiotic chromatin, whereas HSF1 occupies silenced and compacted sex-chromatin before and after the meiotic divisions. The HSF1 and HSF2 localization to condensed chromatin in male germ cells could resemble the situation in stressed cells, where HSFs activate chaperone gene expression when the overall transcription is silenced.

5.6 Genome-wide characterization of HSF1 and HSF2 target sites in cycling and mitotic cells (III)

HSFs are the master regulators of transcription under protein-damaging conditions and capable of inducing gene expression in an environment where the overall transcription is silenced. The ability to maintain homeostasis is fundamental for all organisms; however, the studies on transcription upon heat stress have concentrated on a handful of chaperone genes, mostly in freely cycling cells (Ahlskog *et al*., 2010; Jedlicka *et al*., 1997; Östling *et al*., 2007; Wu, 1984; Yao *et al*., 2006). To understand how transcription is reprogrammed during acute stress, we set out to determine the genome-wide target sites for HSF1 and HSF2, to identify their transcriptional effects at the chromatin and to uncover their ability to direct gene transcription throughout the cell cycle progression. The target sites for HSF1 and HSF2 were characterized by ChIP-seq, which enables genome-wide mapping of protein binding sites in a high resolution and in an unbiased manner (Johnson *et al*., 2007; Park 2009; Pepke *et al*., 2009). Human K562 erythroleukemia cells were chosen as the model system, since in these cells the HSF1 and HSF2 levels and regulatory mechanisms are well characterized, cell synchronization is reliable in relatively large quantities, and chromatin landmarks have been identified by the ENCODE consortium (Consortium EP, 2011; Sarge *et al*., 1993; Sistonen *et al*., 1992; Östling *et al*., 2007). Upon heat stress, HSF1 and HSF2 have been shown localize to the *HSPA1*/*HSP70* promoter with the same kinetics, but after 30 minutes at 42°C, HSF2 is degraded from the *HSPA1*/*HSP70* promoter and from the K562 cells (Ahlskog *et al*., 2010). We, therefore, characterized the HSF1 and HSF2 target sites at 30 minutes exposure to 42°C when both factors are present in these cells. The acute stress also allowed for analysis of genomic loci that rapidly recruit HSFs, providing a view on the transcriptional processes that are instantly adjusted in the stressed cell. The efficiency of the cell cycle arrest was improved and its adverse effects reduced by thymidine-mediated collection of cells in the S-phase prior to nocodazole-induced arrest in G2/M (modified from Whitfield *et al*., 2000). Fluorescent-mediated analyses of the cellular DNA content confirmed the mitotic arrest (5% of cells in G1 and 85% in G2/M) as compared to freely cycling cells (45% in G1 and 15% in G2/M) (III: Figure 1A).

5.6.1 High resolution maps of HSF1 and HSF2 target sites in the human genome

ChIP-seq and annotation of the sequenced reads to the human genome generated high-resolution maps of HSF1 and HSF2 binding sites (III: Dataset S1, GEO access: GSE43579). In unstressed cycling cells, 45 target sites were identified for HSF1 and 148 for HSF2 (III: Figure 1B). Upon acute stress, both HSF1 and HSF2 were rapidly recruited to the genome as indicated by the 1 242 target sites for HSF1 and 899 for HSF2, and by the increased average fold enrichments of the targets (III: Figure 1B, Dataset S1). These results confirmed the previous indications that also HSF2 is a stress-responsive factor, and revealed its rapid binding to a wide set of target loci under protein-damaging conditions. Intriguingly, in mitotic cells, the capacity of HSF1 and HSF2 to interact with the condensed chromatin was dramatically different: HSF2 occupied 50 target sites in non-stressed mitotic cells and 545 sites upon acute stress (III: Figure 1B, Dataset S1). In

contrast, HSF1 interacted only with the promoter of *HSPA1B*/*HSP70.2* in the absence of stress, and with 35 target sites upon heat stress (III: Figure 1B, Dataset S1). At most of the target sites, HSF1 was capable for heat-induced binding in cycling cells only and no HSF1 binding was detected in mitosis (Figure 13a; III: Figure 1D, Dataset S1). However, the 35 target loci that HSF1 occupied in heat-treated mitotic cells included prominent enrichments on a number of promoters such as arsenite-inducible RNA-associated protein (*AIRAP/ZFAND2A*), *HSPA1A*/*HSP70.1*, *HSPD1/HSP60*, *HSPE1/HSP10*, *HSPH1*/*HSP110*, mitochondrial ribosome protein 6 (*MRPS6*), and *DNAJB6* (Figure 13b; III: Figure 1C and Dataset S1).

Previously, the binding of HSF2 to target genes upon stress has been considered to be HSF1-dependent (Östling *et al*., 2007; Sandqvist *et al*., 2009). The unbiased genome-wide analyses, however, identified genes that were specific for HSF1 or HSF2, including glucosidase beta acid (*GBA*), rho guanine nucleotide exchange factor 1 (*ARHGEF1*), myeloid/lymphoid or mixed-lineage leukemia (*MLL*) and lamin-beta (*LMNB*) (Figure 13c; III: Figure 1E-F). The promoter of *MLL*, as well as the introns of *LMNB* and malic enzyme 1, NADP(+)-dependent, cytosolic (*ME1*) manifest mitosis-specific binding of HSF2 (Figure 13c-d; III: Figure 1F). Intriguingly, at certain loci, the localization of HSF1 and HSF2 in respect to each other was dependent on the prevailing condition. For example at *ME1*, HSF1 and HSF2 colocalized at the promoter in heat-treated cycling cells, but occupied distinct loci, separated by 200 kbp, in heat-treated mitotic cells (Figure 13d).

5.6.2 HSF1 and HSF2 display distinct binding profiles in the human genome but bind to similar consensus DNA sequences

Comparison of the genomic coordinates of HSF1 and HSF2 target loci uncovered that in heat-treated cycling cells, HSF1 shared 54% of its target sites with HSF2, corresponding to 74% of HSF2 targets (III: Figure S2A). On the contrary, under optimal growth conditions and in mitosis, HSF1 and HSF2 displayed strikingly different binding profiles in the human genome (III: Figure S2A). Despite their distinct target sites, a virtually identical consensus HSE was identified for HSF1 and HSF2, revealing that these closely related factors recognize strikingly similar nucleotide sequences (III: Figure S2B). The virtually identical consensus sequence also implies that the DNA element alone cannot determine the binding of HSF1 *versus* HSF2 or define the composition of HSF1-HSF2 complex at the chromatin.

The importance of inverted nGAAn pentamers for HSF1 binding has been identified in diverse species including yeast (Hahn *et al*., 2004), fly (Gonsalves *et al*., 2011; Guertin and Lis, 2010) and human (Mendillo *et al*., 2012; Trinklein *et al*., 2004). By this first genome-wide analysis of HSF2 binding sites, we revealed that the striking conservation of the core HSE extends from HSF1 to HSF2 (III: Figure S2A). Previous footprinting studies have shown that HSF2 is able to bind to two inverted nGAAn pentamers, whereas HSF1 requires a more extensive HSE (Kroeger *et al*., 1993; Sistonen *et al*., 1992). The ability of HSF2 to stably bind to short HSEs is supported by genome-wide analyses of HSF1-specific and HSF2-specific target loci (Figure 14a), and by investigating the composition of HSEs on individual target sites (Figure 14b).

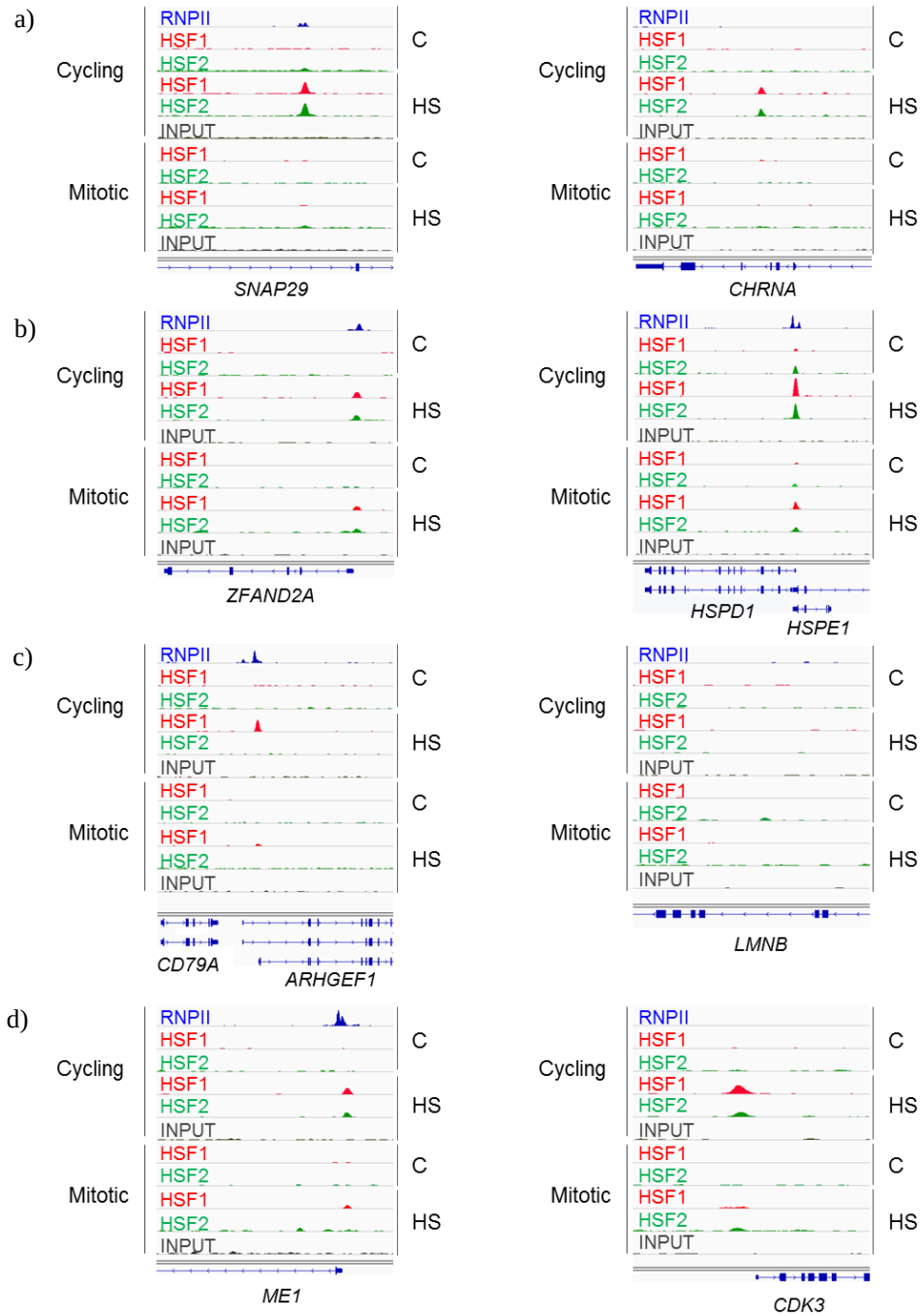
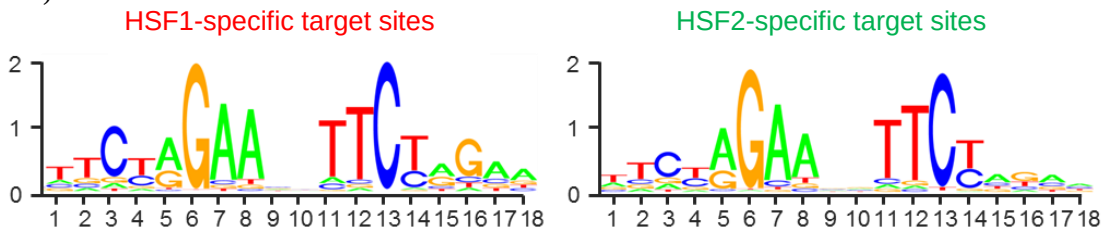


Figure 13. HSF1 and HSF2 enrichments in cycling and mitotic K562 cells. a) Promoters of synaptosomal-associated protein 29kDa (*SNAP29*) and cholinergic receptor, nicotinic, alpha (*CHRNA*) are bound by HSF1 and HSF2 in cycling cells only, whereas b) promoters of *ZFAND2A* and *HSPD1*/*HSPE1* are occupied by HSF1 and HSF2 in cycling and mitotic cells. c) *ARHGEF1* and *LMNB* illustrate HSF1 and HSF2-specific binding sites, respectively. d) Promoter-proximal loci of *ME1* and cyclin-dependent kinase 3 (*CDK3*) contain a different composition of HSF1 and HSF2 in cycling *versus* mitotic cells. C: control, HS: heat shock. Scale of the Y-axis: 0-100.

a)



b)

HSPA1A:	-183	AGTGAATCCCAGAAGACTCTGGAGAGTTCT . . . CGAAACCCCTGGAATATTCCCGACC	-64
AP4B1:	-115	GGCGCTGGAAATCCGGCGGAGGACCTAGAAGTAGCGGGAGCGATCTGTGGGCGGG	-61
DUSP1:	-792	AAAGCACTCCATGCCCTGAACTTTTCTAGAAGAATCCAGGCAGAACATTTGCGCAGG	-736
NEAT1:	-467	GCAGAGATCCCTCCGCCGCCGCCTGGAATTTTCCAGATGTCCTGCCGGCCCTTCTTT	-411
NUDC:	-263	AACCCTGCCAGATTTGAGTCCAGGGCTTTCTAGAAGTCTCCACGGCATCGGGTGACC	-207
UBB:	-260	CCAATCAGTGCCGGCGCTGCAAGGAAGTTTCCAGAGCTTTCGAGGAAGGTTTCTTCA	-224
UBC:	-351	GACACCCAGCTCCTCTGCAGTGTCTCCCCTGGAAAGTTCTCGAGCGTTCCCCAGCTT	-295
TAF7:	-380	GCACCAGATAGGCTAGACTCTAGGGAACATCCGAGATGTTTCCAGAAGCAAAGACG	-325
ARHGEF1:	+138	GGTGGGGAGAGTGCGGAGCCCGAGCGCGGAGGCTTCGGTTCCGGTGGCGGCGAT	+191
DARS:	+13	CCGCAGCTCTCGCGATCTTTCTGGAGCCGCACCTCCACGCGGAGTCCGAGCGCGTG	+68
STX6:	-88	GAAGTCGTGTTTCAGGCGGCGTTCGTGATTGGCCGCGGCCGAGGGTCCCGGGGGTG	-33
MLL:	-593	CTTTTATTCTAGAAGCGTT . . . ACTGGAAGCTGCTCTCCCGGGCTGTTCGATTCCCAG	-458
ADAM9:	+48645	ATTAGTTTCATGACTTTGG . . . AATGGTAGTTCTAGTTTTAGTTCTTTGAAAAATCTT	+48802
HIST2H2AC:	-406	ACGGCTTTCTTGGAGCCCT . . . AGGACCAATGAAAAACTCGTTCTTGACCCCACCCCT	-97

Figure 14. HSF2 can occupy short HSEs whereas HSF1 requires three or more inverted nGAAn pentamers for a stable binding. a) Genome-wide analysis of HSEs on HSF1-specific (left) and HSF2-specific (right) target loci using MEME-ChIP. The y-axis indicates degree of conservation of each nucleotide, the position of which is denoted at the x-axis. b) Examples of HSEs on selected HSF1-specific (red), HSF2-specific (green) and shared (blue) target loci. The inverted GAA and TTC nucleotides are highlighted and the numbers indicate the position from the TSS.

