Characterizing the Regulation of Mitochondrial Nucleoids

Thesis by

Yun Elisabeth Wang

In Partial Fulfillment of the Requirements for the degree

of

Degree of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2014

(Defended September 4th, 2013)

ACKNOWLEDGEMENTS

I am most grateful for the support of my advisor, Dr. David Chan. His direction and suggestions throughout the course of my graduate career have been invaluable as I progressed through my projects. I am also grateful to Dr. Hsiuchen Chen, who has been not only a wonderful source of ideas and technical knowledge and provided me with innumerous protocols, but also stood as an outstanding example of a how to be a scientist.

My fellow graduate student, Georgi Marinov, deserves the deepest of thanks for many years of experiment-based advice, computational work, and emotional support, all of which was indispensible. His truly exceptional scientific ability has been a boon to my graduate work.

I would also like to thank my current and past committee members. Dr. Marianne Bronner, Dr. Pamela Bjorkman, Dr. Angela Stathopoulos, and Dr. Barbara Wold have been incredibly helpful with comments, suggestions, and not least, encouragement. To Dr. Barbara Wold, I owe the deepest respect and gratitude as not only a member of my committee but also as my collaborator. I greatly appreciate our discussions and your hard work on our joint projects; it was a pleasure to be able to work with you.

To all the members of the Chan lab, including Dr. Nickie Chan, Dr. Anne Chomyn, Naama Dror, Oliver Loson, Elizabeth Nelson, Dr. Huu Ngo, Priscilla Lee Wah Tee, Raymond Liu, Shuxia Meng, Dr. Prashant Mishra, Dr. Anh Pham, Rebecca Rojansky, Dr. Anna Salazar, and Dr. Chun-Shik Shin, I owe much appreciation for reagents, insightful

discussions, and support. I fully believe that you are some of the most exceptional people, both scientifically and personally. The field of mitochondrial dynamics is very lucky to have you on their team.

I would not be here without my mother and father. In the past twenty-seven years, they have shown me what it is to be the epitome of loving, supportive parents, and I can only hope that I can make them proud. To my younger brother, David, I thank for keeping me young at heart and for constantly reminding me of the joy of discovery.

Last, but not least, to all of my friends here at Caltech, I thank you for all the support throughout my tenure. Constantine Sideris, thank you for not only computational help but also personal support. To Dr. Gloria Sheng, thank you for being the one who could always see the big picture. To my friends - my hypothesis that such a small school could not hold many extraordinary people was proven so wrong.

ABSTRACT

Mitochondria contain a 16.6 kb circular genome encoding 13 proteins as well as mitochondrial tRNAs and rRNAs. Copies of the genome are organized into nucleoids containing both DNA and proteins, including the machinery required for mtDNA replication and transcription. Although mtDNA integrity is essential for cellular and organismal viability, regulation of proliferation of the mitochondrial genome is poorly understood. To elucidate the mechanisms behind this, we chose to study the interplay between mtDNA copy number and the proteins involved in mitochondrial fusion, another required function in cells. Strikingly, we found that mouse embryonic fibroblasts lacking fusion also had a mtDNA copy number deficit. To understand this phenomenon further, we analyzed the binding of mitochondrial transcription factor A, whose role in transcription, replication, and packaging of the genome is well-established and crucial for cellular maintenance. Using ChIP-seq, we were able to detect largely uniform, non-specific binding across the genome, with no occupancy in the known specific binding sites in the regulatory region. We did detect a single binding site directly upstream of a known origin of replication, suggesting that TFAM may play a direct role in replication. Finally, although TFAM has been previously shown to localize to the nuclear genome, we found no evidence for such binding sites in our system.

To further understand the regulation of mtDNA by other proteins, we analyzed publicly available ChIP-seq datasets from ENCODE, modENCODE, and mouseENCODE for evidence of nuclear transcription factor binding to the mitochondrial genome. We

identified eight human transcription factors and three mouse transcription factors that demonstrated binding events with the classical strand asymmetrical morphology of classical binding sites. ChIP-seq is a powerful tool for understanding the interactions between proteins and the mitochondrial genome, and future studies promise to further the understanding of how mtDNA is regulated within the nucleoid.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

Abf2p. ARS-Binding factor

ActB. Beta-actin

AD. Alzheimer's disease

ADOA. Autosomal dominant optic atrophy

AFM. Atomic force microscopy

ALS. Amyotrophic lateral sclerosis

AP-1. Activator protein 1

bp. basepair

BrDU. Bromo-uridine

CEBPβ. CCAT/enhancer binding protein beta

ChIP. Chromatin immunoprecipitation

CMT. Charcot-Marie-Tooth disease

Cybrid. Cellular hybrid

DAPI. 4',6-diamidino-2-phenylindole

D-Loop. Displacement loop

DNA. Deoxyribonucleic acid

DNM1. Dynamin 1 gene

DRP1. Dynamin-related protein gene

E. coli. Escherichia coli

EtBr. Ethidium bromide

FAIRE-seq. formaldehyde-assisted isolation of regulatory elements

FIS1. Fission 1 gene

FISH. Fluorescence in situ hybridization

FRDA. Friedreich's Ataxia

FZO. Fuzzy onions

GAPDH. Glyceraldehyde 3-phosphate dehydrogenase

GR. Glucocorticoid receptor

H2B. Histone 2B

HD. Huntington disease

hTFAM. Human TFAM

Kb. Kilobasepair (1000 basepairs)

iB5s. Integrin beta-5 subunit

IHF. Integration host factor

MafF. V-maf avian musculoaponeurotic fibrosarcoma oncogene homolog F

MafK. V-maf avian musculoaponeurotic fibrosarcoma oncogene homolog K

mtDNA. Mitochondrial DNA

MEF. Mouse embryonic fibroblast

MERRF. Myoclonic Epilepsy with ragged red fibers

MFN1. Mitofusin 1 gene

MFN2. Mitofusin 2 gene

mGH. Mouse growth hormone

mTFAM. Mouse TFAM

mtHSP70. Mitochondrial heat shock protein 70

MTS. Mitochondrial targeting sequence

NAP. Nuclear associated protein

NCR. Non-coding region

NFE2. Nuclear factor, erythroid 2

OPA1. Optic atrophy 1 gene

O_L. Origin of light strand replication

ORF. Open reading frame

PALM. Photoactivated localization microscopy

PD. Parkinson's disease

PECAM1. Platelet endothelial cell adhesion molecule 1

POLG. Mitochondrial polymerase gamma

POLG2. Mitochondrial polymerase gamma subunit 2

PPARγ2. Peroxisome proliferator-activated receptor gamma 2

PPIF. Peptidylprolyl isomerase F/cyclophilin F

qPCR. Quantitative polymerase chain reaction

RPM. Reads per million

SDS-PAGE. Sodium dodecyl sulfate – polyacrylamide gel electrophoresis

SILAC. Stable isotope labeling by amino acids in cell culture

STAT1/3/5. Signal transducer and activator of transcription 1/3/5

STED. Stimulated emission depletion microscopy

STORM. Stochastic optical reconstruction microscopy

TFAM. Transcription factor A, mitochondrial; gene

TFIB. Transcription factor IB; mitochondrial gene

TFIIB. Transcription factor IIB; mitochondrial gene

TF. Transcription factor

TPM. Tethered particle motion

C h a p t e r   1

INTRODUCTION

I.   Background

Mitochondria most likely arose as a result of an endosymbiotic event (Sagan 1967) between the ancestor of modern eukaryotes and a member of the α-proteobacteria clade (Yang, Oyaizu et al. 1985), and are now present in all known species of eukaryotes. In modern eukaryotes, mitochondria serve as the epicenter of oxidative phosphorylation, which utilizes the ETC to create a proton gradient across the inner mitochondrial membrane to drive production of ATP using ATP synthase. Thus, mitochondria containing functional metabolic machinery are necessary for cellular survival. Interestingly, reflective of the organelle's prokaryotic ancestry, mitochondria retain their own circular DNA genome (Nass, Nass et al. 1965, van Bruggen, Borst et al. 1966), although its size has been greatly reduced in many eukaryotes through transfer of genes to the nucleus. In present-day eukaryotes, transcription and translation of the nuclear components of the mitochondrial proteome occurs in the nucleus and cytoplasm, respectively, and the components are then imported into the mitochondrial to modulate organellar function. These components include that of the separate mitochondrial transcription, replication and regulatory machineries, a number of which retain evidence of their prokaryotic origin (Szklarczyk and Huynen 2010). Given the relatively simple architecture as well as its importance for the cell, mtDNA has been extensively studied for much of the 50 years since its discovery.

However, although we now have some understanding of how some processes modulate mtDNA integrity, much still remains to be understood about the dynamics of the processes underlying these phenotypes.

The mitochondrial genome is an essential component of eukaryotic cells

mtDNA was first visualized by electron microscopy in chick embryos, a finding that was biochemically confirmed by the finding of DNA in the highly purified mitochondria in the eukaryotes *Neurospora crassa* and *S. cerevisiae* (Luck and Reich 1964, Schatz, Haslbrunner et al. 1964). In humans, the 16.6kb mitochondrial genome only encodes for 13 proteins, as well as the 22 mitochondrial tRNAs and 2 mitochondrial rRNAs (Anderson, Bankier et al. 1981). However, these products are necessary and required for the organelle's bioenergetics functions because the proteins, translated using the encoded tRNAs and rRNAs, are essential components of the ETC. Remarkably, although the mitochondrial genome only encodes for 27 products, mass spectrometry analysis of the *S. cerevisiae* mitochondrial proteome has identified approximately 25% of the 1100 proteins that are imported into mitochondria as being involved in the maintenance of mtDNA (Sickmann, Reinders et al. 2003), highlighting the importance of the genome.

It was clear from the first that there were many copies of mtDNA per cell, far beyond that of the nuclear genome. While some initial studies involved lengthy purification of mitochondria followed by quantification of DNA yield, at great risk of

nuclear contamination (Borst and Kroon 1969, Nass 1969), more successful quantitative analysis of copy number was performed utilizing thymidine kinase-negative HeLa cells, which selectively incorporate thymidine into mtDNA. This analysis yielded an estimate of 8800mtDNA/HeLa cell (Bogenhagen and Clayton 1974). Later experiments by Satoh and Kuroiwa in A2780 ovarian carcinoma cells first utilized visualization of mtDNA, through DAPI staining, to effect quantification, using a video-intensified photon counting microscope system, or VIM system. The VIM system was utilized to enhance the low signal intensity of the mtDNA staining, and also to allow for quantitation of the mtDNA by photon counting, using the T4 phage genome as a benchmark. Using this method, about 500 mtDNA copies per cell were counted (Satoh and Kuroiwa 1991). Quantification utilizing other techniques, such as Southern blot (Shmookler Reis and Goldstein 1983, Tang, Schon et al. 2000), ethidium bromide staining (Iborra, Kimura et al. 2004), and qPCR (Legros, Malka et al. 2004, Wai, Ao et al. 2010, Brown, Tkachuk et al. 2011), has resulted in a known range of copy numbers per human cell from the hundreds to hundreds of thousands (Garrido, Griparic et al.).

Although cells exhibit a range of mtDNA levels, it is clear that mtDNA copy number is far in excess to that of nuclear DNA. In order to replicate the large number of copies required in time for cell division (as well as drive the closely linked transcription of the genome), while nuclear DNA replication is tightly timed via the cell cycle, mtDNA replication occurs constantly (Bogenhagen and Clayton 1977, Magnusson, Orth et al. 2003), with some peaks at S and G2 (Lee, Kim et al. 2007, Chatre and Ricchetti 2013). However, mtDNA copy number within cells is tightly controlled, albeit by largely

uncharacterized mechanisms (Clay Montier, Deng et al. 2009). While cells requiring low amounts of energy, such as spleen and endothelial cells, consistently maintain a low copy number, cells requiring high levels of ATP, such as neurons and muscle cells, consistently exhibit relatively high levels of mtDNA (Williams 1986, Moyes, Battersby et al. 1998). Cells thus regulate mtDNA copy number in order to meet energy needs, and the requirement for such tight regulation becomes evident when observing the effects of loss of effective mtDNA copy number in cells and organisms.

Pathologies of mtDNA instability

When effective mtDNA copy number is perturbed, whether through mutation, depletion, or deletion, an expected phenotype is not always observed. This is due to a unique quality of mitochondrial disorders resulting from mutations or deletions in the genome: the phenotypic "threshold effect." (Rossignol, Faustin et al. 2003) The number of genomes per cell can number between the hundreds and the thousands (Bogenhagen and Clayton 1974, Satoh and Kuroiwa 1991), leading to a large number of templates from which transcription and translation may take place. Furthermore, mitochondria constantly exchange their contents through fusion and fission. Thus, there is enough wild-type protein product distributed across mitochondria to maintain normal function until a certain threshold of mutated DNA is reached, upon which there is a precipitous decline in a number of measures of mitochondrial activity, such as respiration rates, ATP production, or phenotype of interest (Rossignol, Faustin et al. 2003, Chen, Chomyn et al.

2005). This threshold has been assayed in culture both by analysis of cybrids containing varying levels of mtDNA mutations as well as by analyzing primary patient samples. For example, in cells containing the A8344G tRNA^{Leu} mutation causing MERRF, respiration was found to be impacted only when mutant mtDNA levels were between 85-95% of the total mtDNA per cell (Shoffner, Lott et al. 1990, Zhou, Chomyn et al. 1997, Moslemi, Tulinius et al. 1998), values typical of that found with other mtDNA mutations, with the usual range being between 60-95% (Rossignol, Faustin et al. 2003).

While the phenotypic threshold effect means that any single mtDNA mutation event does not lead to pathology, the accumulation of these mutations has significant phenotypic effects. Inherited mtDNA mutations lead to cellular respiratory defects, typically in the tissues with the highest rate of aerobic metabolism. This results in the phenotypes of neuromuscular impairment, movement disorders, myopathy, eye disorders, and other systemic manifestations, albeit often with high symptom variability between patients with the same mutation (Zeviani and Di Donato 2004). For example, a point mutation in tRNA^{leu} (Goto, Nonaka et al. 1990, DiMauro and Schon 2003), A3243G, can cause the symptoms associated with multiple disorders, such as mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes, progressive external ophthalmoplegia, isolated myopathy, or cardiomyopathy (Chinnery and Turnbull 1999, Leonard and Schapira 2000). Such heterogeneity in phenotype of a single mutation is primarily postulated to be due to differential heteroplasmy between tissues in individuals resulting from unequal segregation upon mitosis (Lombes, Aure et al. 2014).

Damage or depletion of mtDNA may also causes numerous inherited disorders, including Alpers' Disease, ataxia neuropathy spectrum, progressive external ophthalmoplegia, and mitochondrial depletion syndrome (Suomalainen and Isohanni 2010, Stumpf, Saneto et al. 2013). Furthermore, loss and damage to mtDNA has been implicated in cardiovascular disease (Sugiyama, Hattori et al. 1991, Ide, Tsutsui et al. 2001, Karamanlidis, Nascimben et al. 2010, Karamanlidis, Bautista-Hernandez et al. 2011), diabetes (Maassen, LM et al. 2004, Simmons, Suponitsky-Kroyter et al. 2005, Gauthier, Wiederkehr et al. 2009), neurodegenerative disorders such as Alzheimer's (Coskun, Beal et al. 2004, Coskun, Wyrembak et al. 2012, Rice, Keeney et al. 2014), and aging (Corral-Debrinski, Shoffner et al. 1992, Trifunovic and Larsson 2008). Strikingly, a ~40% increase in mtDNA copy number in aging transgenic mice overexpressing human transcription factor A leads to significantly improved motor learning memory as assayed by latency to fall on the rotorod test, and to improved working memory as measured by the number of errors during the radial arm water maze (Ikeuchi, Matsusaka et al. 2005, Hayashi, Yoshida et al. 2008). Similar promotion of cell survival or function has been seen with increasing mtDNA copy number in other models of disease associated with decreased mtDNA abundance, such as diabetes (Suarez, Hu et al. 2008, Gauthier, Wiederkehr et al. 2009), Alzheimer's (Xu, Zhong et al. 2009), and Parkinson's (Keeney, Quigley et al. 2009, Piao, Kim et al. 2012), demonstrating the strong relationship between defective mtDNA copy number and pathology. Considering that loss of mtDNA leads to pathology, it is critical to understand how mtDNA copy number and integrity are maintained.