5.6.3. Distribution of HSF1 and HSF2 binding sites in the human genome

HSF1 and HSF2 binding have been extensively studied at certain gene promoters, such as *HSP70.1/HSPA1A* and *HSP70.2/HSPA1B*. To search for the distribution of HSFs in the human genome, the proportions of binding sites in distinct genomic regions were analyzed. Over half of the HSF1 and HSF2 binding sites in unstressed cycling cells occurred within intergenic regions (III: Figure S3A). Upon heat stress, however, the proportion of intergenic regions declined and the proportion of gene promoters and exons increased (III: Figure S3A). For example, in unstressed cycling cells 13% of HSF1 and 12% of HSF2 target loci were found within promoters, and upon heat stress, the share of promoters was 19% for HSF1 and 22% for HSF2

(III: Figure S3A). In mitotic cells, the few HSF1-occupied loci were highly concentrated on promoters (III: Figure S3A). In the human genome, less than three percent of the base pairs constitute promoters or exons (Human Genome Project), which indicates that HSF1 and HSF2 were clearly enriched on gene promoters and protein-coding sequences (see III: Figure S3A). Although the average fold enrichments of HSF1 and HSF2 targets were higher on promoter regions than on coding sequences (III: Figure S3B), prominent HSF1 and HSF2 enrichments occurred also within introns, exons, and intergenic regions as illustrated with introns of protein kinase C alpha and beta (*PRKCA* and *PRKCB*, respectively), exon of calcium-activated channel N1 (*KCNN1*) and intergenic region upstream of *HSPB1*/*HSP27* (III: Figure S3C).

5.7 HSF1 and HSF2 coordinate core cellular processes in cycling cells (III)

5.7.1 HSF1 and HSF2 bind to diverse groups of target genes and colocalize with RNPII at the target promoters

To identify the biological processes that HSF1 and HSF2 regulate *via* their target genes, gene ontology analyses with DAVID (Dennis *et al*., 2003) were performed. Since heat shock initiates a rapid transcriptional response, we investigated whether RNPII occupies the HSF target sites prior to stress and if the occupancy of RNPII is characteristic for gene groups that are associated with a specific function. Utilizing ChIP-seq data of RNPII, provided by the ENCODE consortium (wgEncodeEH000529, Iyer Laboratory, University of Texas), 90% of the HSF1- or HSF2-targeted loci on promoters were identified to contain RNPII (III: Figure 4A). At the gene bodies, approximately half of the HSF1 or HSF2 targets harbored RNPII prior to stress (III: Figure 4A).

Besides targeting RNPII-containing genomic regions, majority of target promoters for HSFs were bound by both HSF1 and HSF2. In heat-shocked cycling cells, HSF1 and HSF2 co-occupied promoters of chaperones and cochaperones, transcriptional and translational regulators, and mediators of the cell cycle progression (III: Figure 4A and Dataset S2). On the gene bodies, HSF1 and HSF2 targets were more diverse. While HSF1 occupied RNPII-containing loci of genes that mediate transcriptional repression, methylation and inhibition of metabolism, HSF2 localized to RNPII-deficient coding sequences of genes involved in activation of a variety of core cellular processes. These included genes that activate transcription, proliferation as well as metabolic processes (III: Figure 4A and Dataset S2).

5.7.2 HSF1 induces expression of polyubiquitin genes

Ubiquitin is a versatile signaling molecule that is required for directing damaged or misfolded proteins for proteasomal degradation (reviewed in Finley, 2009). In human, ubiquitin is encoded by four genes. *Ribosomal protein 27A* (*RPS27A*) and *ubiquitin A52 ribosomal protein* (*UBA52*) are so called monoubiquitin genes that encode precursors for one ubiquitin and one ribosomal protein. *Ubiquitin B* (*UBB*) and *ubiquitin C* (*UBC*) are polyubiquitin genes that code for three and nine ubiquitin proteins, respectively (reviewed in Kimura and Tanaka 2010; schematically

illustrated in III: Figure 3B). Each of these protein precursors are co-translationally cleaved into individual ubiquitin or ribosome proteins as soon as the polypeptides emerge from the ribosome. Heat-inducible expression of polyubiquitin genes has been previously demonstrated in several species and the *UBB* and *UBC* promoters are known to contain HSEs (Bond and Schlesinger, 1989; Finley *et al*., 1987; Fornace *et al*., 1986; Lee *et al*., 1988). In accordance, we detected HSF1 and HSF2 occupancy at the polyubiquitin gene promoters upon heat stress, whereas the monoubiquitin genes lacked both HSFs (III: Figure 3A). Furthermore, only polyubiquitin genes were found to be heat-induced (III: Figure 3C). Analyses of the expression of polyubiquitin genes in the presence and absence of HSF1 and HSF2 uncovered that HSFs are direct regulators of *UBB* and *UBC* and that the induction of polyubiquitin mRNA is dependent on HSF1 (III: Figure 3D).

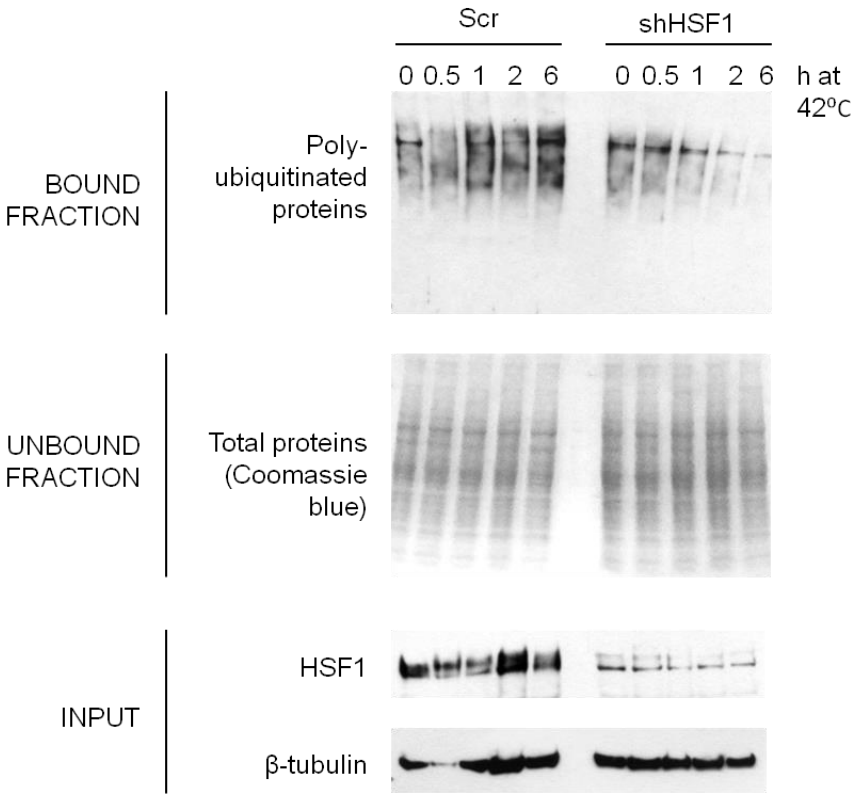


Figure 15. Protein ubiquitination in heat-treated human cells. Protein-damaging stress leads to accumulation of polyubiquitinated proteins in Scrambled (Scr) transfected cells, whereas in cells deficient of HSF1 (shHSF1), the heat-induced accumulation of polyubiquitinated proteins is remarkably diminished. No clear difference in the amount of non-ubiquitinated proteins is detected in Scr *versus* shHSF1 transfected cells. BOUND FRACTION contains polyubiquitinated proteins that were isolated using tandem ubiquitin binding entities (TUBEs). UNBOUND FRACTION contains proteins that did not bind to the TUBEs. INPUT indicates the expression of HSF1 and β-tubulin in total lysates.

Generation of multiple ubiquitin molecules per translation cycle permits rapid and effective production of ubiquitin in cells that express UBB and UBC. To investigate whether cells that lack HSF1 display reduced capacity to ubiquitinate the proteome, we analyzed the levels of polyubiquitinated proteins and accumulation of non-ubiquitinated proteins in heat-treated cells. To isolate polyubiquitinated proteins from the total protein pool, agarose-conjugated TUBEs that have high affinity for ubiquitin chains (Hjerpe *et al.*, 2009), were used. The pool of soluble proteins that did not bind to TUBEs was collected for analyses of non-ubiquitinated proteins. Increased protein ubiquitination has been shown upon heat stress, which we confirmed in Scr transfected cells (Figure 15). Intriguingly, in cells deficient of HSF1, the polyubiquitination of the proteome was markedly reduced (Figure 15, bound fraction). The inability for efficient ubiquitination in the absence of HSF1 could cause accumulation of misfolded and aggregated proteins that are unmarked by ubiquitin. However, no clear difference in the amount of soluble proteins that were unmarked by ubiquitin was observed during 6 hours of heat stress (Figure 15, unbound fraction). Protein ubiquitination is a multistep process which involves chaperone machinery, ubiquitin activating and conjugating enzymes as well as ubiquitin ligases (reviewed in Finley, 2009; Hartl *et al.*, 2011). Although our results do not allow dissecting the causal relationships of ubiquitin expression, chaperoning capacity and the cascades of protein ubiquitination, these results argue that HSF1-induced transcription of *UBB* and *UBC* balances the increased need for protein degradation under stress conditions.

5.7.3 HSF1 and HSF2 define the expression of chaperone complexes in stressed cycling cells

The ChIP-seq enabled an unbiased analysis of HSF1 and HSF2 distribution at every chaperone gene in the human genome, and investigations on the specific constituents of the chaperone complexes that are expressed in heat-stressed cells. By ChIP-seq, occupancy of HSF1 and HSF2 was detected on 70% of *HSP*, 90% of *chaperonin* and 13% of *DNAJ* genes upon acute heat stress (III: Tables I and SI). Promoter-proximally paused RNPII was originally identified at the *HSP70* promoter and is currently known to poise several *HSP* genes for rapid or synchronous activation (Rasmussen and Lis, 1993; Rougvie and Lis, 1988; reviewed by Adelman and Lis, 2012). To elucidate the mechanisms for HSF-mediated transcriptional regulation, we investigated the presence of paused RNPII at the chaperone genes using existing ChIP-seq data (wgEncodeEH000616, Snyder Laboratory, Yale University). Although ChIP cannot determine transcriptional processes, recent global-run-on sequencing (GRO-seq) has indicated that the most promoter-associated RNPII is transcriptionally engaged (Core *et al.*, 2012). This allowed us to search for promoters where RNPII enrichment at least five times exceeds the overall signal on the gene and consider them to be transcriptionally paused (III: Tables 1 and S1).

A majority of the HSF-targeted chaperone genes was bound by both HSF1 and HSF2 upon heat stress. Moreover, the HSF1 and HSF2 binding sites were localized on promoters that contained paused RNPII (III: Tables I and SI). Exceptions were the small HSPs *HSPB2*/*HSP27-2*, *HSPB5*/*CRYAB* and *HSPB9* where no paused RNPII was found (III: Table I). Intriguingly, beyond the classical foldases, such as *HSP70* and *HSP40*, HSF1 and HSF2 bound to promoters of

HSPH1/*HSP110*, *HSPH2*/*HSP105*, and *DNAJB6* which have recently been identified as disaggregases or inhibitors of aggregate formation in human cells (Hageman *et al*., 2010; Rampelt *et al*., 2012). Moreover, HSF1 and HSF2 targeted several chaperonin genes including *HSPE1*/*HSP10* and members of the *CCT* family (III: Table 1). The impact of HSF1 and HSF2 on transcription was investigated by depleting the cells from HSF1 or HSF2 and analyzing the target gene mRNA expression (III: Figures 2A-B, S4A). In freely cycling cells, *HSPH1*/*HSP110*, *HSPH2*/*HSP105*, *DNAJB6* and *CCT1* were identified as novel HSF-regulated chaperone genes whose heat-induced mRNA expression was dependent on HSF1 (III: Figures 2B and S4A). Particularly, the expression of HSPH1/HSP110 was rapidly increased to 7-fold, while the levels of HSPH2/HSP105, DNAJB6 and CCT1 doubled during a 2-hour heat shock (III: Figure 2B). Also the promoter of *HSPA8*/*HSC70*, was targeted by HSF1 and HSF2 (III: Table 1) and a 2-fold induction was found (Figure 16). The profound enrichment of HSF1 and HSF2 at the promoter of *HSPE1*/*HSP10*, did not lead to induction of HSPE1 mRNA during 2 hours of heat stress (Figure 16). Intriguingly, Mendillo and coworkers have shown a considerable increase in HSPE1/HSP10 mRNA expression after a 1-hour heat shock followed by a 2-hour recovery (Mendillo *et al*., 2012). These results indicate either cell type-specific regulatory mechanisms, or that HSPE1/HSP10 is required after stress when the cellular functions are readjusted for conditions that support proliferation.

5.7.4 HSF1 is required for heat-induced cochaperone expression

DNAJB6 has been identified as the most potent chaperone in inhibiting protein aggregation, and HSPH1/HSP110 and HSPH2/HSP105 have been shown to direct the HSP70-HSP40 machinery to solubilize and refold aggregated proteins in metazoan cells (Hageman *et al*., 2010; Rampelt *et al*., 2012). These results indicate a considerably wider repertoire of HSF-induced chaperones than earlier anticipated and highlight the importance of trancriptionally controlling the whole chaperone machinery. Since virtually every function of a chaperone complex is directed by cochaperones (Craig *et al*., 2006; Li *et al*., 2012; Zuehlke and Johnson, 2010, and references therein), we investigated the localization of HSF1 and HSF2 to the cochaperone genes and analyzed the presence or absence of paused RNPII.

Besides binding to a large number of chaperone genes, HSF1 and HSF2 co-occupied many cochaperone genes, which have not previously been demonstrated to be under the control of HSFs (III: Figure 2C and Dataset S1). Especially, HSF1 and HSF2 showed a strong tendency for heat-induced binding to cochaperone gene promoters that contained paused RNPII (III: Figure 2C). These genes included cochaperones that direct several aspects of the HSP70 and HSP90 complexes, including the client recognition, folding efficacy and release of the client (III: Figure 2C). We investigated the transcriptional impact of HSF1 and HSF2 on HSP90 cochaperones *AHSA1* (encoding AHA1), *CDC37*, and *PTGES3* (encoding p23) (III: Figure 2D). AHA1 is the most potent activator of the HSP90 protein complex, whereas CDC37 and p23 are involved in client protein recruitment and maturation, and are inhibitory to the HSP90 ATPase cycle (Ali *et al*., 2006; Li *et al*., 2012; Retzlaff *et al*., 2010; Taipale *et al*., 2012). A 3-fold heat-induced increase of AHA1 mRNA and a 2-fold increase of p23 mRNA were observed during a 2-hour

heat shock in the presence of HSF1 (III: Figure 2D). The levels of *CDC37* were not affected by the heat stress or depletion of HSFs (III: Figure 2D). Curiously, *CDC37* was the only cochaperone gene where HSF1 and HSF2 binding was not detected at the promoter but, instead, at the first intron (III: Figure 2C). As conclusion, our results revealed that HSF1 and HSF2 control a wide repertoire of chaperone and cochaperone genes upon heat stress, determining the composition of whole chaperone machinery that maintains protein homeostasis in stressed cells.

5.7.5 HSFs coordinate the expression of transcriptional regulators, cell cycle determinants and signaling molecules

In heat-stressed cells, the overall transcription and translation are silenced, cytoskeleton reorganized and the cell cycle progression stalled (reviewed in Kourtis and Tavernakis, 2011; Richter *et al*, 2010; Toivola *et al*., 2010). Despite the silenced transcriptional environment, acute heat stress induced binding of HSF1 and HSF2 to hundreds of target sites in freely cycling cells. These HSF target sites contained genomic regions that code for transcriptional and translational regulators, determinants of the cell cycle progression, signaling molecules and chromatin organizers (III: Figure 4A, Dataset S2). We verified the mRNA levels of a set of HSF1 and HSF2 target genes, including transcriptional regulators TATA-box associated factor 7 (*TAF7*), *MLL* and CCCTC-binding factor (*CTCF*). TAF7, which has been shown to function as a transcriptional check-point for RNPII-initiated transcription (Gegonne *et al*., 2006; 2008), showed a 3-fold HSF1-dependent increase upon heat stress (III: Figure 4B). Instead, no heat-induced expression of chromatin modifier and insulator CTCF was detected (III: Figure 4B). *MLL* is a trithorax homolog, methyltransferase and a mitotic factor that is involved in the epigenetic maintenance of transcriptional memory (Blobel *et al*., 2009; Tyagi *et al*., 2007). Curiously, *MLL* was identified as an HSF2-specific target gene (III: Figures 1F, Dataset S1) and the absence of HSF2 changed the kinetics of MLL expression upon stress (III: Figure 4B). Of the translational components, the expression of eukaryote elongation factor 1 gamma (*EEF1G*), aspartyl-tRNA synthetase (*DARS*) and *MRPS6* was analyzed. While the levels of EEF1G mRNA doubled during stress, the expression of DARS and MRPS6 remained unchanged (Figure 16; III: Figure 4B).

HSF1 and HSF2 bound also key regulators of mitosis, including nuclear distribution C homolog (*NUDC*) and barrier to autointegration factor 1 (*BANF1*). NUDC is involved in nuclear movement and kinetochore assembly (Aumais, *et al*., 2003; Nishino *et al*. 2006) and showed a 1.5-fold induction upon heat stress (III: Figure 4B). BANF1, which mediates the assembly of nuclear envelope and progression of cytokinesis (Bradley *et al*., 2005; Haraguchi *et al*., 2008), displayed elevated expression in the absence of HSF2 (III: Figure 4B). The diverge effects of HSFs on *NUDC* and *BANF1* indicate distinct roles and promoter-specific mechanisms for HSF1 and HSF2 when coordinating regulators of mitosis.

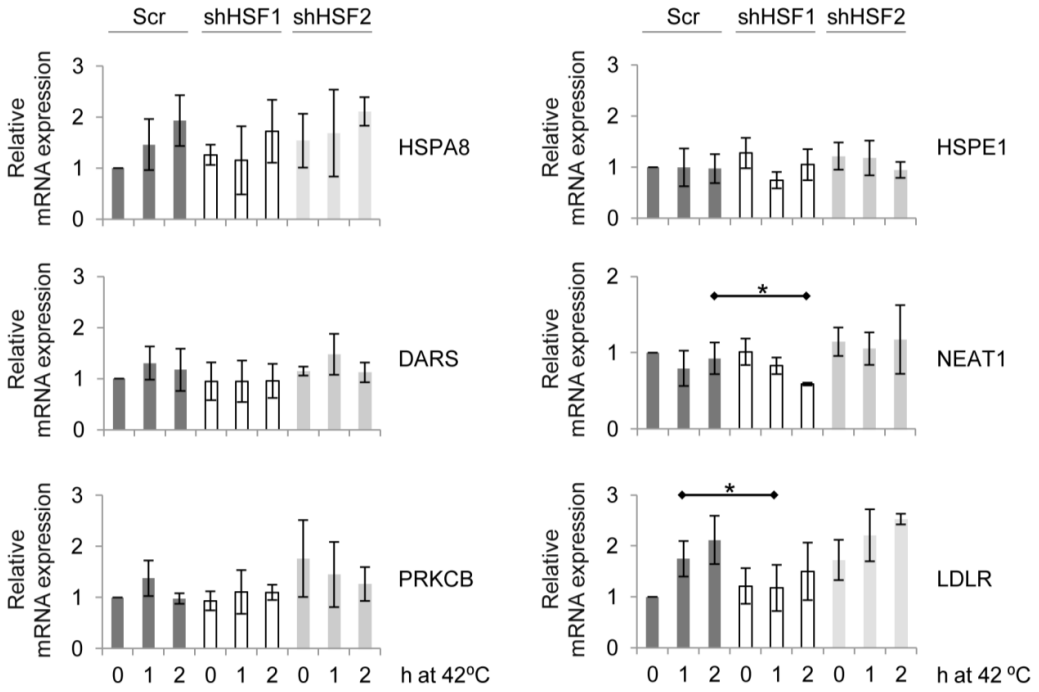


Figure 16. Expression of HSF1 and HSF2 target genes during 2 hours of heat stress. HSF1 or HSF2 was depleted from the cells and the mRNA levels of target genes were determined during 2 hours of heat stress. Scale bars denote standard deviations. *P < 0.1.

Cellular stress responses are highly coordinated processes in which multiple cellular functions are reorganized and synergistically regulated for an efficient protection against the hostile conditions. HSF1 and HSF2 binding was detected at coding sequences of several genes that encode core components of signaling cascades, such as *PRKCA* and *PRKCB* (III: Figure S3C, Dataset S1). The members of PRKC family are transducers of signaling cascades and have vital functions in several physiological processes, including responses to osmotic chock, mechanical shearing and elevated temperature (reviewed in Mellor and Parker, 1998). Despite clear binding, neither HSF1 nor HSF2 was identified to induce the PRKC mRNA expression during the 2-hour heat treatment (Figure 16; III: Figure 4B). However, lack of HSF2 allowed for more than a 2-fold increase in the PRKCA levels in non-stressed cells (III: Figure 4B). The function of PRKCs upon elevated temperatures has been suggested to occur independently of HSE-mediated responses (Kamada *et al*., 1995), but the binding of HSF1 and HSF2 to introns of *PRKC* genes (III: Figure S3C, Table S1) indicates a cross-talk between these distinct pathways. At low density lipoprotein receptor (*LDLR*), HSF1 and HSF2 colocalized at two distinct loci; at a promoter region ~400 bp upstream of the TSS and at an exon-intron-boundary several kbp downstream of the TSS. The two distinct binding sites at a single gene could indicate regulation of several transcriptional steps or reflect a higher-order structure where two distinct loci colocalize with a single HSF cluster. *LDLR* encodes a cell surface protein that mediates cholesterol uptake and lipid metabolism (reviewed in Jeon and Blacklow, 2005) and, surprisingly, heat stress caused a 2-fold increase in

LDLR mRNA levels (Figure 16). The observed induction was dependent on HSF1 but not HSF2, which is in line with previous reports where HSF1 has been identified to control metabolic processes of healthy and malignant cells (Dai *et al*., 2007; Hahn *et al*., 2004; Jin *et al*., 2011).

Besides binding to protein-coding genes, HSF1 and HSF2 displayed prominent enrichments on non-protein coding genes, such as long non-coding RNA00304 (*LINC00304*) and nuclear paraspeckle assembly transcript 1 (*NEAT1*) (III: Dataset S1). *LINC00304* contained the most enriched target loci detected for HSF1 and HSF2, but we were not able to generate primers to analyze the expression of LINC00340 mRNA. We, however, analyzed the expression of NEAT1, which mediates the formation of nuclear structures that are involved in lncRNA-mediated regulation of gene expression (reviewed in Naganuma and Hirose, 2013). Upon heat stress, the levels of NEAT1 remained unchanged, but depletion of HSF1 caused a 40% reduction in its expression (Figure 16), suggesting that HSF1 could be involved in maintaining the basal level of NEAT1 transcription during heat exposure.

In conclusion, our results revealed that HSF1 and HSF2 coordinate the expression of a versatility of genes in freely cycling human cells that are exposed to protein-damaging conditions. HSF1 was required for the heat-induced expression of mitotic factors, metabolic mediators, as well as transcriptional and translational regulators, whereas HSF2 was mainly involved in controlling transcription in non-stressed conditions (Figure 16; III: Figures 2-4). In fact, depletion of HSF2 did not hamper the heat-induced expression of any of the studied genes, rather allowed for faster induction of the mRNA expression. However, the absence of HSF2 disrupted the expression kinetics of MLL (III: Figure 4B) and caused more than a 2-fold increase in mRNA levels of BANF1 and PRKCA under optimal growth conditions (III: Figure 4B). The HSF2-mediated effects at chromatin might partially be regulated by HSF1, particularly at the target loci where these factors colocalize. All-in-all, our results revealed promoter-specific mechanisms for HSF1 and HSF2 and uncovered their distinct mechanisms for orchestrating transcriptional processes.

5.8 HSF1 has dramatically impaired capacity for chromatin binding and transcriptional activation in mitosis (III)

5.8.1 Mitosis inhibits transcriptional activation and renders chromatin inaccessible for HSF1

In heat-treated mitotic cells, HSF1 was unable to bind to 1207 loci that it occupied in cycling cells exposed to the same conditions (III: Figure 5B). Among the genes that lacked a clear HSF1 binding in mitosis were the polyubiquitin genes (III: Figure S4B), and subsequently, UBB and UBC expression remained unchanged in stressed mitotic cells (III: Figure S4C). However, certain genomic regions did recruit HSF1 in heat-treated mitotic cells (Figure 13b; III: Figures 1B-C and 5A-B, Datasets S1 and S2). These HSF1 target loci consisted mainly of promoters of chaperones and translational components, including *HSPA8*/*HSC70*, *HSPE1*/*HSP10*, *HSPH1/HSP110*, *HSPH2*/*HSP105*, *NUDC*, *DNAJB6*, *MRPS6* and *ZFAND2A*/*AIRAP* (III: Dataset S1). We depleted the cells of HSF1 or HSF2 (III: Figure 5C), confirmed that the lack of HSFs did not interfere with

synchronizing the cells in mitosis (III: Figure 5C-D), and investigated whether HSFs can induce transcription in cells where the chromatin is condensed for division (III: Figures 5E and S4D). Despite clear heat-induced binding of HSF1 and HSF2, no transcriptional activation of *HSPA8*/*HSC70*, *HSPH1/HSP110*, *HSPH2*/*HSP105*, *NUDC*, *DNAJB6* or *MRPS6* was detected in mitosis (III: Figures 5E and S4D, Dataset S1). Considering that HSF1 was a potent activator of *HSPA8*/*HSC70*, *HSPH1/HSP110*, *HSPH2*/*HSP105*, *NUDC*, and *DNAJB6* in cycling cells (Figure 16; III: Figures 2 and 4B), these results indicate that the occupancy of HSF1 at target promoters does not induce transcription in the mitotic chromatin environment, where the promoters likely lack the transcriptional machinery and the condensed chromatin generates barriers for transcriptional initiation and elongation (reviewed in Alabert and Groth, 2012; Delcuve *et al*., 2008).

5.8.2 Mitotic cells are highly susceptible to heat-induced stress

Mitotic cells are highly sensitive for stress, which is manifested by increased apoptosis and chromosomal abnormalities upon heat exposure (Hut *et al*., 2005; Martínez-Balbás *et al*., 1995). The impaired ability of mitotic cells to maintain homeostasis upon exposure to protein-damaging stress could be explained by the limited DNA-binding activity of HSF1 and the subsequent inability of the cells to express genes that are vital for counteracting the stress conditions. To confirm the mitotic susceptibility to heat stress in the human K562 cells, we compared the proportion of cells with fragmented DNA, indicative of cell death, in cycling and mitotic cells (III: Figure S4E). While the proportion of cycling cells with fragmented DNA remained constant during 6 hours of heat stress, the percentage of mitotic cells with fragmented DNA doubled during a 2-hour, and tripled during a 6-hour exposure to heat (III: Figure S4E). These results confirm the impaired thermotolerance in mitotic cells, and indicate that the binding of HSF2 to mitotic chromatin cannot rescue the inability of HSF1 to transactivate gene expression and to protect the proteome.

5.9 HSF2 directs gene transcription throughout the cell cycle progression (III)

5.9.1 HSF2 is an epigenetic regulator of gene expression throughout the cell cycle progression

The radically limited capacity of HSF1 to interact with the chromatin in mitosis was contrasted by the ability of HSF2 to bind to chromatin in both non-stressed and heat-stressed mitotic cells (III: Figures 1B and 5B, Dataset S1). Particularly, the 545 target loci that HSF2 occupied in stressed mitotic cells illustrate the profoundly distinct features of these closely related transcription factors. A staggering finding was that HSF2 localized to a distinct set of targets in cycling and mitotic cells (III: Figure 5A-B, Datasets S1 and S2), which indicates cell cycle phase-dependent plasticity of HSF2 in transcriptional regulation. Besides binding to the promoters of chaperone genes, HSF2 localized to genes encoding cell cycle regulators, translational

components and mediators of cell adhesion (III: Figure 5A, Dataset S2). However, the inability of HSF2 to induce gene expression in cycling and mitotic cells (Figure 16; III: Figures 2-5) raised a question on the role of HSF2 at the target chromatin.