Organization of the Mitochondrial Genome

Because mtDNA is relatively simply organized, the genome has been studied extensively. Mitochondrial genes are densely packed along the genome, with the notable exception of the non-coding D-loop (Tang, Schon et al. 2000), which is located within the ~1kb NCR. Transcription initiates in the D-loop, is carried out by the mitochondrial-specific RNA polymerase POLRMT, and results in long polycistronic transcripts on each strand (the H-strand and the L-strand), from the light strand promoter LSP and two heavy strand promoters, HSP1 and HSP2 (Cantatore and Attardi 1980, Montoya, Christianson et al. 1982). The D-loop also contains two putative origins of leading strand replication (O_H and Ori-b), while another putative origin of replication, the lagging strand origin of replication (O_L), is located downstream of the ND2 ORF (Crews, Ojala et al. 1979, Clayton 1982, Yasukawa, Yang et al. 2005). Thus, the D-loop is considered the primary regulatory region of the mitochondrial genome.

Given the 16.6 kb size of the human mitochondrial genome, which translates to approximately 5μm in diameter, and that mitochondria are only ~2μm in diameter, it was predicted that mtDNA would be compacted by a specific mechanism in order to pack into the mitochondrial matrix, a task presumably performed by proteins. The first visualization of mtDNA localized to discrete loci within mitochondria was through DAPI staining of HeLa cells (Satoh and Kuroiwa 1991). Subsequent experiments utilizing PicoGreen, EtBr, BrdU, and FISH for detection of DNA have confirmed this finding (Davis and Clayton

1996, Bereiter-Hahn and Vöth 1997, Margineantu, Gregory Cox et al. 2002, Alam, Kanki et al. 2003, Garrido, Griparic et al. 2003, Magnusson, Orth et al. 2003, Gilkerson, Schon et al. 2008). Use of super-resolution microscopy, including STED, STORM, and PALM, has allowed for measurement of the nucleoid at ~100nm in mammals using fluorescently labeled TFAM, or anti-DNA, anti-TFAM, or anti-BrdU immunostaining (Brown, Tkachuk et al. 2011, Kukat, Wurm et al. 2011). Subtracting the volume of the antibodies used in labeling, the final diameter of each nucleoid is estimated to be ~70nm – amenable to the mitochondrial matrix.

Quantification of mtDNA copy number as discussed previously, in conjunction with immunocytochemistry, has yielded estimates of 1 − 10 copies of nucleoids per cell in various cell types (Satoh and Kuroiwa 1991, Cavelier, Johannisson et al. 2000, Iborra, Kimura et al. 2004, Legros, Malka et al. 2004, Gilkerson, Schon et al. 2008). The first estimate of mtDNA/nucleoid in human cells, using DAPI staining in A2780 cells, used the assumption of a stepwise function of fluorescence intensity per mtDNA molecule to estimate copy number in each nucleoid, and found an average of 1.4 genomes per nucleoid, or about 4.6 mtDNA per organelle (Satoh and Kuroiwa 1991). However, this study, as well as subsequent papers using this technique as well as studies estimating by dividing average copy number per cell by average nucleoid number per cell, have all suffered from problems stemming from the detection limit for nucleoids in immunocytochemistry. Using STED, which detects ~1.6-fold more nucleoids due to increased sensitivity beyond that of the detection limit of conventional confocal microscopy, it has been estimated that there are an average of 1.4 genomes per nucleoid in human primary fibroblasts, meaning that the

majority of nucleoids contain only one genome (Kukat, Wurm et al. 2011). Calculations in 3T3 mouse fibroblast cells have estimated that nucleoids occupy an average of 8.3 x 10^5 nm^3 (Brown, Tkachuk et al. 2011). Given that Brown et al. estimates that each nucleoid contains 3 mtDNAs, only 6.3% of the volume of each nucleoid is composed of DNA; the remainder is occupied by protein. Thus, the vast majority of the mitochondrial nucleoid is composed of the proteins that may modulate and interact with the mitochondrial genome.

The Protein Components of the Mitochondrial Nucleoid

Eukaryotic cellular viability is dependent upon the proliferation of mitochondria and of its essential genome, which must be replicated constantly so that the copy number per cell may roughly double in order to be ready for cell division (Magnusson, Orth et al. 2003). It is safe to presume that such activities are effected by proteins that associate with the genome. Due to analogous paradigms of the yeast mtDNA as well as the ancestral bacterial nucleoid (Delius and Worcel 1974), it was long postulated that proteins would associate with the mammalian nucleoid. However, while mtDNA-protein complexes had been purified previously (Barat, Rickwood et al. 1985), it was only ten years after the first visualization of mtDNA puncta within cells that proteins were visually localized to the nucleoids, when an eGFP-tagged form of the mitochondrial helicase Twinkle was found to colocalize with an anti-DNA antibody (Spelbrink, Li et al. 2001). Soon afterward, TFAM, mtSSB, and POLG/POLG2 were also identified as components of the nucleoid via colocalization with Twinkle-eGFP and with BrdU incorporation (Garrido, Griparic et al.

2003). The particularly strong interactions between TFAM and mtSSB and mtDNA have been exploited in order to further identify other proteins associated with the nucleoids in HeLa cells via immunoprecipitation followed by mass spectrometry (Wang and Bogenhagen 2006), or by formaldehyde crosslinking, sedimentation, and fractionation, followed by liquid chromatography-tandem mass spectrometry (Bogenhagen, Rousseau et al. 2008), yielding more large-scale identification of associated proteins. From this data, Bogenhagen et al. have proposed a "layered" model of protein organization within the nucleoid, where the "core" proteins (e.g., TFAM, mtSSB, TFB1M, TFB2M, mTERF), which were identified via both the immunoaffinity and the harsher formaldehyde methods, are tightly associated with the nucleoid, and are surrounded by more loosely-associated nucleoid proteins which were not captured by the formaldehyde method (e.g., HSP60, PHB1/2, ATAD3A).

In addition to the proteins identified by Bogenhagen et al., a majority of which can be related directly to maintenance and regulation of mtDNA and its associated proteins, further studies of mitochondrial proteins have identified a number that colocalize to nucleoids by immunocytochemistry. While mtRNA polymerase, TFB1M, TFB2M, mTERF1, topoisomerase 1 and Twinkle, all members of the transcription machinery, were identified by Bogenhagen et al., mTERF2 and mTERF3 have separately been identified as binding to the genome (Park, Asin-Cayuela et al. 2007, Pellegrini, Asin-Cayuela et al. 2009). Other classes of proteins, such as helicases (Wang, Shu et al. 2009, Szczesny, Borowski et al. 2010), RNA binding proteins (Sondheimer, Fang et al. 2010), chaperone proteins (Kaufman, Kolesar et al. 2003), proteases (Fu and Markovitz 1998), ribosomal

proteins (Rorbach, Richter et al. 2008), enzymes involved in lipid metabolism (Wang and Bogenhagen 2006, Bogenhagen, Rousseau et al. 2008), and even cytoskeletal attachment proteins (Reyes, He et al. 2011), have been identified as associating with the nucleoid. It requires little stretch of the imagination to postulate about the function of the localization some of these protein classes, such as RNA binding proteins and ribosomal proteins, to the nucleoid. However, the purpose of the enzymes involved in lipid metabolism and cytoskeletal attachment proteins, for example, is less obvious and remains an area to be explored. Interestingly, the proteins PHB1 (Kasashima, Sumitani et al. 2008) and ATAD3A (He, Mao et al. 2007, Holt, He et al. 2007), found at the periphery of the nucleoid, have already been implicated in mtDNA stability and nucleoid morphology, although the exact mechanisms still remain to be explored.