The HSF2-mediated effects on transcription were mainly detected in non-stressed cells where the deficiency of HSF2 allowed for increased target gene expression (III: Figures 3D and 4B). These findings suggest that HSF2 either inhibits transcription or modifies the chromatin environment at the target loci. Particularly the expression of MLL was affected by the absence of HSF2 both in cycling and mitotic cells (III: Figures 4B and 5E). To elucidate the role of HSF2 in mitosis, when the overall transcription is silenced (Alabert and Groth, 2012; Delcuve *et al*., 2008), we analyzed the expression kinetics of MLL after releasing the non-stressed cells from mitotic arrest and allowing the cells to enter G1 and S phases (III: Figure 6A). In cells expressing HSF2, the levels of MLL mRNA gradually doubled during the cell cycle progression from mitosis to G1 and S phases (III: Figure 6A). These results are in agreement with a previous study where MLL protein levels were shown to increase in post-mitotic HeLa cells (Liu *et al*., 2007). On the contrary, cells that were deficient of HSF2 were unable to induce the MLL expression during cell cycle progression from mitosis to G1 (III: Figure 6A). Especially in early G1 (after a 2-hour progression from mitosis), the levels of MLL were significantly lower (p-value 0.012) in HSF2-deficient cells than in cells where HSF2 was expressed (III: Figure 6A). As a control gene, we examined *DUSP1*, which was bound by HSFs in cycling cells only (III: Figure 1D). *DUSP1* is heat-inducible (Keyse and Emslie, 1992) and, accordingly, DUSP1 was expressed at a constant level during the cell cycle progression, both in the presence and absence of HSF2 (III: Figure 6A). Although we cannot rule out the possibility that HSF2 regulates *MLL* also outside mitosis, HSF2 binding to the *MLL* promoter in non-stressed mitotic but not cycling cells (III: Figure 1F) suggests mitosis-specific regulation. In conclusion, the ability of HSF2 to bind to the *MLL* promoter in mitosis (III: Figure 1F), and to regulate its post-mitotic expression (III: Figure 6A), indicates the involvement of HSF2 in restoring gene expression after the mitotic silencing.

5.10 HSF2 interacts with mitotic and meiotic chromatin while HSF1 is efficiently excluded from the dividing genome (III)

5.10.1 HSF2 binds to open and closed chromatin during mitosis

The ability of HSF2 to bind to mitotic chromatin (III: Figures 1B, 5B and S2A) argued that a great number of HSEs are accessible for transcriptional regulators during cell division. To analyze the chromatin environment at the HSF target sites, we compared the coordinates of HSF1 and HSF2 binding sites to DNaseI hypersensitive regions that have been mapped in mitotic K562 cells (wgEncodeEH003472, Crawford Laboratory, Duke University). Open chromatin was found to be a prerequisite for HSF1 binding, since 97% of its mitotic targets occurred within the DNaseI hypersensitive regions (III: Figure 6B). In contrast, 42% of HSF2 target loci in mitosis

were found within the DNaseI hypersensitive sites, indicating that HSF2 is capable for interacting with open and closed chromatin (III: Figure 6B).

5.10.2 Mitotic chromatin is inaccessible for HSF1

The localization of HSF2 to open and closed chromatin did not, however, explain why HSF1 was absent from the 195 DNaseI hypersensitive regions that contained HSF2. To address whether HSF1 was excluded from the mitotic chromatin via inhibiting its intrinsic DNA-binding capacity, we analyzed whether HSF1 can associate with an introduced HSE-containing oligonucleotide. By utilizing streptavidin-mediated pull-down of biotin-conjugated oligonucleotides, both HSF1 and HSF2 were detected in complex with the introduced HSE in heat-treated cycling and mitotic cells (III: Figure 6C). Curiously, HSF2 was able to bind to the HSE also in untreated cells (III: Figure 6C) and its levels declined in mitosis (III: Figure 6C). Hyperphosphorylation of HSF1 has been associated with its heat-induced activation and considered to be an indicator of transcriptionally active HSF1 (reviewed in Anckar and Sistonen, 2011). Surprisingly, hyperphosphorylation of HSF1 was detected in mitosis regardless of the presence or absence of stress stimulus (III: Figure 6C). Although the specific amino acid residues that are post-translationally modified in mitosis remain to be characterized, these PTMs could be involved in coordinating the activity of HSF1 in dividing cells. Taken together, the ability of HSF1 to bind to exogenous DNA demonstrates that HSF1 indeed is capable of binding to DNA in mitosis. However, its inability to interact with the genome in mitosis implies that the access of HSF1 to the dividing chromatin is dramatically impaired.

5.10.3 HSF2 localizes to dividing chromatin during meiosis I and II

The ability of HSF1 to bind to introduced HSE, and its inability to localize to target sites in mitosis resembles the situation in male germ cells: In mouse testis, HSF1 has been shown to bind to introduced HSE upon heat treatment (Sarge *et al*., 1995), but has not been detected to bind the *HSP* promoters in stressed or non-stressed male germ cells (I: Table S1; II: Table S1; Nakai *et al*., 2000; Vydra *et al*., 2006; Widlak *et al*., 2007). Our studies revealed high HSF1 expression outside the dividing chromatin both during meiosis I and meiosis II of mouse gametogenesis (Figure 12; II: Figures 1 and S1; III: Figure 6D). On the contrary, abundant HSF2 localization was detected at the chromatin in meiotic divisions (Figure 12; III: Figure 6D). The localization of HSF2 but not HSF1 to mitotic and meiotic chromatin indicates highly coordinated division of labor between these closely related transcription factors and suggests a common mechanism that selectively allows HSF2 to bind to dividing chromatin from which HSF1 is efficiently excluded.

5.11 Conclusions and future perspectives on HSF1 and HSF2 as regulators of cellular stress responses in freely cycling and mitotic human cells (III)

HSFs are well characterized orchestrators of HSP expression under protein-damaging conditions (reviewed in Richeter *et al*., 2010). Heat stress, however, affects virtually every molecule in the cell, induces adaptation of several elemental cellular processes and, therefore, suggests a wider transcriptional adjustment than the mere induction of protein folding chaperones. To characterize the HSF1- and HSF2-mediated transcriptional programming in stressed human cells, we identified the genome-wide target sites for HSF1 and HSF2 and utilized computational and biochemical analyses to investigate HSF-mediated transcriptional programming. Our results revealed a myriad of genes that HSF1 and HSF2 bind to and a versatility of processes that they regulate in stressed freely cycling cells (summarized in Figure 17; original data in Figures 13-16; III: Figures 1-4, Datasets S1-2). Owing to the highly dynamic nature of the chromatin, transcription factors meet remarkably different landscapes in different phases of the cell cycle. In G1 phase, the chromatin is relatively accessible and also S phase provides a window of opportunity for transcription factors to reach the DNA. Instead, mitosis has been described as transcriptionally inert (Alabert and Groth, 2012; Delcuve *et al*., 2008; Probst *et al*., 2009), although the mitotic chromatin has been suggested to be dynamic (Nishino *et al*., 2012) and accessible for a subset of transcriptional regulators (Chen *et al*., 2005; Kadauke and Blobel, 2012). Since heat poses a severe threat to the cell, we questioned whether mitotic cells are able to mount transcriptional responses when challenged by proteotoxicity. Strikingly, the ability of HSF1 to bind to the chromatin was nearly abolished in mitosis, but hundreds of target sites were identified for HSF2 (III: Figure 1 and Dataset S1). Despite HSF1 or HSF2 binding, no evidence for heat-induced transcription in mitosis was found, supporting the prevailing view that mitosis renders chromatin transcriptionally silent (see Figure 17 for a model of HSF1 and HSF2 in cycling and mitotic cells). The efficient inhibition of HSF1 to bind to the chromatin in mitosis could reflect the severity of the threat to the integrity of the DNA if transcription was allowed in mitosis. Alternatively, the chromatin landscape in mitosis, which likely is devoid of RNPII and PIC, could lack the signals that are required for homing HSF1 to the correct genomic regions.

Given the inefficiency of HSF1 to localize to the dividing chromatin, the discovery of hundreds of target sites for HSF2 was puzzling (III: Figures 1 and 5). Particularly, the highly similar consensus HSE for HSF1 and HSF2 (Figure 14; III: Figure S2B) suggested their similar recruitments to the target sites. However, the diverging DNA-binding profiles of HSF1 and HSF2 (III: Figure S2) indicated that factors beyond the DNA sequence contribute to the rapid localization of HSFs to their target loci. To date, the mechanisms of how HSF1 *versus* HSF2 recognize the target loci, and how the stoichiometry of HSF1-HSF2-complex is determined, remain unknown. The cells' need for factor-specific regulation of HSF1 and HSF2 is partially explained by the profoundly different means by which HSF1 and HSF2 coordinate transcription: Whereas HSF1 has been described as a potent transactivator by a number assays in different laboratories, the action mechanisms of HSF2 have remained enigmatic. We and others have shown that the lack of HSF2 causes diminished expression of target genes in development, but the

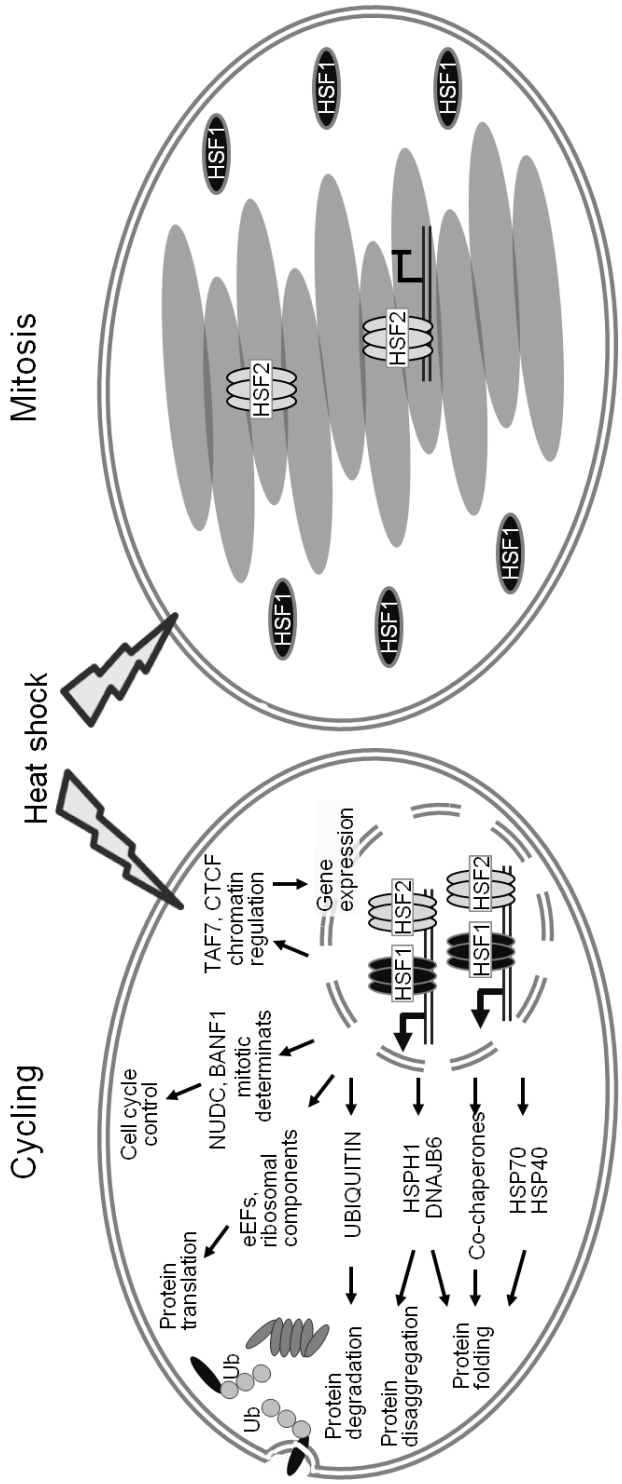


Figure 17. Localization and target genes of HSF1 and HSF2 in heat-stressed cycling and mitotic cells. In freely cycling K562 cells (left) HSF1 and HSF2 bind to a wide set of target genes and regulate the expression of chaperones, co-chaperones, ubiquitin, components of translation machinery, regulators of cell cycle progression and coordinators of chromatin. In dividing K562 cells (right), HSF2 localizes to the condensed chromatin and binds to target genes, whereas HSF1 is efficiently excluded from the dividing chromatin. Instead of activating transcription, HSF2 is able to mark target genes for rapid activation in G1. Ub: ubiquitin.

developmental models have not allowed for determining whether HSF2 is a direct transactivator, or whether HSF2, instead, affects the chromatin landscape and recruitment of other factors that are the primary activators of gene expression. Indeed, we were not able to couple HSF2 to transcriptional activation when measured by the levels of target gene mRNA (Figure 16; III: Figures 3-4). Instead, several lines of evidence points to a role for HSF2 as a regulator of the chromatin environment. For example, the mitosis-specific binding of HSF2 to the promoter of *MLL* suggests that HSF2 affects the epigenetic state and allows for a rapid transcriptional reactivation in early G1 (III: Figures 1F and 6A). Similar results were obtained by analyzing the kinetics of HSPA1/HSP70.1 expression in freely cycling cells: Given that HSF2 is degraded from the HSP70 promoter after exposure to 30 minutes heat shock at 42°C (Ahlskog *et al*., 2010), the HSF2-dependent peak on HSP70 mRNA levels after 2 hours at 42°C (III: Figure S4F) argue that HSF2 carries its effects at the chromatin during acute stress which, subsequently, contributes to the target gene expression in the later phases. This type of chromatin modification could well be a mechanism for HSF2 also in developmental processes, during which the epigenetic features of terminally differentiated cells are determined.