While the methods of protein-associated nucleoid identification described above have most probably identified the most common proteins found in the nucleoid, less abundant and more transiently associated proteins are more difficult to capture. One such set of proteins that has been identified as associating with mtDNA through other methods is the canonically nuclear transcription factors, a few of which may directly interact with mtDNA. The first such transcription factors identified were GR, p53, and p43, which were shown to localize to within mitochondria (Demonacos, Tsawdaroglou et al. 1993, Caelles, Helmberg et al. 1994, Wrutniak, Cassar-Malek et al. 1995). However, it was not until later, when new techniques for elucidating *in vivo* binding of proteins to DNA, that the mitochondrial roles of these transcription factors, independent of nuclear effect, could be determined. For example, CREB is present in mitochondrial fractions following subcellular

subfractionation, localizes to the mitochondria via immunoelectron microscopy (Cammarota, Paratcha et al. 1999), and has been shown to be imported into isolated rat liver mitochondria via the TOM translocation complex with mtHSP70 as an associated chaperone (Lee, Kim et al. 2005, De Rasmo, Signorile et al. 2009), providing compelling evidence for mitochondrial localization. Furthermore, expression of a mitochondrially targeted form of CREB results in increased expression of the ND2, ND4, and ND5 on the genome; conversely, a dominant negative form of CREB decreases expression of the mitochondrial genes and causes a concomitant decrease in ETC complex I activity (Lee, Kim et al. 2005), suggesting that CREB independently modulates expression of the mtDNA. Finally, *in vivo* ChIP and *in vitro* footprinting data have shown that CREB binds to three predicted binding sites in the D-loop (Lee, Kim et al. 2005). Similar verification of correct localization, independent mitochondrial modulation, and mtDNA-specific binding has also been shown for p43 (Wrutniak, Cassar-Malek et al. 1995, Enriquez, Fernandez-Silva et al. 1999, Fernandez-Vizarra, Enriquez et al. 2008), ERα/β (Casas, Rochard et al. 1999, Monje and Boland 2001, Chen, Delannoy et al. 2004, Jazbutyte, Kehl et al. 2009, Milanesi, Vasconsuelo et al. 2009), and p53 (Heyne, Mannebach et al. 2004, Achanta, Sasaki et al. 2005), with some evidence for mitochondrial-specific action by STAT3 (Wegrzyn, Potla et al. 2009). Such modulation of both mitochondrial and nuclear genomes could be favorable, allowing the cell to simultaneously regulate both genomes in response to a single stimulus. Nonetheless, it is apparent that interactions between the mitochondrial genome and canonically nuclear transcription factors are not unprecedented, and are phenomena whose implications remain to be fully understood.

TFAM is a Site-Specific Mitochondrial Transcription Factor

Of the many proteins identified to date as interacting with the mitochondrial genome, the HMG-box containing protein TFAM is the most abundant, best characterized, and arguably one of the most crucial. Because of the prokaryotic origin of mitochondria, TFAM is particularly important due to its role not only in its site-specific role as a transcription factor, but also in its nonspecific, "scaffolding" capabilities in the nucleoid. Clues as to the function of TFAM come from the ancestral prokaryotic and fungal systems. Bacterial chromosomes such as that of *E. coli* , which spans 4.6 million bp, lack the sophisticated histone-based organization of eukaryotic nuclear genomes. Instead, a class of proteins called NAPs, such as HU and IHF, distribute across the genome and effect an architectural role, performing the compacting function required to fit the genome within the cell and forcing DNA into a 160° turn (Rice, Yang et al. 1996, Mouw and Rice 2007, Dillon and Dorman 2010). The HU family of proteins, in particular, may further compact the genome via its ability to dimerize with itself (Rice, Yang et al. 1996). The first protein in eukaryotes found to have a similar function was the yeast Abf2p, an HMG box family protein which was found to localize to mitochondria, to be required for mtDNA integrity (Diffley and Stillman 1991), and to bind to DNA with a footprint of ~27bp (Diffley and Stillman 1992). In addition, AFM of Abf2p in conjunction with DNA has shown that Abf2p packages and compacts DNA (Brewer, Friddle et al. 2003). The mammalian ortholog of Abf2p is TFAM, which has a role not only in the packaging of mtDNA similar to that of the NAPs and Abf2p, but also in mitochondrial transcription.

TFAM, which consists of two DNA-binding HMG boxes separated by a linker and followed by a C-terminal tail, was first purified and identified by Fisher and Clayton (Fisher and Clayton 1985, Fisher and Clayton 1988, Parisi and Clayton 1991). The transcriptional function of TFAM relies on its binding specifically to the LSP and HSP promoter regions of the D-loop, where the HSP1, HSP2, and LSP transcripts originate. Initial DNase I footprinting experiments revealed that TFAM bound preferentially to a region close to the LSP transcription start site, in a part of the promoter sequences that was found by systemic deletional analysis to be necessary for augmented efficiency of transcriptional initiation (Chang and Clayton 1984). TFAM was also localized to the HSP1 promoter, yielding a ~20-30 bp footprint at -12 to -39 of the respective transcription start sites (Fisher and Clayton 1988); no footprint upstream of the HSP2 start site has been detected to date. Interestingly, there is significant homology between the primary binding sites of TFAM, LSP and HSP1, only when the HSP1 sequence is inverted in direction, suggesting that TFAM must either act bidirectionally, or bend ~180° to contact the polymerase at either LSP or HSP. Finally, a single, specific TFAM binding site has been localized in rat mtDNA upstream of the O_L (Gadaleta, D'Elia et al. 1996, Cingolani, Capaccio et al. 1997, Pierro, Capaccio et al. 1999), and is postulated to tie into the role of TFAM in replication.

The classic view of the minimal mtDNA transcriptional machinery is that it contains three essential components: TFAM, POLRMT, and TFB2M. Although it is a homolog of the self-initiating prokaryotic T7 RNA polymerase, the mitochondria-specific POLRMT cannot initiate transcription on its own. Indeed, purified TFAM and POLRMT

together are also insufficient to activate transcription in *in vitro* experiments (Falkenberg, Gaspari et al. 2002). However, transcription can be reconstituted using purified TFAM and a partially purified mitochondrial fraction *containing* POLRMT*,* indicating that a protein associated with POLRMT is also required for activation (Falkenberg, Gaspari et al. 2002). This observation led to the identification of TFB1M and TFB2M, which have significant homology to the yeast essential transcription factor MTF1 as well as the prokaryotic rRNA dimethyltransferases. While both TFB1M and TFB2M, in the presence of both TFAM and POLRMT, were able to induce *in vitro* transcription in a run-off assay, TFB2M produced ~10-fold greater transcript levels (Falkenberg, Gaspari et al. 2002, Litonin, Sologub et al. 2010). Given the stronger homology of TFB1M to the bacterial rRNA dimethyltransferases, it was hypothesized that while TFB1M could partially supplant the role of TFB2M as a transcription factor as necessary, its primary role is in 12S rRNA methylation needed for translation. The hypothesized far stronger methyltransferase activity of TFB1M would later be demonstrated to be indeed correct *in vitro* (Seidel-Rogol, McCulloch et al. 2003, Cotney and Shadel 2006) as well as in drosophila, mouse, and human systems (Matsushima, Garesse et al. 2004, Matsushima, Adan et al. 2005, Cotney, Wang et al. 2007, Cotney, McKay et al. 2009, Metodiev, Lesko et al. 2009).

A commonly accepted theory of transcriptional initiation involves initial site-specific binding of TFAM to a locus in the promoter. The sharp U-turn that TFAM induces in DNA aligns its C-terminal tail to the transcriptional machinery (Ngo, Kaiser et al. 2011, Rubio-Cosials, Sidow et al. 2011, Ngo, Lovely et al. 2014), and the

POLRMT/TFB2M complex, which also maintains some site-specific recognition (Falkenberg, Gaspari et al. 2002, Gaspari, Larsson et al. 2004, Shi, Dierckx et al. 2012), is then recruited to the TFAM binding site, causing initiation of transcription. Beyond direct biochemical evidence discussed previously, this view of initiation is supported by several pieces of evidence. First, mutations to the specific TFAM binding sites leads to inactivation of promoter activity (Dairaghi, Shadel et al. 1995). Additionally, the C-terminal tail of TFAM is particularly important for induction of transcription; ablation leads to loss of ability to initiate transcription (Kanki, Ohgaki et al. 2004), due in part to decreased ability to preferentially distort DNA (Malarkey, Bestwick et al. 2012), or perhaps due to inability to associate with TFB2M and form the complete initiation complex (McCulloch and Shadel 2003). Of note, more recent evidence suggests that TFAM may even directly interact with POLRMT to form the initiation complex (Yakubovskaya, Guja et al. 2014). Because of this domain-specific effect, the ~180° turn that TFAM imparts on DNA is crucial for initiation at LSP but not at HSP1 (Malarkey, Bestwick et al. 2012, Ngo, Lovely et al. 2014), where TFAM is oriented in reverse to its position in LSP. While TFAM bends DNA similarly at HSP1 (Ngo, Lovely et al. 2014), because the C-terminus is located after the proximal HMG box domain, it remains close to the transcriptional machinery despite the distortion. Indeed, deficits in DNA bending by TFAM lead to selective loss of transcription from LSP, but not from HSP1 (Ngo, Kaiser et al. 2011). Finally, when the distance between TFAM binding site and transcriptional machinery binding is lengthened, efficiency of transcription initiation is greatly disrupted, demonstrating the importance of the specificity of initial binding of TFAM for transcriptional nucleation (Dairaghi, Shadel et al. 1995). Taken together, the

site- and structure-specific properties of TFAM binding at the LSP and HSP1 promoters appear to be essential for initiation of transcription.

Interestingly, there has been some recent debate about whether TFAM is truly required for transcriptional activation. Using an *in vitro* system where all proteins were *E. coli*-derived and not epitope-tagged, it was demonstrated that presence of TFB2M and POLRMT alone were able to drive transcription from both HSP1 and LSP start sites, with transcription from HSP1 being just as or more efficient (Shutt, Lodeiro et al. 2010). Interestingly, similar results have been observed by another group as well, although they were not addressed in the publication (Litonin, Sologub et al. 2010). Although there is extensive research, discussed above, showing that the basal transcriptional machinery requires TFAM in addition to POLRMT and TFB2M, Shutt et al. suggest that the use of recombinant, non-epitope tagged proteins which were not co-expressed in insect cells, as per previous experiments, as well as HSP1-specific analysis of transcriptional initiation, contributed to the novel results.

Given the significant implications for the field, a direct rebuttal of the Shutt et al. work was shortly forthcoming. Utilizing an analogous *in vitro* transcription system, Shi et al. replicate the canonically accepted results, demonstrating that specific TFAM binding upstream of the start site is necessary to initiate transcription, whether the protein source is mitochondrial extract, epitope-tagged purified proteins, or recombinant non-epitope tagged proteins (Shi, Dierckx et al. 2012). In addition, the authors tested parameters which could lead to the results produced by Shutt et al. Namely, because it is thought that

part of TFAM's function is to aid in "melting" DNA in order to permit POLRMT/TFB2M binding, conditions that have previously been shown to allow for promoter breathing, e.g., decreased salt concentration (Metzler and Ambjornsson 2005) and negatively supercoiled templates (Parvin and Sharp 1993), were tested in the *in vitro* system. Strikingly, when NaCl levels were decreased to less than 12mM, transcription was increasingly enabled from both LSP and, to a greater extent, HSP1, mirroring the previous results. Additionally, a comparison of transcription using a linear template as compared to a negatively supercoiled, circular template, demonstrated that negative supercoiling was also transcription permissive. This effect could be abolished by treatment with mitochondrial topoisomerase I, which relaxes the supercoils. Similar permissiveness in negatively supercoiled DNA has also previously been demonstrated in the nuclear transcription with RNA polymerase II (Parvin and Sharp 1993). Given that Shi et al. performed their experiments in buffer conditions lacking NaCl, it is reasonable to conclude is that while TFAM is usually a necessary component of the basal transcriptional system, conditions in which DNA is more "melted" can result in relaxed requirements for TFAM binding.

Regardless of whether TFAM is absolutely required for transcriptional activation, it still plays an important role in transcription, and thus in the dynamics of differential transcription from the transcriptional initiation site. Initial *in vitro* footprinting assays found that much higher concentrations of TFAM were required to visualize footprinting at HSP1 as compared to LSP (Fisher and Clayton 1988), possibly due to additional DNA-protein contacts at LSP that do not exist when TFAM binds to either HSP1 or nonspecific

DNA (Ngo, Lovely et al. 2014). Subsequent work has also shown that, in the presence of POLRMT and TFB2M, increasing the concentration of TFAM very quickly maximizes transcription levels from LSP, while much higher concentrations are required to drive transcription from HSP1 (Falkenberg, Gaspari et al. 2002), in line with the footprinting results. Interestingly, binding affinity studies have consistently shown that TFAM has a lower K_D at LSP than at HSP via FRET and SPR (Kaufman, Durisic et al. 2007, Malarkey, Bestwick et al. 2012). However, when TFAM concentrations rise further, transcript levels drop precipitously, a phenomenon that is potentially attributable to TFAM's nonspecific binding properties (Dairaghi, Shadel et al. 1995, Shutt, Lodeiro et al. 2010). Furthermore, it is possible that under certain permissive conditions, when TFAM levels are extremely low, transcription progresses at a low level from both promoters (Shutt, Lodeiro et al. 2010); there is some *in vivo* evidence for this, where transcription levels in mice where mouse TFAM has been replaced by the largely transcription-incompetent human version still show low levels of transcription (Freyer, Park et al. 2010), although this may be attributed to any residual transcription factor activity on the part of human TFAM. Other *in vivo* experiments related to TFAM-induced transcription levels are mixed, with some showing that overexpression of TFAM leads to preferential transcription from LSP (Ekstrand, Falkenberg et al. 2004), while others see transcription from both promoters (Gensler, Weber et al. 2001, Garstka, Schmitt et al. 2003). However, *in vivo* data is complicated by the inability to control initial concentrations of proteins as well as the presence of any transcriptional modulators, rendering such experiments difficult to interpret. Given the current evidence, however, it

is certainly probable that TFAM levels could in part help regulate levels of expression from the promoters.

A possible model of TFAM regulation attempting to reconcile the existing evidence (Shutt, Bestwick et al. 2011) speculates that low levels of TFAM enable low levels of transcription from both LSP HSP1, allowing for sufficient gene expression for maintenance. Higher levels of TFAM would preferentially activate LSP, driving relatively higher levels of replication through creation of the requisite primer. Even higher levels would drive maximal transcription from both promoters for high-level transcription and replication, while saturating levels would lead to transcriptional inhibition at both primers. Such a model where TFAM modulates the genome would only make sense in the context of the cell's ability to regulate TFAM levels between nucleoids. Indeed, there is significant evidence that this occurs. Analysis of nucleoids via microscopy reveals that TFAM is heterogeneously distributed across nucleoids, with some nucleoids containing levels below the detection limit (Wang and Bogenhagen 2006). Additionally, measurement of replication in individual nucleoids has demonstrated that actively replicating nucleoids have generally less TFAM than those that were less active (Wai, Teoli et al. 2008). Interestingly, in the yeast transcription system, mtDNA became less compacted and relative Abf2p ratios decreased when yeast were placed under growth-permissive conditions (Kucej, Kucejova et al. 2008), further lending credence to the theory that TFAM modulates regulation of levels of transcription. However, more work needs to be performed to confirm the role of TFAM at different

concentrations and under different conditions, ideally with maximum physiological relevance, in order to elucidate the intricacies of this regulatory system.

A third promoter, HSP2, also exists within the D-loop; however, its transcriptional activity is poorly characterized compared to that of LSP and HSP1. Indeed, although analysis of mitochondrial RNA molecules had previously identified HSP2 as a transcriptional start site (Montoya, Christianson et al. 1982, Montoya, Gaines et al. 1983), attempts at further characterization of transcription from the promoter had inconsistent results (Martin, Cho et al. 2005, Litonin, Sologub et al. 2010), with an *in vitro* transcription system capable of inducing transcription at LSP unable to produce an HSP2 transcript. However, it was found that upon isolation of HSP2 from the HSP1 promoter, transcription could be reconstituted using both recombinant proteins and cell mitochondrial extracts, albeit at ~100-fold reduced levels compared to HSP1 (Lodeiro, Uchida et al. 2012, Zollo, Tiranti et al. 2012). Interestingly, addition of TFAM at levels where LSP and HSP1 are activated, even to concentrations far below a 1:1 ratio between TFAM and POLRMT/TFB2M, results in repression of transcription at HSP2. The mechanisms behind this differential response to TFAM are largely unknown, but are dependent on residues 220-236 of the TFAM C-terminal tail. This new knowledge about the function of TFAM in regulating transcription of the HSP2 promoter lends further credence to the hypothesis that TFAM levels within nucleoids are regulated in order to control expression of the mitochondrial genome.