The results presented in this thesis uncover the versatility of HSFs as transcriptional coordinators and identify their profoundly distinct mechanisms as regulators of gene expression. However, several essential aspects of HSFs and transcriptional regulation in general remain to be established. Since cellular stress responses have mainly been studied in cultured cell lines, comprehension on how different cell types respond to stress in intact organisms remains to be established. Several studies in the worm *C. elegans* point to a systemic control over cellular stress reponses and HSF (reviewed in Gidalevitz *et al*., 2011), but whether HSFs are under organismal control also in mammals is completely unexplored. The rapidly induced binding of HSFs to their target genes upon stress also raises a question on the mechanisms by which transcription factors are able to localize a couple of nucleotide long DNA elements in the growded nucleus of a human cell. Indeed, the cellular stress responses in the context of nuclear arcitechture and the genome-wide changes in the chromatin landscape during and after stress remain unknown. Sequencing of the human genome (Venter *et al*., 2001) initiated an avalanche of genome-wide investigations that have identified the multitude of coding and non-coding sequences and revealed the importance of the epigenetic state and the coordination of the genome by regulatory factors. Attempts to understand the full complexity of the human genome include the ENCODE project where chromatin marks such as DNaseI hypersensitive sites, histone modifications and occupancies of regulatory proteins are characterized in different cell types. For the future, the ever growing number of chromatin landmarks that have been identified in K562 cells will allow for detailed analysis of the chromatin that is targeted by HSF1 and HSF2. The characterization of chromatin landscape, together with *de novo* motif searches and proteomic analyses, is likely to provide a comprehensive view on the transcriptional regulation at the heat-responsive genomic regions, as well as elucidate how HSF1 and HSF2 are recruited to their target loci. To date, the ENCODE consortium has focused solely on non-stressed conditions and, therefore, unraveling the global changes in the chromatin state in response to stress will require mapping of the sequential recruitment and removal of DNA-bound proteins, histone modifications and interacting chromosomal regions during stress response. This daunting task could be facilitated

by the recently described high-throughput ChIP-seq (Yan *et al*., 2013) or digital genomic footprinting (Hesselberth *et al*., 2009), where hundreds of chromatin landmarks in a single cell type were mapped or DNaseI hypersensitivity sites utilized to reveal footprints of DNA-binding proteins, respectively.

The finesse of ChIP-seq is its independency from probe design and hybridization which, in ChIP-chip, limit the genomic regions included in the studies and blur the resolution of obtained protein binding sites. For example, the promoter ChIP-chip studies that we conducted in mouse testis (I and II) were devoid of all the information of HSF1 and HSF2 binding sites in exons and introns, RNA coding genes, as well as intergenic regions. Indeed, the unbiased mapping of protein-binding sites by ChIP-seq revealed that the majority of HSF1 or HSF2 binding sites in K562 cells locate to genomic regions that do not code for proteins. These regions could be distal enhancers or other regulatory elements, or code for RNA species that recently were identified in the human genome (The ENCODE Consortium, 2013). Appreciation on how HSF1 and HSF2 regulate the non-protein coding regions would require sequencing of all the RNA-species present in the cell, following the elongation of RNPII at the genome, and mapping the changes in chromatin landscape during the heat shock response. A subset of HSF1 and HSF2 binding sites are likely to occur in chromatin regions that colocalize for an efficient and coordinated usage of cellular machinery, such as transcription apparatus. 5C chromatin capture (reviewed in Dekker *et al*., 2013) and visualization techniques could be highly informative in revealing the three-dimensional structure and interacting genomic regions upon stress. Besides extending the investigations on cellular architecture to membranes, cytoskeleton and organelles, imaging techniques would also allow for analyses of the repetitive DNA sequences which currently are neglegted from ChIP-seq and other genome-wide studies that require mapping of sequences to a reference genome. In summary, the rapidly induced transcriptional reprogramming upon stress will continue to foster ground breaking insights into the mechanism of transcriptional coordination and provide a robust model for investigating gene expression programs in space and time, and in the context of the dynamic cellular architechture.

CONCLUDING REMARKS

Since the start of my PhD thesis project, the advances in genome-wide analyses have fueled a change from a gene-centric view to analyses of transcriptional programs and chromatin states in a whole genome scale. We joined the development by characterizing HSF1 and HSF2 target genes in mouse spermatogenesis as well as in freely cycling and mitotic human K562 cells. The results presented in this thesis and its original publications, together with results from several other laboratories, have led to a change in our understanding of HSF biology. For several decades, the rapid induction of HSPs upon stress has provided the transcriptional researchers excellent tools for investigating inducible transcription and fostered ground breaking insights into the mechanisms of gene expression. The recent years, however, have enabled broadening the view to genome-wide transcriptional responses that HSFs mediate in different physiological models. The genome-wide analyses by us and others have confirmed the previous observations that HSFs control a variety of physiological processes and that they coordinate a distinct set of target genes in development, cancer progression and cellular stress responses (summarized in Figure 18). Taken together, my PhD thesis has investigated fascinating processes of life, including the delicate choreography of male germ cell development and the rapidly induced transcriptional reprogramming in acutely stressed cells. HSF1 and HSF2 have provided an excellent model of closely related transcriptional regulators whose cellular functions have proven to be highly intertwined but molecular actions exerted through profoundly distinct mechanisms. The future years hold great promises and challenges for understanding transcriptional regulation and cellular processes in the context of cell type, cellular architecture and intact organisms.

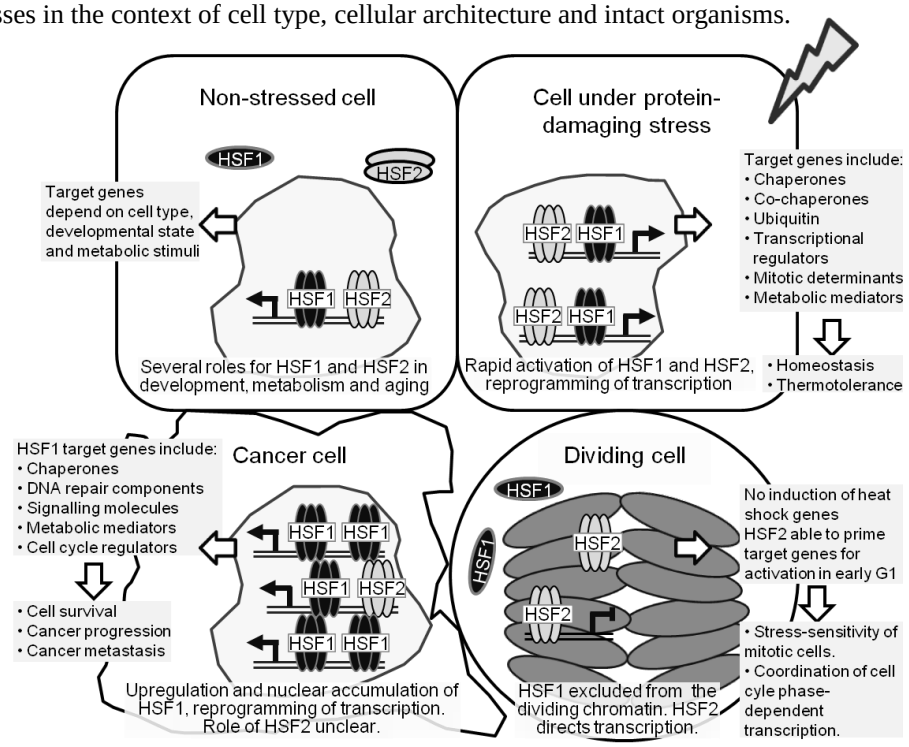


Figure 18. Functions of HSF1 and HSF2 in distinct physiological processes. The activity and target genes of HSF1 and HSF2 in the context of cell type, cell cycle phase and stimuli.

ACKNOWLEDGEMENTS

The work for this PhD thesis was carried out at the Department of Biosciences, Åbo Akademi University and Centre for Biotechnology Turku (CBT), University of Turku and Åbo Akademi University. I wish to thank Professors John Eriksson, Erik Bonsdorff and Kid Törnquist at the Department of Biosciences and Professor Riitta Lahesmaa at CBT for providing outstanding research facilities with state of art device, energetic athmosphere, dedicated researchers and a skilled staff. Professor Olli Lassila and Nina Widberg are thanked for the encouragement, funding, travel opportunities and for the memorable time that I have had as a member of Turku Doctoral Programme of Biomedical Sciences (TuBS). I wish to extend my deep gratitude to all my colleagues at Åbo Akademi, CBT and TuBS for the fun moments together and for sharing the wonders of life sciences.

I am deeply grateful to Professor John T. Lis for accepting the invitation to act as the opponent of my PhD thesis. The impressive work by Professor Lis inspires striving to excellence and to develop tools where their lack hinders scientific progress. Professor Jorma Palvimo and Docent Maria Vartiainen are warmly acknowledged for thorough pre-examination of this thesis and for their helpful advice in its improving. I am thankful for the members of my thesis committee, Maria Vartiainen and Professor Noora Kotaja, whose expertise in the fields of cell biology and gametogenesis has been invaluable, and whose great humor and advice in life in general have been elemental in directing my future in biomedical sciences.

The grounding force of this thesis has been my supervisor, Professor Lea Sistonen. Lea's profound interest towards science and her belief in her students create an inspiring and encouraging athmosphere where cutting edge science is made possible. I am thankful for the many good years that I'm priviledged to have spent in Lea's lab. I, particularly, thank Lea for her expertise, never-ending endeavor to bring out the best of her students, generous support, confidence to let me lead my own projects and for expecting nothing but excellence.

I wish to warmly thank Professor John Eriksson for the many important roles that he has had during my time as a PhD student. Besides acting as custos, John is a driving force in many scientific communities and get-to-gethers in and outside the lab. John's visions and great attitude towards science has made a deep impact on me.

Science is a way of life that I have been priviledged to share with skillfulled and warm colleagues. Lea's and John's labs are a collection of wonderful personalities and dedicated scientists whose positive attitude towards life has impressed me throughout the years. I wish to collectively thank all the past and present members of the Sistonen-Eriksson laboratories for sharing great discussions, long days in the lab, travels, parties and for showing so much support. With you the work in science has been fun and inspiring, and I feel truly priviledged to be part of this scientific family.

This thesis is result of collaborations: I wish to thank my co-authors Malin Blom, Elisabeth Christians, Annie Conter, Alexandra Elsing, Eva Henriksson, Karoliina Klavér (née Rautoma), Noora Kotaja, Asta Laiho, Pia Roos-Mattjus, Christian Sergelius, Lea Sistonen, Jenni Vasara and Malin Åkerfelf for their contributions and enthusiasm. I would especially like to thank Eva

Henriksson and Malin Åkerfelt for sharing the intriguing projects and discoveries of HSFs in male gametogenesis; as well as Malin Blom, Christian Sergelius and Jenni Vasara whose careful work was elemental in verifying the results of ChIP-sequencing. Jada Balaji, Malin Blom, Johanna Manninen, Christopher Pratt and Hanser Seijas are thanked for being outstanding students and for engaging into the various, some times grazy projects. I wish especially thank Hanser Seijas whose Venezuelan smile outlasted the long Finnish winter and whose positivity was invaluable during the year of optimizing and preparing samples for ChIP-seq. Jari Korhonen and Jouko Sandholm are acknowledged for introducing me to the art of microscopy, and Perttu Terho for technical assistance with flow cytometry. Helena Saarento, Beata Paziewska and Gunilla Henriksson are thanked for their kind assistance in and ouside the lab.

I would like to warmly thank Dr. Michael J. Guertin for introducing me to the analyses of large data sets, sharing his scripts and expertise and for opening the door to the world of programming. Mike's hands-on advice is currently paving the way for years to come. I wish also to thank Professor Elisabeth Christians who took me in her lab and shared her expertise, research interests and showed great hospitality during my stay in Toulouse. Warm thanks also to Aïcha Metchat for introducing me to Toulouse and percussion arts, as well as to Christophe Audouard and Florent Le Masson for making sure I feel home in the lab and find my way to the luch restaurant every day.

My sincerest thanks to all my friends with whom I've had the priviledge to grow up with and share the joys and challenges of life. I have been fortunate to have many great team mates, as well as friends who have not been obsessed by either science or sports. Your contribution to taking me out to the real world has been invaluable.

Throughout my life, my parents Erja and Lauri, and my brother Ville have shown endless support and care, and have helped me realize that the world is full of possibilities. This thesis has its basis on your love which has given me the courage to go after my dreams.

Finally, I wish to acknowledge my family, Sari, Aaron and Erin whose love and support are the essence of my life.

TuBS, Cancer Foundations, Finnish Cultural Foundation, The Research Institute of Åbo Akademi University, Åbo Akademi Rector as well as Joe, Tor and Pentti Borg's Foundation are gratefully acknowledged for financial support.

Turku, May 2014

Anniina Vihervaara

REFERENCES

Abane R, Mezger V. Roles of heat shock factors in gametogenesis and development. (2010). FEBS J **277**: 4150-4172.

Abravaya K, Myers MP, Murphy SP, Morimoto RI. (1992). The human heat shock protein hsp70 interacts with HSF, the transcription factor that regulates heat shock gene expression. Genes Dev **6**: 1153-1164.

Adelman K, Lis JT. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet **13**: 720-731.

Ahlskog JK, Björk JK, Elsing AN, Aspelin C, Kallio M, Roos-Mattjus P, Sistonen L. (2010). Anaphase-promoting complex/cyclosome participates in the acute response to protein-damaging stress. Mol Cell Biol **30**: 5608-5620.

Åkerfelt M, Morimoto RI. and Sistonen L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. Nat Rev Mol Cell Biol **11:** 545-555.

Akhtar A, Gasser SM. (2007). The nuclear envelope and transcriptional control. Nat Rev Genet **8**: 507-517.

Alabert C, Groth A. (2012). Chromatin replication and epigenome maintenance. Nat Rev Mol Cell Biol **13**: 153-167.

Alastalo TP, Lönnström M, Leppä S, Kaarniranta K, Pelto-Huikko M, Sistonen L, Parvinen M. (1998). Stage-specific expression and cellular localization of the heat shock factor 2 isoforms in the rat seminiferous epithelium. Exp Cell Res **240**: 16-27.

Alföldi JE. (2008). Sequence of the Mouse Y Chromosome (Ph.D. Thesis). Cambridge: Massachusetts Institute of Technology. 93 p.

Ali MM, Roe SM, Vaughan CK, Meyer P, Panaretou B, Piper PW, Prodromou C, Pearl LH. (2006). Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. Nature **440**: 1013-1017.

Amin J, Anathan J, Voellmy R. (1988). Key features of heat shock regulatory elements. Mol Cell Biol **8**: 3761-3769.

Anckar J, Hietakangas V, Denessiouk K, Thiele DJ, Johnson MS, Sistonen L. (2006). Inhibition of DNA binding by differential sumoylation of heat shock factors. Mol Cell Biol **26**: 955-964.

Anckar J, Sistonen L. (2007). SUMO: Getting it on. Biochem Soc Trans **35**: 1409-1413.

Anckar J, Sistonen L. (2011). Regulation of HSF1 function in the heat stress response: implications in aging and disease. Annu Rev Biochem **80**: 1089-1115.

Andrés V, González JM. (2009). Role of A-type lamins in signaling, transcription, and chromatin organization. J Cell Biol **187**: 945-957.

Andrulis ED, Guzmán E, Döring P, Werner J, Lis JT. (2000). High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. Genes Dev **14**: 2635-2649.