Beyond its role as a transcription factor, TFAM is also essential for normal replication. The mechanism by which mtDNA is replicated is hotly debated, with competing theories postulating that replication follows the strand-asymmetric model, strand-symmetric model, and/or RITOLS model of replication (Kasamatsu and Vinograd 1973, Clayton 1982, Holt, Lorimer et al. 2000, Brown, Cecconi et al. 2005, Yasukawa, Reyes et al. 2006). Regardless of the specific mechanism by which mtDNA replicates, however, it is well-accepted that the minimum *in vitro* replisome consists of the subunits of Polγ, helicase Twinkle, and mtSSB (Spelbrink, Li et al. 2001, Falkenberg, Larsson et al. 2007, Falkenberg and Larsson 2009, Lee, Kennedy et al. 2009, Milenkovic, Matic et al. 2013). However, even with the minimum machinery, replication from O_H is dependent on the LSP transcript, and RNA primer formation is also required at O_L for light strand replication to occur. As formation of significant amounts of RNA transcripts in mitochondria most probably requires TFAM, it is essential for replication (Chang and Clayton 1984, Chang and Clayton 1985).

Due to its role in replication, TFAM is required for normal levels of mtDNA. Mice heterozygous for a knockout of TFAM exhibit not only an expected reduction of 22% in mitochondrial transcript levels in the heart and kidney, but also a universal 34% reduction in mtDNA copy number across all assayed tissues. Furthermore, homozygous knockout mice have no detectable levels of mtDNA and die during embryogenesis (Larsson, Wang et al. 1998), highlighting the importance of TFAM in maintenance of mtDNA levels and in cellular and organismal viability. Interestingly, overexpression of TFAM at low levels leads to an increase in mtDNA copy number (Ekstrand, Falkenberg

et al. 2004), suggesting that upregulation of TFAM alone can singlehandedly enhance replication and thus maintenance of the genome.

TFAM Functions as a Nonspecific Histone-Like Protein for mtDNA

Beyond the role of TFAM as a transcription factor, TFAM also acts as a scaffolding protein for the genome in much the same way that Abf2p does in the yeast system, and that the NAPs do in the bacterial. Recently, the crystal structure of TFAM has been solved, and has been shown to indeed bind to ~22bp of DNA (Ngo, Kaiser et al. 2011, Rubio-Cosials, Sidow et al. 2011). It also bends the DNA into a 180° turn, similar to that of bacterial IHF, which compacts at 160° (Rice, Yang et al. 1996). Although these structures were derived from binding of TFAM to its LSP binding site, an instance of specific binding, TFAM is also thought to organize the rest of the genome by binding nonspecifically in a similar manner. Beyond numerous studies colocalizing TFAM and nucleoids, TFAM has been shown to be able to bind to multiple DNA sequences, both by methylation interference studies (Fisher, Parisi et al. 1989) and by the crystal structure of TFAM bound to a nonspecific sequence, which has recently been solved (Ngo, Lovely et al. 2014). Furthermore, estimates of the average number of TFAM molecules per unit of mtDNA have suggested that it is sufficiently abundant to coat the genome completely (Alam, Kanki et al. 2003, Ekstrand, Falkenberg et al. 2004, Kaufman, Durisic et al. 2007). Although one dissenting study does estimate that there may only be an average of 35 TFAM molecules per mtDNA in HeLa cells (Takamatsu, Umeda et al. 2002), far less

than what is required to coat the genome, it is possible that this outlier datapoint may be due to cell growth condition and quantification differences. TFAM has also been shown to bind to nonspecific DNA with affinity approaching that of LSP and HSP1 (Kaufman, Durisic et al. 2007, Wong, Rajagopalan et al. 2009, Malarkey, Bestwick et al. 2012). Finally, DNase I footprinting (Fisher, Lisowsky et al. 1992), *in organello* footprinting (Ghivizzani, Madsen et al. 1994), and chromatin-immunoprecipitation data (Ohgaki, Kanki et al. 2007) have shown that TFAM binds to the mitochondrial genome outside of its promoter binding sites. Therefore, from colocalization studies, its known DNA-binding properties, its abundance in the cell, and direct binding evidence, TFAM most likely binds across the mitochondrial genome.

Although the role of TFAM in transcription and thus replication is well established and known to be independently required for cellular and organismal viability, loss of the nonspecific binding function of TFAM alone is also sufficient to cause mtDNA instability. Fortunately, the two functions of TFAM may be largely isolated thanks to the nature of the mouse form of TFAM, which, although recognizing the corresponding -12 to -39 bp upstream of the mouse mtDNA transcription start site, shares less than 50% homology with the human counterpart at DNA contact sites and is a poor activator of transcription in the human system despite similar binding affinities as assayed by EMSA (Ekstrand, Falkenberg et al. 2004, Gaspari, Larsson et al. 2004). This disassociation between the transcription factor and nonspecific binding properties of TFAM has allowed for some very interesting experiments in mouse models. In a mouse expressing hTFAM as well as endogenous levels of mTFAM, mtDNA copy numbers

increase to levels proportionate to the amount of additional hTFAM expression, despite poor transcriptional activation. Although levels of ND6 (from the LSP transcript) were elevated in the higher-expressing animals, levels of transcripts from HSP were unchanged. Furthermore, respiratory chain function and mitochondrial mass also remain constant upon expression of hTFAM. Interestingly, although hTFAM expression was unable to rescue the embryonic lethality of mTFAM-/- mice, mtDNA levels were at ~30% of WT, indicating that hTFAM alone, despite its highly impaired transcriptional activity in the mouse system, is capable of preserving some levels of mtDNA, most probably due to its nonspecific binding properties. A second mouse model, a heart-specific knockout of mTFAM, experiences rescue of mtDNA levels as well as phenotype when hTFAM is overexpressed despite the transcription defect, further illustrating the ability of TFAM to stabilize mtDNA through its nonspecific binding properties (Freyer, Park et al. 2010). Finally, expression of a TFAM variant lacking the C-terminal tail that is required for transcriptional activation but which is still competent in DNA binding in HeLa cells, where endogenous TFAM levels are knocked down by RNAi, results in prevention of mtDNA depletion (Kanki, Ohgaki et al. 2004), a result replicated in chicken cells as well (Wong, Rajagopalan et al. 2009). The ability of TFAM to increase mtDNA copy number through a non-transcriptionally active form suggests that its function as the mitochondrial histone may increase genomic stability by protecting mtDNA from turnover or deleterious damage.

Although TFAM is required to protect mtDNA, too much TFAM is deleterious. A significant increase in TFAM:mtDNA ratio has been shown to lead to decrease in

mtDNA copy number and transcription (Pohjoismaki, Wanrooij et al. 2006, Matsushima, Goto et al. 2010). Because the increased TFAM levels seem to be associated directly with the genome instead of elsewhere within the cell, and because transcription levels also seem to be perturbed, it appears that this depletion in copy number is due to TFAM interacting with the genome. If TFAM were exclusively a site-specific transcription factor, one would expect that increasing TFAM levels would result in increasing transcripts from LSP, resulting in proportionally increasing mtDNA levels. However, given that this only holds true up to a certain threshold ~2x greater than endogenous levels, this further validates the theory that TFAM usually nonspecifically binds to mtDNA, and at supersaturating levels prevents progression of transcription and thus of replication of the genome through structural over-compaction; indeed, as discussed previously, mtDNA with higher levels of TFAM have been shown to be less active in replication (Wai, Teoli et al. 2008). Therefore, the histone-like properties of TFAM may serve not only to stabilize mtDNA, but also to regulate copy number beyond the transcription factor capabilities. One could certainly envision a quiescent cell line requiring mtDNA copy number stasis utilizing such TFAM-mediated control to limit replication.

The nonspecific, architectural role of TFAM in mtDNA appears to rely on three main properties of the protein: its ability to bend DNA, its ability to bind cooperatively, and its ability to form homodimers. The ability of TFAM to bend DNA was initially suggested by EMSA data showing altered migration of a TFAM-DNA complex through the gel, suggestive of a bend in DNA (Fisher, Lisowsky et al. 1992), and by STEM data

showing that binding of *Xenopus* TFAM at LSP results in a sharp bend in the DNA (Antoshechkin, Bogenhagen et al. 1997). While original estimates placed the bending of DNA at ~100° by AFM (Kaufman, Durisic et al. 2007), the four crystal structures for TFAM solved to date - two while associated with LSP DNA (Ngo, Kaiser et al. 2011, Rubio-Cosials, Sidow et al. 2011), one while associated with HSP1, and the last while associated with nonspecific DNA (Ngo, Lovely et al. 2014) – all suggest a ~180° angle, indicating that the proximate bend is significantly greater than the observed end-to-end bend. FRET assays to further elucidate the bend of DNA have shown that the decrease in end-to-end distance for a 25-mer bound to TFAM is similar for LSP and HSP1 at ~22Å, which is interesting given the slightly lower binding affinity for HSP1, but smaller for nonspecific DNA at ~15Å (Malarkey, Bestwick et al. 2012). The site specificity and the level of bend are dependent on the C-terminal tail, with ablation rendering the decrease in end-to-end distance uniform across all templates at ~10-12Å. Thus, TFAM bends nonspecific DNA, albeit bending at a lower angle, with the C-terminus playing an important role in the degree of bend. As discussed earlier, naked mtDNA requires compaction because is far larger than the diameter of the mitochondrial matrix. Given that this feature of TFAM alone is capable of compacting DNA to a significant extent (Ngo, Lovely et al. 2014), it appears that one of the primary functions of TFAM's bending of nonspecific DNA may be to aid in compaction of the genome into the nucleoid.

A second property of TFAM is its ability to bind DNA with positive cooperativity, with a single binding event nucleating that of subsequent molecules. Such

cooperativity was first qualitatively identified via AFM studies where TFAM was shown

to heterogeneously bind to DNA, preferentially localizing near loci where a previous

binding event had occurred, leading to an estimated Hill coefficient of ~2 (Kaufman,

Durisic et al. 2007). Other studies using fluorescence anisotropy have also shown similar

results, with a Hill coefficient of ~2.5 (Wong, Rajagopalan et al. 2009). Additionally,

TPM experiments using fluorescently labeled TFAM have shown that TFAM slides

along DNA despite its bending properties, is stopped by collision, and forms patches on

DNA in a cooperative manner, finding that TFAM is ~100x more likely to bind next to

an extant molecular than to bare DNA (Farge, Laurens et al. 2012). It has been suggested

that initial binding results in template "melting" and in an increase in flexibility,

facilitating further binding events in the area of affect. Such cooperativity has

implications for the binding of TFAM outside of the promoters. Given the increased

affinity for the LSP site, nucleation from LSP and HSP to the rest of the genome could be

expected to occur.   Indeed, evidence of such "phasing" has already been observed

(Fisher, Lisowsky et al. 1992, Ghivizzani, Madsen et al. 1994).

The third major property of TFAM is its ability to form homodimers. There has

been significant debate over the stoichiometry of TFAM due to conflicting evidence.

From early immunoprecipitation studies, there was some initial evidence that TFAM

could interact with itself, although the exact nature of the interactions was unclear

(Antoshechkin, Bogenhagen et al. 1997). Subsequent experiments using size exclusion

chromatography, surface plasmon resonance (Kaufman, Durisic et al. 2007), and

analytical ultracentrifugation (Wong, Rajagopalan et al. 2009) have all furthered the

theory that TFAM binds as a dimer; there is even isothermal titration calorimetry (Malarkey, Bestwick et al. 2012), and heterogeneous subunit assay (Gangelhoff, Mungalachetty et al. 2009) data suggesting that a TFAM dimer binds at LSP as well as at HSP1 and nonspecific sequences (Malarkey, Bestwick et al. 2012) despite a DNase I footprint insufficient to accommodate two TFAM molecules. Further analytical ultracentrifugation experiments suggest that the C-terminal tail is required for dimerization (Wong, Rajagopalan et al. 2009). However, there is also evidence for TFAM binding as a monomer from size exclusion chromatography with multi-angle light scattering analysis (Ngo, Kaiser et al. 2011) and single molecule fluorescence microscopy (Farge, Laurens et al. 2012). The four recent TFAM crystal structures, while all depict monomeric TFAM binding to DNA, have intriguingly exhibited dimerization at an interface between HMG box A and HMG box B, an interesting coincidence that was further explored by FRET, which shows that TFAM-TFAM interactions do occur in the presence of DNA (Ngo, Lovely et al. 2014), although not at the C-terminus as previous evidence suggested. Most striking, however, is evidence from a TPM assay demonstrating that mutation of the novel homodimer interface leads to decreased compaction as compared to wild-type TFAM, while maintaining wild-type levels of transcriptional activation and binding affinity (Ngo, Lovely et al. 2014). This strongly suggests that this new interface, and not the previously identified C-terminal region, is the region involved in dimerization. Future work will need to be performed to elucidate the properties of this novel interaction, but it appears that while TFAM does not directly bind as a dimer to DNA, the dimerization properties do serve to increase the compaction

of the mitochondrial genome, possibly through a looping mechanism previously observed by AFM (Kaufman, Durisic et al. 2007).

Dynamics of the Mitochondrial Nucleoid

Because mtDNA is required for production of essential components of the ETC, it is predicted that its regulation must be dynamic in order to meet the varying needs of the cell. It is clear that the mitochondrial nucleoid contains a variety of proteins involved in transcription, translation, replication, and regulation of mtDNA, which is essential for cellular and organismal viability. Central to these functions is TFAM, which is the most abundant, and perhaps the best-studied, protein in the mitochondrial nucleoid. TFAM's function in the nucleoid involves not only its site-specific role as a transcription factor probably required for initiation of transcription, but also as a non-specific architectural histone-like protein required for protection of mtDNA stability. Therefore, many models of exactly how TFAM may regulate the main processes of the nucleoid via its biochemical properties have been proposed in order to account for these sometimes contradictory findings. Interestingly, in the ancestral bacterial system, the NAPs are involved in genome organization through not only DNA bending, but also looping and other higher-order structures (Dillon and Dorman 2010, Wang, Li et al. 2011), and are somewhat analogous in role to TFAM's nonspecific histone-like function. Within the NAPs, the HU family of proteins bends prokaryotic DNA for compaction in a strikingly similar way to that of TFAM (Ngo, Kaiser et al. 2011). Interestingly, the expression of

the various NAPs varies between that of the exponential and stationary growth phases (Dame, Espeli et al. 2012), with some NAPs, including FIS, HU, and H-NS, expressed at high levels during exponential growth while others such as Dps and CbpA are expressed during stationary phases of growth (Ali Azam, Iwata et al. 1999). Given the differential binding properties of TFAM at various concentrations, there certainly exists a rationale for a TFAM-based system of regulation within the nucleoid. Indeed, TFAM levels are heterogeneous between nucleoids and correlate with different levels of replication (Wai, Teoli et al. 2008), suggesting some level of regulation by TFAM as well as some control of internucleoid protein distribution within the cell. Furthermore, semi-quantitative analysis of TFAM binding to mtDNA in the frontal cortex, soleus, and liver of aged rats reveals substantial differences in TFAM quantities and binding patterns as compared to young or calorie-restricted rats (Picca, Fracasso et al. 2013, Picca, Pesce et al. 2013, Picca, Pesce et al. 2014), suggesting that TFAM binding and levels can be modulated within cells.

Given a model of mtDNA regulation, how would TFAM itself be regulated to effect control? A prime candidate is the Lon protease, which is known to degrade TFAM (Matsushima, Goto et al. 2010), a property especially important given that TFAM mRNA levels are not always tied to protein levels (King and Attardi 1996, Micol, Fernandez-Silva et al. 1997). Overexpression of Lon protease leads to depletion of TFAM (Matsushima, Goto et al. 2010), while knockdown increases stability, but because Lon protease also degrades other proteins (Bender, Lewrenz et al. 2011), it would be both faster and more advantageous for more specific methods of regulation to also exist. In

fact, similarly to the nuclear histones, TFAM has been shown to be post-translationally modified by glycosylation (Suarez, Hu et al. 2008), phosphorylation (Dephoure, Zhou et al. 2008, Chen, Yang et al. 2009, Lu, Lee et al. 2013), acetylation (Dinardo, Musicco et al. 2003), and ubiquitination (Shi, Chan et al. 2011, Wagner, Beli et al. 2011). Strikingly, when TFAM is phosphorylated by mitochondrial cAMP-dependent protein kinase, its ability to bind DNA and initiate translation are impaired, and it is selectively degraded by Lon protease. Further studies remain to be performed to fully understand the implications of this and other modifications on the properties of TFAM.

Conclusion

In the nearly 50 years since the discovery of the mitochondrial genome, much work has been done to characterize its organization as well as the proteins regulating its replication as well as its transcription and translation products, all essential for proper ETC function. While many proteins have been shown to be components of the DNA-protein nucleoid, key amongst these is TFAM, with many studies demonstrating its functions both as a site-specific transcription factor and as a nonspecific histone-like architectural protein as being essential for genome stability. However, while many models exist which attempt to explain the existing evidence for how TFAM interacts with and regulates the genome, much work remains to be done to validate and further understand current findings. Finally, while how TFAM binds to the genome may be well studied, very little is known about how these binding events are controlled between and

within nucleoids. Understanding the pathways that control the TFAM-mediated regulation of mtDNA stability will serve as a powerful tool, even with potential for therapeutic applications in disorders where mtDNA stability has been observed, such as Alzheimer's, Parkinson's, and aging.