Aumais JP, Williams SN, Luo W, Nishino M, Caldwell KA, Caldwell GA, Lin SH, Yu-Lee LY. (2003). Role for NudC, a dynein-associated nuclear movement protein, in mitosis and cytokinesis. J Cell Sci **116**: 1991-2003.

Bachtrog D. (2013). Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet **14**: 113-124.

Baler R, Welch WJ, Voellmy R. (1992). Heat shock gene regulation by nascent polypeptides and denatured proteins: hsp70 as a potential autoregulatory factor. J Cell Biol **117**: 1151-1159.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell **125**: 315-326.

Bhowmick BK, Satta Y, Takahata N. (2007). The origin and evolution of human ampliconic gene families and ampliconic structure. Genome Res **17**: 441-450.

Björk JK, Sandqvist A, Elsing AN, Kotaja N, Sistonen L. (2010). miR-18, a member of Oncomir-1, targets heat shock transcription factor 2 in spermatogenesis. Development **137**: 3177-3184.

Björk JK, Sistonen L. (2010). Regulation of the members of the mammalian heat shock factor family. FEBS J **277**: 4126-4139.

Blobel GA, Kadauke S, Wang E, Lau AW, Zuber J, Chou MM, Vakoc CR. (2009). A reconfigured pattern of MLL occupancy within mitotic chromatin promotes rapid transcriptional reactivation following mitotic exit. Mol Cell **36**: 970-983.

Bond U, Schlesinger MJ. (1989). The chicken ubiquitin gene contains a heat shock promoter and expresses an unstable mRNA in heat-shocked cells. Mol Cell Biol **6**: 4602-4610.

Boulon S, Westman BJ, Hutten S, Boisvert FM, Lamond AI. (2010). The nucleolus under stress Mol Cell **40**: 216-227.

Bradley CM, Ronning DR, Ghirlando R, Craigie R, Dyda F. (2005). Structural basis for DNA bridging by barrier-to-autointegration factor. Nat Struct Mol Biol **12**: 935-936.

Braun RE. (2001). Packaging paternal chromosomes with protamine. Nat Genet **28**: 10-12.

Breitling R, Armengaud P, Amtmann A, Herzyk P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett **573**: 83-92.

Brès V, Yoh SM, Jones KA. (2008). The multi-tasking P-TEFb complex. Curr Opin Cell Biol **20**: 334-340.

Brown SA, Imbalzano AN, Kingston RE. (1996). Activator-dependent regulation of transcriptional pausing on nucleosomal templates. Genes Dev **10**: 1479-1490.

Brown SA, Kingston RE. (1997). Disruption of downstream chromatin directed by a transcriptional activator. Genes Dev **11**: 3116-3121.

Brownell JE, Allis CD. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Curr Opin Genet Dev **6**: 176-184.

Burgoyne PS, Mahadevaiah SK, Turner JM. (2009). The consequences of asynapsis for mammalian meiosis. Nat Rev Genet **10**: 207-216.

Cantone I, Fisher AG. (2013). Epigenetic programming and reprogramming during development. Nat Struct Mol Biol **20**: 282-289.

Castillo-Davis CI, Hartl DL. (2003). GeneMerge--post-genomic analysis, data mining, and hypothesis testing. Bioinformatics **19**: 891-892.

Chang Y, Östling P, Åkerfelt M, Trouillet D, Rallu M, Gitton Y, El Fatimy R, Fardeau V, Le Crom S, Morange M, Sistonen L, Mezger V. (2006). Role of heat-shock factor 2 in cerebral cortex formation and as a regulator of p35 expression. Genes Dev **20**: 836-847.

Chen D, Dundr M, Wang C, Leung A, Lamond A, Misteli T, Huang S. (2005). Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. J Cell Biol **168**: 41-54.

Christians E, Davis AA, Thomas SD, Benjamin IJ. (2000). Maternal effect of Hsf1 on reproductive success. Nature **407**: 693-694.

Chu B, Soncin F, Price BD, Stevenson MA, Calderwood SK. (1996). Sequential phosphorylation by mitogen-activated protein kinase and glycogen synthase kinase 3 represses transcriptional activation by heat shock factor-1. J Biol Chem **271**: 30847-30857.

Cleveland WS, Devlin SJ. (1988). Locally-weighted regression: an approach to regression analysis by local fitting. J Am Stat Assoc **83**: 596-610.

Cocquet J, Ellis PJ, Mahadevaiah SK, Affara NA, Vaiman D, Burgoyne PS. (2012). A genetic basis for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. PLoS Genet **8**: e1002900.

Consortium EP. (2011). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol **9**: e1001046.

Cook PR. (1999). The organization of replication and transcription. Science **284**: 1790-1795.

Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. (2012). Defining the status of RNA polymerase at promoters. Cell Rep **2**: 1025-1035.

Craig EA, Huang P, Aron R, Andrew A. (2006). The diverse roles of J-proteins, the obligate Hsp70 co-chaperone. Rev Physiol Biochem Pharmacol **156**: 1-21.

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn CD, Lehmann E, Leike K, Sydow JF, Vannini A. (2008). Structure of eukaryotic RNA polymerases. Annu Rev Biophys **37**: 337-352.

Darnell JE Jr. (2013). Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture. RNA **19**: 443-460.

Dai C, Whitesell L, Rogers AB, Lindquist S. (2007). Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. Cell **130**: 1005-1018.

Dekker J, Marti-Renom MA, Mirny LA. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet **14**: 390-403.

Delcuve GP, He S, Davie JR. (2008). Mitotic partitioning of transcription factors. J Cell Biochem **105**: 1-8.

Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol **4**: P3.

Dohle GR, Smit M, Weber RF. (2003). Androgens and male fertility. World J Urol **21**: 341-345.

Edwards RG, Beard HK. (1997). Oocyte polarity and cell determination in early mammalian embryos. Molecular Human Reproduction **3**: 863-905.

Ellis PJ, Clemente EJ, Ball P, Touré A, Ferguson L, Turner JM, Loveland KL, Affara NA, Burgoyne PS. (2005). Deletions on mouse Yq lead to upregulation of multiple X- and Y-linked transcripts in spermatids. Hum Mol Genet **14**: 2705-2715.

Erickson RP. (1973). Haploid gene expresion versus meiotic drive: the relevance of intercellular bridges during spermatogenesis. Nat New Biol **243**: 210-212.

Farnham PJ. (2009). Insights from genomic profiling of transcription factors. Nat Rev Genet **10**: 605-616.

Feklistov A. (2013). RNA polymerase: in search of promoters. Ann N Y Acad Sci **1293**: 25-32.

Felsenfeld G, Groudine M. (2003). Controlling the double helix. Nature **421**: 448-453.

Feng J, Liu T, Qin B, Zhang Y, Liu XS. (2012). Identifying ChIP-seq enrichment using MACS. Nat Protoc **7**: 1728-1740.

Fijak M, Meinhardt A. (2006). The testis in immune privilege. Immunol Rev **213**: 66-81.

Finley D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu Rev Biochem **78**: 477-513.

Finley D, Ozkaynak E, Varshavsky A. (1987). The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. Cell **48**: 1035-1046.

Fiorenza MT, Farkas T, Dissing M, Kolding D, Zimarino V. (1995). Complex expression of murine heat shock factors. Nucleic Acids Res **23**: 467-474.

Fornace AJ Jr, Alamo I Jr, Hollander MC, Lamoreaux E. (1986). Ubiquitin mRNA is a major stress-induced transcript in mammalian cells. Nucleic Acids Res **17**: 1215-1230.

Fuda NJ, Ardehali MB, Lis JT. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature **461**: 186-192.

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

Fujimoto M, Hayashida N, Katoh T, Oshima K, Shinkawa T, Prakasam R, Tan K, Inouye S, Takii R, Nakai A. (2010). A novel mouse HSF3 has the potential to activate nonclassical heat-shock genes during heat shock. Mol Biol Cell **21**: 106-116.

Fujimoto M, Izu H, Seki K, Fukuda K, Nishida T, Yamada S, Kato K, Yonemura S, Inouye S, Nakai A. (2004). HSF4 is required for normal cell growth and differentiation during mouse lens development. EMBO J **23**: 4297-4306.

Fujimoto M, Nakai A. (2010). The heat shock factor family and adaptation to proteotoxic stress. FEBS J **277**: 4112-4125.

Fujimoto M, Takaki E, Takii R, Tan K, Prakasam R, Hayashida N, Iemura S, Natsume T, Nakai A. (2012). RPA assists HSF1 access to nucleosomal DNA by recruiting histone chaperone FACT. Mol Cell **48**: 182-194.

Fussner E, Djuric U, Strauss M, Hotta A, Perez-Iratxeta C, Lanner F, Dilworth FJ, Ellis J, Bazett-Jones DP. (2011). Constitutive heterochromatin reorganization during somatic cell reprogramming. EMBO J **30**: 1778-1789.

Gallo GJ, Prentice H, Kingston RE. (1993). Heat shock factor is required for growth at normal temperatures in the fission yeast Schizosaccharomyces pombe. Mol Cell Biol **13**: 749-761.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. (2000). Genomic expression programs in the response of yeast cells to environmental changes Mol Biol Cell **11**: 4241-4257.

Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S. (2010). From meiosis to postmeiotic events: the secrets of histone disappearance. FEBS J **277**: 599-604.

Gegonne A, Weissman JD, Lu H, Zhou M, Dasgupta A, Ribble R, Brady JN, Singer DS. (2008). TFIID component TAF7 functionally interacts with both TFIIH and P-TEFb. Proc Natl Acad Sci U S A **105:** 5367-5372.

Gegonne A, Weissman JD, Zhou M, Brady JN, Singer DS. (2006). TAF7: a possible transcription initiation check-point regulator. Proc Natl Acad Sci U S A **103**: 602-607.

Gibcus JH, Dekker J. (2013). The hierarchy of the 3D genome. Mol Cell **49**: 773-782.

Gidalevitz T, Prahlad V, Morimoto RI. (2011). The Stress of Protein Misfolding: From Single Cells to Multicellular Organisms. Cold Spring Harb Perspect Biol **3**: a009704.

Gilmour DS, Fan R. (2008). Derailing the locomotive: transcription termination. J Biol Chem **283**: 661-664.

Gilmour DS, Lis JT. (1984). Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. Proc Natl Acad Sci U S A **81**: 4275-4279.

Gilmour DS, Lis JT. (1985). In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. Mol Cell Biol **5**: 2009-2018.

Gilmour DS, Pflugfelder G, Wang JC, Lis JT. (1986). Topoisomerase I interacts with transcribed regions in Drosophila cells. Cell **44**: 401-407.

Gonsalves SE, Moses AM, Razak Z, Robert F, Westwood JT. (2011). Whole-genome analysis reveals that active heat shock factor binding sites are mostly associated with non-heat shock genes in Drosophila melanogaster. PLoS One **6**: e15934.

Gottardo R. (2009). Modeling and analysis of ChIP-chip experiments. Methods Mol Biol **567**: 133-143.

Grosse R, Vartiainen MK. (2013). To be or not to be assembled: progressing into nuclear actin filaments. Nat Rev Mol Cell Biol **14**: 693-697.

Guertin MJ, Lis JT. (2010). Chromatin landscape dictates HSF binding to target DNA elements. PLoS Genet **6**: e1001114.

Guertin MJ, Petesch SJ, Zobeck KL, Min IM, Lis JT. (2010). Drosophila heat shock system as a general model to investigate transcriptional regulation. Cold Spring Harb Symp Quant Biol **75**: 1-9.

Guettouche T, Boellmann F, Lane WS, Voellmy R. (2005). Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. BMC Biochem **6**: 4.

Guibert S, Forné T, Weber M. (2009). Dynamic regulation of DNA methylation during mammalian development. Epigenomics. **1**: 81-98.

Guo Y, Guettouche T, Fenna M, Boellmann F, Pratt WB, Toft DO, Smith DF, Voellmy R. (2001). Evidence for a mechanism of repression of heat shock factor 1 transcriptional activity by a multichaperone complex. J Biol Chem **276**: 45791-45799.

Hageman J, Rujano MA, van Waarde MA, Kakkar V, Dirks RP, Govorukhina N, Oosterveld-Hut HM, Lubsen NH, Kampinga HH. (2010). A DNAJB chaperone subfamily with HDAC-dependent activities suppresses toxic protein aggregation. Mol Cell **37**: 355-369.

Hahn JS, Hu Z, Thiele DJ, Iyer VR. (2004). Genome-wide analysis of the biology of stress responses through heat shock transcription factor. Mol Cell Biol **24**: 5249-5256.

Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R, Surani MA. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature **452**: 877-881.

Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science **329**: 78-82.

Hamperl S, Wittner M, Babl V, Perez-Fernandez J, Tschochner H, Griesenbeck J. (2013). Chromatin states at ribosomal DNA loci. Biochim Biophys Acta **1829**: 405-417.

Hanahan D, Weinberg RA. (2011). Hallmarks of cancer: the next generation. Cell **144**: 646-674.

Haraguchi T, Kojidani T, Koujin T, Shimi T, Osakada H, Mori C, Yamamoto A, Hiraoka Y. (2008). Live cell imaging and electron microscopy reveal dynamic processes of BAF-directed nuclear envelope assembly. J Cell Sci **121**: 2540-2554.

Hartl UF, Bracher A, Hayer-Hartl M. (2011). Molecular chaperones in protein folding and proteostasis. Nature **475**: 324-332.

Hess RA, de Franca RL. (2008). Spermatogenesis and cycle of the seminiferous epithelium. Adv Exp Med Biol **636**: 1-15.

Hesselberth JR, Chen X, Zhang Z, Sabo PJ, Sandstrom R, Reynolds AP, Thurman RE, Neph S, Kuehn MS, Noble WS, Fields S, Stamatoyannopoulos JA. (2009). Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. Nat Methods **6**: 283-289.

Hietakangas V, Ahlskog JK, Jakobsson AM, Hellesuo M, Sahlberg NM, Holmberg CI, Mikhailov A, Palvimo JJ, Pirkkala L, Sistonen L. (2003). Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. Mol Cell Biol **23**: 2953-2968.

Hietakangas V, Anckar J, Blomster HA, Fujimoto M, Palvimo JJ, Nakai A, Sistonen L. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification. Proc Natl Acad Sci U S A **103**: 45-50.

Hjerpe R, Aillet F, Lopitz-Otsoa F, Lang V, England P, Rodriguez MS. (2009). Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. EMBO Rep **10**: 1250-1258.

Holaska JM, Kowalski AK, Wilson KL. (2004). Emerin caps the pointed end of actin filaments: evidence for an actin cortical network at the nuclear inner membrane. PLoS Biol **2**: E231.

Holmberg CI, Hietakangas V, Mikhailov A, Rantanen JO, Kallio M, Meinander A, Hellman J, Morrice N, MacKintosh C, Morimoto RI, Eriksson JE, Sistonen L. (2001). Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. EMBO J **20**: 3800-3810.

Hut HM, Kampinga HH, Sibon OC. (2005). Hsp70 protects mitotic cells against heat-induced centrosome damage and division abnormalities. Mol Biol Cell **16**: 3776-3785.

Iborra FJ, Pombo A, Jackson DA, Cook PR. (1996). Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. J Cell Sci **109**: 1427-1436.

Inouye S, Izu H, Takaki E, Suzuki H, Shirai M, Yokota Y, Ichikawa H, Fujimoto M, Nakai A. (2004). Impaired IgG production in mice deficient for heat shock transcription factor 1. J Biol Chem **279**: 38701-38709.

Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M, Brown PO. (2001). Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. Nature **409**: 533-538.

Izu H, Inouye S, Fujimoto M, Shiraishi K, Naito K, Nakai A. (2004). Heat shock transcription factor 1 is involved in quality-control mechanisms in male germ cells. Biol Reprod **70**: 18-24.

Jakobsen BK, Pelham HR. (1988). Constitutive binding of yeast heat shock factor to DNA in vivo. Mol Cell Biol **8**: 5040-5042.

Jedlicka P, Mortin MA, Wu C. (1997). Multiple functions of Drosophila heat shock transcription factor in vivo. EMBO J **16**: 2452-2462.

Jeon H, Blacklow SC. (2005). Structure and physiologic function of the low-density lipoprotein receptor. Annu Rev Biochem **74**: 535-562.

Jin X, Moskophidis D, Mivechi NF. (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.

Kadauke S, Blobel GA. (2012). "Remembering" tissue-specific transcription patterns through mitosis. Cell Cycle **11**: 3911-3912.

Kadonaga JT. (2012). Perspectives on the RNA polymerase II core promoter. Wiley Interdiscip Rev Dev Biol **1**: 40-51.

Kallio M, Chang Y, Manuel M, Alastalo T-P, Rallu M, Gitton Y, Pirkkala L, Loones MT, Paslaru L, Larney S, Hiard S, Morange M, Sistonen L, Mezger V. (2002). Brain abnormalities, defective meiotic chromosome synapsis and female subfertility in HSF2 null mice. EMBO J **21**: 2591-2601.

Kampinga HH, Hageman J, Vos MJ, Kubota H, Tanguay RM, Bruford EA, Cheetham ME, Chen B, Hightower LE. (2009). Guidelines for the nomenclature of the human heat shock proteins. Cell Stress Chap **14**: 105-111.

Kaplan CD, Morris JR, Wu C, Winston F. (2000). Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster. Genes Dev **14**: 2623-2634.

Keyse SM, Emslie EA. (1992). Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase. Nature **359**: 644-647.

Kimura Y, Tanaka K. (2010). Regulatory mechanisms involved in the control of ubiquitin Homeostasis. J Biochem **147**: 793-798.

Khalil AM, Boyar FZ, Driscoll DJ. (2004). Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis. Proc Natl Acad Sci U S A **101**: 16583-16587.

Kimmins S, Kotaja N, Davidson I, Sassone-Corsi P. (2004). Testis-specific transcription mechanisms promoting male germ-cell differentiation. Reproduction **128**: 5-12.

Kimmins S, Sassone-Corsi P. (2005). Chromatin remodelling and epigenetic features of germ cells. Nature **434**: 583-589.

Kinoshita K, Shinka T, Sato Y, Kurahashi H, Kowa H, Chen G, Umeno M, Toida K, Kiyokage E, Nakano T, Ito S, Nakahori Y. (2006). Expression analysis of a mouse orthologue of HSFY, a candidate for the azoospermic factor on the human Y chromosome. J Med Invest **53**: 117-122.

Kleene KC. (2003). Patterns, mechanisms, and functions of translation regulation in mammalian spermatogenic cells. Cytogenet Genome Res **103**: 217-224.

Kline MP, Morimoto RI. (1997). Repression of the heat shock factor 1 transcriptional activation domain is modulated by constitutive phosphorylation. Mol Cell Biol **17**: 2107-2115.

de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. (2011). Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet **7**: e1002384.

Kotaja N, Kimmins S, Brancorsini S, Hentsch D, Vonesch JL, Davidson I, Parvinen M, Sassone-Corsi P. (2004). Preparation, isolation and characterization of stage-specific spermatogenic cells for cellular and molecular analysis. Nat Methods **1**: 249-254.

Kotaja N, Sassone-Corsi P. (2007). The chromatoid body: a germ-cell-specific RNA-processing centre. Nat Rev Mol Cell Biol **8**: 85-90.

Kornblihtt AR. (2007). Coupling transcription and alternative splicing. Adv Exp Med Biol **623**: 175-189.

Kourtis N, Tavernakis N. (2011). Cellular stress response pathways and ageing: intricate molecular relationships. EMBO J **30**: 2520-2531.

Kroeger PE, Morimoto RI. (1994). Selection of new HSF1 and HSF2 DNA-binding sites reveals difference in trimer cooperativity. Mol Cell Biol **14**: 7592-7603.

Kroeger PE, Sarge KD, Morimoto RI. (1993). Mouse heat shock transcription factors 1 and 2 prefer a trimeric binding site but interact differently with the HSP70 heat shock element. Mol Cell Biol **13**: 3370-3383.

Kühl NM, Rensing L. (2000). Heat shock effects on cell cycle progression. Cell Mol Life Sci **57**: 450-463.

Kumano G. (2012). Polarizing animal cells via mRNA localization in oogenesis and early development. Develop Growth Differ **54**: 1-18.

Langmead B, Trapnell C, Pop M, Salzberg SL. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol **10**: R25.

Le Masson F, Razak Z, Kaigo M, Audouard C, Charry C, Cooke H, Westwood JT, Christians ES. (2011). Identification of heat shock factor 1 molecular and cellular targets during embryonic and adult female meiosis. Mol Cell Biol **31**: 3410-3423.

de Lanerolle P, Serebryannyy L. (2011). Nuclear actin and myosins: life without filaments. Nat Cell Biol **13**: 1282-1288.

Lee HS, Simon JA, Lis JT. (1988). Structure and expression of ubiquitin genes of Drosophila melanogaster. Mol Cell Biol **8**: 4727-4735.

Lewis M, Helmsing PJ, Ashburner M. (1975). Parallel changes in puffing activity and patterns of protein synthesis in salivary glands of Drosophila. Proc Natl Acad Sci U S A **72**: 3604-3608.

Li J, Soroka J, Buchner J. (2012). The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones. Biochim Biophys Acta **1823**: 624-635.

Lindquist S. (1981). Regulation of protein synthesis during heat shock. Nature **293**: 311-314.

Lis JT, Mason P, Peng J, Price DH, Werner J. (2000). P-TEFb kinase recruitment and function at heat shock loci. Genes Dev **14**: 792-803.

Liu H, Cheng EH, Hsieh JJ. (2007). Bimodal degradation of MLL by SCFSkp2 and APCCdc20 assures cell cycle execution: a critical regulatory circuit lost in leukemogenic MLL fusions. Genes Dev **21**: 2385-2398.

Littlefield O, Nelson HC. (1999). A new use for the 'wing' of the 'winged' helix-turn-helix motif in the HSF-DNA cocrystal. Nat Struct Biol **6**: 464-470.

Loison F, Debure L, Nizard P, le Goff P, Michel D, le Dréan Y. (2006). Up-regulation of the clusterin gene after proteotoxic stress: implication of HSF1-HSF2 heterocomplexes. Biochem J **395**: 223-231.

Martínez-Balbás MA, Dey A, Rabindran SK, Ozato K, Wu C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. Cell **83**: 29-38.

Machanic P, Bailey TL. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics **27**: 1696-1697.

Maston GA, Evans SK, Green MR. (2006). Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet **7**: 29-59.

McKee BD, Handel MA. (1993). Sex chromosomes, recombination, and chromatin conformation. Chromosoma **102**: 71-80.

McMillan DR, Xiao X, Shao L, Graves K, Benjamin IJ. (1998). Targeted disruption of heat shock transcription factor 1 abolishes thermotolerance and protection against heat-inducible apoptosis. J Biol Chem **273**: 7523-7528.

Meikar O, Da Ros M, Kotaja N. (2012). Epigenetic regulation of male germ cell differentiation. Subcell Biochem **61**: 119-138.

Mellor H, Parker PJ. (1998). The extended protein kinase C superfamily. Biochem J **332**: 281-292.

Mendillo ML, Santagata S, Koeva M, Bell GW, Hu R, Tamimi RM, Fraenkel E, Ince TA, Whitesell L, Lindquist S. (2012). HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. Cell **150**: 549-562.

Metchat A, Åkerfelt M, Bierkamp C, Delsinne V, Sistonen L, Alexandre H and Christians ES. (2009). Mammalian heat shock factor 1 is essential for oocyte meiosis and directly regulates Hsp90alpha expression. J Biol Chem **284**: 9521-9528.

Métivier R, Reid G, Gannon F. (2006). Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. EMBO Rep **7**: 161-167.

Miller D, Brinkworth M, Iles D. (2010). Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics. Reproduction **139**: 287-301.

Min JN, Han MY, Lee SS, Kim KJ, Park YM. (2000). Regulation of rat heat shock factor 2 expression during the early organogenic phase of embryogenesis. Biochim Biophys Acta **1494**: 256-262.

Mitchell JA, Fraser P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. Genes Dev **22**: 20-25.

Mohn F, Schübeler D. (2009). Genetics and epigenetics: stability and plasticity during cellular differentiation. Trends Genet **25**: 129-136.

Morimoto RI. (1988). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. Genes Dev **12**: 3788-3796.

Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JM. (2008). The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. Nat Genet **40**: 794-799.

Naganuma T, Hirose T. (2013). Paraspeckle formation during the biogenesis of long non-coding RNAs. RNA Biol **10**: 456-461.

Nakai A, Tanabe M, Kawazoe Y, Inazawa J, Morimoto RI, Nagata K. (1997). HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. Mol Cell Biol **17**: 469-481.

Nakai A, Suzuki M, Tanabe M. (2000). Arrest of spermatogenesis in mice expressing an active heat shock transcription factor 1. EMBO J **19**: 1545-1554.

Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD, Lee JT. (2006). Postmeiotic sex chromatin in the male germline of mice. Curr Biol **16**: 660-667.

Nishino M, Kurasawa Y, Evans R, Lin SH, Brinkley BR, Yu-Lee LY. (2006). NudC is required for Plk1 targeting to the kinetochore and chromosome congression. Curr Biol **16**: 1414-1421.

Nishino Y, Eltsov M, Joti Y, Ito K, Takata H, Takahashi Y, Hihara S, Frangakis AS, Imamoto N, Ishikawa T, Maeshima K. (2012). Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. EMBO J **31**:1644-1653.

Oakberg EF. (1956). A description of spermiogenesis in the mouse and its use in analysis of the cycle of the seminiferous epithelium and germ cell renewal. Am J Anat **99**: 391-413.

Ohno S, Kaplan WD, Kinosita R. (1959). On the end-to-end association of the X and Y chromosomes of Mus musculus. Exp Cell Res **18**: 282-290.

Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet **36**: 1065-1071.

Östling P, Björk JK, Roos-Mattjus P, Mezger V, 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.

Pabo CO, Sauer RT. (1992). Transcription factors: structural families and principles of DNA recognition. Annu Rev Biochem **61**: 1053-1095.

Park PJ. (2009). ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet **10**: 669-680.

Park JM, Werner J, Kim JM, Lis JT, Kim YJ. (2001). Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. Mol Cell **8**: 9-19.

Paronetto MP, Sette C. (2010). Role of RNA-binding proteins in mammalian spermatogenesis. Int J Androl **33**: 2-12.

Parvinen M. (2005). The chromatoid body in spermatogenesis. Int J Androl **28**: 189-201.

Parvinen M, Hecht NB. (1981). Identification of living spermatogenic cells of the mouse by transillumination-phase contrast microscopic technique for 'in situ' analyses of DNA polymerase activities. Histochemistry **71**: 567-579.

Pearson JC, Lemons D, McGinnis W. (2005). Modulating Hox gene functions during animal body patterning. Nat Rev Genet **6**: 893-904.

Pelham HR. (1982). A regulatory upstream promoter element in the Drosophila hsp 70 heat-shock gene. Cell **30**: 517-528.

Pepke S, Wold B, Mortazavi A. (2009). Computation for ChIP-seq and RNA-seq studies. Nat Methods **6**: S22-S32.

Pennisi E. (2012). ENCODE project writes eulogy for junk DNA. Science **337:** 1159-1161.

Perry J, Palmer S, Gabriel A, Ashworth A. (2001). A short pseudoautosomal region in laboratory mice. Genome Res **11**: 1826-1832.

Peterson CL, Laniel MA. (2004). Histones and histone modifications. Curr Biol **14**: R546-551.

Petesch SJ, Lis JT. (2008). Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. Cell **134**: 74-84.

Pirkkala L, Nykänen P, Sistonen L. (2001). Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. FASEB J **15**: 1118-1131.

Piotrowska K, Zernicka-Goetz M. (2002). Early patterning of the mouse embryo: contributions of sperm and egg. Development **129**: 5803-5813.

Pogany GC, Corzett M, Weston S, Balhorn R. (1981). DNA and protein content of mouse sperm. Implications regarding sperm chromatin structure. Exp Cell Res **136**: 127-136.

Pope BD, Aparicio OM, Gilbert DM. (2013). SnapShot: Replication timing. Cell **152**: 1390-1390.

Probst, AV, Dunleavy E, Almouzni G. (2009). Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol **10**: 192-206.

Qu H, Fang X. (2013). A brief review on the Human Encyclopedia of DNA Elements (ENCODE) project. Genomics Proteomics Bioinformatics **11**: 135-141.

Rabindran SK, Haroun RI, Clos J, Wisniewski J, Wu C. (1993). Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science **259**: 230-234.

Rallu M, Loones M, Lallemand Y, Morimoto R, Morange M, Mezger V. (1997). Function and regulation of heat shock factor 2 during mouse embryogenesis. Proc Natl Acad Sci U S A **94**: 2392-2397.

Rampelt H, Kirstein-Miles J, Nillegoda NB, Chi K, Scholz SR, Morimoto RI, Bukau B. (2012). Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. EMBO J **31**: 4221-4235.

Rasmussen EB, Lis JT. (1993). *In vivo* transcriptional pausing and cap formation on three *Drosophila* heat shock genes. Proc Natl Acad Sci U S A **90**: 7923-7927.

Reddy PM, Stamatoyannopoulos G, Papayannopoulou T, Shen CK. (1994). Genomic footprinting and sequencing of human beta-globin locus: Tissue specificity and cell line artifact. Journal of Biological Chemistry **269**: 8287-8295.

Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA. (2000). Genome-wide location and function of DNA binding proteins. Science **290**: 2306-2309.

Retzlaff M, Hagn F, Mitschke L, Hessling M, Gugel F, Kessler H, Richter K, Buchner J. (2010). Asymmetric activation of the hsp90 dimer by its cochaperone aha1. Mol Cell **37**: 344-354.

Richard P, Manley JL. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev **23**: 1247-1269.

Richter K, Haslbeck M, Buchner J. (2010). The heat shock response: life on the verge of death. Mol Cell **40**: 253-266.

Riel JM, Yamauchi Y, Sugawara A, Li HY, Ruthig V, Stoytcheva Z, Ellis PJ, Cocquet J, Ward MA. (2013). Deficiency of the multi-copy mouse Y gene Sly causes sperm DNA damage and abnormal chromatin packaging. J Cell Sci **126**: 803-813.

Riethoven JJ. (2010). Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. Methods Mol Biol **674**: 33-42.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. (2011). Integrative Genomics Viewer. Nat Biotechnol **29**: 24-26.

Ritossa FM. (1962). A new puffing pattern induced by a temperature shock and DNP in Drosophila. Experientia **18**: 571-573.

Roeder RG. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett **579**: 909-915.

Roegiers F, Jan YN. (2004). Asymmetric cell division. Curr Opin Cell Biol **16**: 195-205.

Rougvie AE, Lis JT. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell **54**: 795-804.

Rousseaux S, Caron C, Govin J, Lestrat C, Faure AK, Khochbin S. (2005). Establishment of male-specific epigenetic information. Gene **345**: 139-153.

Russell LD, Ettlin RA, SinhaHikim AP, Clegg ED. (1990). Histological and histopathological evaluation of the testis. Cache River Press 1st Ed.

Sakkas D, Moffatt O, Manicardi GC, Mariethoz E, Tarozzi N, Bizzaro D. (2002). Nature of DNA damage in ejaculated human spermatozoa and the possible involvement of apoptosis. Biol Reprod **66**: 1061-1067.

Salmand PA, Jungas T, Fernandez M, Conter A, Christians ES. (2008). Mouse heat-shock factor 1 (HSF1) is involved in testicular response to genotoxic stress induced by doxorubicin. Biol Reprod **79**: 1092-1101.

Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell **146**: 353-358.

Sandqvist A, Björk JK, Åkerfelt M, Chitikova Z, Grichine A, Vourc'h C, Jolly C, Salminen TA, Nymalm Y, 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.

Santagata S, Hu R, Lin NU, Mendillo ML, Collins LC, Hankinson SE, Schnitt SJ, Whitesell L, Tamimi RM, Lindquist S, Ince TA. (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.

Santagata S, Mendillo ML, Tang YC, Subramanian A, Perley CC, Roche SP, Wong B, Narayan R, Kwon H, Koeva M, Amon A, Golub TR, Porco JA Jr, Whitesell L, Lindquist S. (2013). Tight coordination of protein translation and HSF1 activation supports the anabolic malignant state. Science **341**: 1238303.

Sarge KD, Bray AE, Goodson ML. (1995). Altered stress response in testis. Nature **374**: 126.

Sarge KD, Murphy SP, Morimoto RI. (1993). Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization, acquisition of DNA-binding activity, and nuclear localization and can occur in the absence of stress. Mol Cell Biol **13**: 1392-1407.

Schlick T, Hayes J, Grigoryev S. (2012). Toward convergence of experimental studies and theoretical modeling of the chromatin fiber. J Biol Chem **287**: 5183-5191.

Schmeing TM, Ramakrishnan V. (2009). What recent ribosome structures have revealed about the mechanism of translation. Nature **461**: 1234-1242.

Schneider CA, Rasband WS, Eliceiri KW. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature Methods **9**: 671-675.

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. (2007). Genome regulation by polycomb and trithorax proteins. Cell **128**: 735-745.

Selth LA, Sigurdsson S, Svejstrup JQ. (2010). Transcript elongation by RNA polymerase II. Annu Rev Biochem **79**: 271-293.

Shi Y, Mosser DD, Morimoto RI. (1998). Molecular chaperones as HSF1-specific transcriptional repressors. Genes Dev **12**: 654-666.

Shilatifard A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem **75**: 243-269.

Shinkawa T, Tan K, Fujimoto M, Hayashida N, Yamamoto K, Takaki E, Takii R, Prakasam R, Inouye S, Mezger V, 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.

Simon DN, Zastrow MS, Wilson KL. (2010). Direct actin binding to A- and B-type lamin tails and actin filament bundling by the lamin A tail. Nucleus **1**: 264-272.

Sistonen L, Sarge KD, Abravaya K, Phillips B, Morimoto RI. (1992). Activation of heat shock factor 2 (HSF2) during hemin-induced differentiation of human erythroleukemia cells. Mol Cell Biol **12**: 4104-4111.

Sistonen L, Sarge KD, Morimoto RI. (1994). Human heat shock factors 1 and 2 are differentially activated and can synergistically induce hsp70 gene transcription. Mol Cell Biol **14**: 2087-2099.

Solari AJ. (1974). The behavior of the XY pair in mammals. Rev Cytol **38**: 273-317.

Solimini NL, Luo J, Elledge SJ. (2007). Non-oncogene addiction and the stress phenotype of cancer cells. Cell **130**: 986-988.

Solomon MJ, Larsen PL, Varshavsky A. (1988). Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell **53**: 937-947.

Sorger PK, Nelson HC. (1989). Trimerization of a yeast transcriptional activator via a coiled-coil motif. Cell **59**: 807-813.

Sorger PK, Pelham HRB. (1988). Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. Cell **54**: 855-864.

Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, Barthès P, Kokkinaki M, Nef S, Gnirke A, Dym M, de Massy B, Mikkelsen TS, Kaessmann H. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep **3**: 2179-2190.

Spradling A, Penman S, Pardue ML. (1975). Analysis of drosophila mRNA by in situ hybridization: sequences transcribed in normal and heat shocked cultured cells. Cell **4**: 395-404.

Sullivan EK, Weirich CS, Guyon JR, Sif S, Kingston RE. (2001). Transcriptional activation domain of heat shock factor 1 recruit human SWI/SNF. Mol Cell Biol **21**: 5826-5837.

Sutherland H, Bickmore WA. (2009). Transcription factories: gene expression in unions? Nat Rev Genet **10**: 457-466.

Szalay MS, Kovács IA, Korcsmáros T, Böde C, Csermely P. (2007). Stress-induced rearrangements of cellular networks: Consequences for protection and drug design. FEBS Lett **581**: 3675-3680.

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

Taipale M, Krykbaeva I, Koeva M, Kayatekin C, Westover KD, Karras GI, Lindquist S. (2012). Quantitative analysis of hsp90-client interactions reveals principles of substrate recognition. Cell **150**: 987-1001.

Taulli R, Loretelli C, Pandolfi PP. (2013). From pseudo-ceRNAs to circ-ceRNAs: a tale of cross-talk and competition. Nat Struct Mol Biol **20**: 541-543.

Tessari A, Salata E, Ferlin A, Bartoloni L, Slongo ML, Foresta C. (2004). Characterization of HSFY, a novel AZFb gene on the Y chromosome with a possible role in human spermatogenesis. Mol Hum Reprod **10**: 253-258.

Tissières A, Mitchell HK, Tracy UM. (1974). Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J Mol Biol **84**: 389-398.

Toivola DM, Strnad P, Habtezion A, Omary MB. (2010). Intermediate filaments take the heat as stress proteins. Trends Cell Biol **20**: 79-91.

Tollervey JR, Lunyak VV. (2012). Epigenetics: judge, jury and executioner of stem cell fate. Epigenetics **7**: 823-840.

Topol J, Ruden DM, Parker CS. (1985). Sequences required for in vitro transcriptional activation of a Drosophila hsp70 gene. Cell **42**: 527-537.

Touré A, Clemente EJ, Ellis P, Mahadevaiah SK, Ojarikre OA, Ball PA, Reynard L, Loveland KL, Burgoyne PS, Affara NA. (2005). Identification of novel Y chromosome encoded transcripts by testis transcriptome analysis of mice with deletions of the Y chromosome long arm. Genome Biol **6**: R102.

Touré A, Szot M, Mahadevaiah SK, Rattigan A, Ojarikre OA, Burgoyne PS. (2004). A new deletion of the mouse Y chromosome long arm associated with the loss of Ssty expression, abnormal sperm development and sterility. Genetics **166**: 901-912.

Treangen TJ, Salzberg SL. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet **13**: 36-46.

Tres LL. (1977). Extensive pairing of the XY bivalent in mouse spermatocytes as visualized by whole-mount electron microscopy. J Cell Sci **25**: 1-15.

Trinklein ND, Murray JI, Hartman SJ, Botstein D, Myers RM. (2004). The role of heat shock transcription factor 1 in the genome-wide regulation of the mammalian heat shock response. Mol Cell Biol **15**: 1254-1261.

Turner JM. (2007). Meiotic sex chromosome inactivation. Development **134**: 1823-1831.

Tyagi S, Chabes AL, Wysocka J, Herr W. (2007). E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. Mol Cell **27**: 107-119.

Vannini A, Cramer P. (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol Cell **45**: 439-446.

Ventelä S, Toppari J, Parvinen M. (2003). Intercellular organelle traffic through cytoplasmic bridges in early spermatids of the rat: mechanisms of haploid gene product sharing. Mol Biol Cell **14**: 2768-2780.

Venter JC, *et al*. (2001). The sequence of the human genome. Science **291**: 1304-1351.

Vigh L, Nakamoto H, Landry J, Gomez-Munoz A, Harwood JL, Horvath I. (2007). Membrane regulation of the stress response from prokaryotic models to mammalian cells. Ann N Y Acad Sci **1113**: 40-51.

Vihervaara A. (2007). Värmechockfaktor 2 som reglerare av cytokiners expression under musens spermatogenes. Master's Thesis. Turku Centre for Biotechnology, Åbo Akademi University.

Vihervaara A, Sistonen L. (2014). HSF1 at a glance. J Cell Sci: in press.

Vydra N, Malusecka E, Jarzab M, Lisowska K, Glowala-Kosinska M, Benedyk K, Widlak P, Krawczyk Z, Widlak W. (2006). Spermatocyte-specific expression of constitutively active heat shock factor 1 induces HSP70i-resistant apoptosis in male germ cells. Cell Death Differ **13**: 212-222.

Wang G, Ying Z, Jin X, Tu N, Zhang Y, Phillips M, Moskophidis D, Mivechi NF. (2004). Essential requirement for both hsf1 and hsf2 transcriptional activity in spermatogenesis and male fertility. Genesis **38**: 66-80.

Wang G, Zhang J, Moskophidis D, Mivechi NF. (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.

Westerheide SD, Anckar J, Stevens SM Jr, Sistonen L, Morimoto RI. (2009). Stress-inducible regulation of heat shock factor 1 by the deacetylase SIRT1. Science **323**: 1063-1066.

Whitcomb SJ, Basu A, Allis CD, Bernstein E. (2007). Polycomb Group proteins: an evolutionary perspective. Trends Genet **23**: 494-502.

Whitfield ML, Zheng LX, Baldwin A, Ohta T, Hurt MM, Marzluff WF. (2000). Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. Mol Cell Biol **20**: 4188-4198.

Widlak W, Winiarski B, Krawczyk A, Vydra N, Malusecka E, Krawczyk Z. (2007). Inducible 70 kDa heat shock protein does not protect spermatogenic cells from damage induced by cryptorchidism. Int J Androl **30**: 80-87.

Wiederrecht G, Shuey DJ, Kibbe WA, Parker CS. (1987). The Saccharomyces and Drosophila heat shock transcription factors are identical in size and DNA binding properties. Cell **48**: 507-515.

Woodcock CL, Ghosh RP. (2010). Chromatin higher-order structure and dynamics. Cold Spring Harb Perspect Biol **2**: a000596.

Wu C. (1984). Activating protein factor binds in vitro to upstream control sequences in heat shock gene chromatin. Nature **311**: 81-84.

Wu C, Wilson S, Walker B, Dawid I, Paisley T, Zimarino V, Ueda H. (1987). Purification and properties of Drosophila heat shock activator protein. Science **238**: 1247-1253.

Wu CH, Yamaguchi Y, Benjamin LR, Horvat-Gordon M, Washinsky J, Enerly E, Larsson J, Lambertsson A, Handa H, Gilmour D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes Dev **17**: 1402-1414.

Xiao H, Perisic O, Lis JT. (1991). Cooperative binding of Drosophila heat shock factor to arrays of a conserved 5 bp unit. Cell **64**: 585-593.

Xiao X, Zuo X, Davis AA, McMillan DR, Curry BB, Richardson JA, Benjamin IJ. (1999). HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. EMBO J **18**: 5943-5952.

Xu YM, Huang DY, Chiu JF, Lau AT. (2012). Post-translational modification of human heat shock factors and their functions: a recent update by proteomic approach. J Proteome Res **11**: 2625-2634.

Yan J, Enge M, Whitington T, Dave K, Liu J, Sur I, Schmierer B, Jolma A, Kivioja T, Taipale M, Taipale J. (2013). Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. Cell **154**: 801-813.

Yao J, Munson KM, Webb WW, Lis JT. (2006). Dynamics of heat shock factor association with native gene loci in living cells. Nature **442**: 1050-1053.

Yoshima T, Yura T, Yanagi H. (1998). Heat shock factor 1 mediates hemin-induced *hsp70* gene transcription in K562 erythroleukemia cells. J Biol Chem **273**: 25466-25471.

Zentner GE, Henikoff S. (2013). Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol **20**: 259-266.

Zobeck KL, Buckley MS, Zipfel WR, Lis JT. (2010). Recruitment timing and dynamics of transcription factors at the Hsp70 loci in living cells. Mol Cell **40**: 965-975.

Zou J, Guo Y, Guettouche T, Smith DF, Voellmy R. (1998). Repression of heat shock transcription factor HSF1 activation by HSP90 that forms a stress-sensitive complex with HSF1. Cell **94**: 471-480.

Zuo J, Baler R, Dahl G, Voellmy R. (1994). Activation of the DNA-binding ability of human heat shock transcription factor 1 may involve the transition from an intramolecular to an intermolecular triple-stranded coiled-coil structure. Mol Cell Biol **14**: 7557-7568.

ISBN 978-952-12-3080-6
Painosalama Oy – Turku, Finland 2014