II.   <u>Overview of Thesis</u>

<u>Chapter 2</u>

Mitochondria are dynamic organelles whose membranes undergo the opposing processes of fission and fusion. Several of the biochemical factors involved in dynamics have been explored and studied. The functional implications of dysfunction in dynamics have also been analyzed at both the cellular and whole-organism level. However, it was unclear how perturbation of the balance between fission and fusion cause these observed effects. We demonstrate in mouse embryonic fibroblasts that the outer membrane mitochondrial fusion proteins, Mfn1 and Mfn2, and the inner mitochondrial membrane fusion protein Opa1 are required for maintenance of the mitochondrial genome.

<u>Chapter 3</u>

The transcription factor TFAM is critical for initiation of transcription and replication of the genome, and is also thought to perform a packaging function. Although specific binding sites required for initiation of transcription have been identified in the D-loop, little is known about the characteristics of TFAM binding in its nonspecific packaging state. In addition, it is unclear whether TFAM also plays a role in the regulation of nuclear gene expression. Therefore, to capture a high-resolution profile of TFAM-mtDNA interactions across the entire mitochondrial and nuclear genomes in

various cellular states, we performed chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-seq) for TFAM in human HeLa cells. We directly localize TFAM binding to DNA in human cells, demonstrating that TFAM uniformly coats the whole mitochondrial genome, with no evidence of robust TFAM binding to the nuclear genome. One specific binding site upstream of the O_L localizes to the same area as a previously identified site of TFAM binding to the rat mitochondrial genome. This represents the first direct assessment of TFAM binding on a genome-wide scale in human cells.

Chapter 4

While several classically nuclear transcription factors have been previously reported to localize to mitochondria and to bind to the D-loop to effect transcriptional modulation, there has been no comprehensive analysis of such transcription factors across the mitochondrial genome. Here, we analyze existing ChIP-seq data from ENCODE, mouseENCODE and modENCODE consortia for potential interactions on human, mouse, and *C. elegans* genomes, and identify human and mouse transcription factors with strong localized enrichment outslide the NCR that are usually associated with the corresponding recognition sequence motif. We further confirm these finding by localization of the identified factor MafK to the mitochondria via immunocytochemistry. This represents the

first large-scale, genome-wide characterization of canonically nuclear transcription factors

binding to the mitochondrial genome.

BIBLIOGRAPHY

Achanta, G., R. Sasaki, L. Feng, J. S. Carew, W. Lu, H. Pelicano, M. J. Keating and P. Huang (2005). "Novel role of p53 in maintaining mitochondrial genetic stability through interaction with DNA Pol gamma." EMBO J **24**(19): 3482-3492.

Alam, T. I., T. Kanki, T. Muta, K. Ukaji, Y. Abe, H. Nakayama, K. Takio, N. Hamasaki and D. Kang (2003). "Human mitochondrial DNA is packaged with TFAM." Nucleic Acids Res **31**(6): 1640-1645.

Ali Azam, T., A. Iwata, A. Nishimura, S. Ueda and A. Ishihama (1999). "Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid." J Bacteriol **181**(20): 6361-6370.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. Smith, R. Staden and I. G. Young (1981). "Sequence and organization of the human mitochondrial genome." Nature **290**(5806): 457-465.

Antoshechkin, I., D. F. Bogenhagen and I. A. Mastrangelo (1997). "The HMG-box mitochondrial transcription factor xl-mtTFA binds DNA as a tetramer to activate bidirectional transcription." EMBO J **16**(11): 3198-3206.

Barat, M., D. Rickwood, C. Dufresne and J. C. Mounolou (1985). "Characterization of DNA-protein complexes from the mitochondria of Xenopus laevis oocytes." Exp Cell Res **157**(1): 207-217.

Bender, T., I. Lewrenz, S. Franken, C. Baitzel and W. Voos (2011). "Mitochondrial enzymes are protected from stress-induced aggregation by mitochondrial chaperones and the Pim1/LON protease." Mol Biol Cell **22**(5): 541-554.

Bereiter-Hahn, J. and M. Vöth (1997). "Distribution and dynamics of mitochondrial nucleoids in animal cells in culture." Experimental Biology Online **1**(4): 1-17.

Bogenhagen, D. and D. A. Clayton (1974). "The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid." J Biol Chem **249**(24): 7991-7995.

Bogenhagen, D. and D. A. Clayton (1977). "Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle." Cell **11**(4): 719-727.

Bogenhagen, D. F., D. Rousseau and S. Burke (2008). "The layered structure of human mitochondrial DNA nucleoids." J Biol Chem **283**(6): 3665-3675.

Borst, P. and A. M. Kroon (1969). "Mitochondrial DNA: physicochemical properties, replication, and genetic function." Int Rev Cytol **26**: 107-190.

Brewer, L. R., R. Friddle, A. Noy, E. Baldwin, S. S. Martin, M. Corzett, R. Balhorn and R. J. Baskin (2003). "Packaging of single DNA molecules by the yeast mitochondrial protein Abf2p." Biophys J **85**(4): 2519-2524.

Brown, T. A., C. Cecconi, A. N. Tkachuk, C. Bustamante and D. A. Clayton (2005). "Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism." Genes Dev **19**(20): 2466-2476.

Brown, T. A., A. N. Tkachuk, G. Shtengel, B. G. Kopek, D. F. Bogenhagen, H. F. Hess and D. A. Clayton (2011). "Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction." Mol Cell Biol **31**(24): 4994-5010.

Caelles, C., A. Helmberg and M. Karin (1994). "p53-dependent apoptosis in the absence of transcriptional activation of p53-target genes." Nature **370**(6486): 220-223.

Cammarota, M., G. Paratcha, L. R. Bevilaqua, M. Levi de Stein, M. Lopez, A. Pellegrino de Iraldi, I. Izquierdo and J. H. Medina (1999). "Cyclic AMP-responsive element binding protein in brain mitochondria." J Neurochem **72**(6): 2272-2277.

Cantatore, P. and G. Attardi (1980). "Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA." Nucleic Acids Res **8**(12): 2605-2625.

Casas, F., P. Rochard, A. Rodier, I. Cassar-Malek, S. Marchal-Victorion, R. J. Wiesner, G. Cabello and C. Wrutniak (1999). "A variant form of the nuclear triiodothyronine receptor c-ErbAalpha1 plays a direct role in regulation of mitochondrial RNA synthesis." Mol Cell Biol **19**(12): 7913-7924.

Cavelier, L., A. Johannisson and U. Gyllensten (2000). "Analysis of mtDNA copy number and composition of single mitochondrial particles using flow cytometry and PCR." Exp Cell Res **259**(1): 79-85.

Chang, D. D. and D. A. Clayton (1984). "Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA." Cell **36**(3): 635-643.

Chang, D. D. and D. A. Clayton (1985). "Priming of human mitochondrial DNA replication occurs at the light-strand promoter." Proc Natl Acad Sci U S A **82**(2): 351-355.

Chatre, L. and M. Ricchetti (2013). "Prevalent coordination of mitochondrial DNA transcription and initiation of replication with the cell cycle." Nucleic Acids Res **41**(5): 3068-3078.

Chen, H., A. Chomyn and D. C. Chan (2005). "Disruption of fusion results in mitochondrial heterogeneity and dysfunction." J Biol Chem **280**(28): 26185-26192.

Chen, J. Q., M. Delannoy, C. Cooke and J. D. Yager (2004). "Mitochondrial localization of ERalpha and ERbeta in human MCF7 cells." Am J Physiol Endocrinol Metab **286**(6): E1011-1022.

Chen, R. Q., Q. K. Yang, B. W. Lu, W. Yi, G. Cantin, Y. L. Chen, C. Fearns, J. R. Yates, 3rd and J. D. Lee (2009). "CDC25B mediates rapamycin-induced oncogenic responses in cancer cells." Cancer Res **69**(6): 2663-2668.

Chinnery, P. F. and D. M. Turnbull (1999). "Mitochondrial DNA and disease." Lancet **354 Suppl 1**: SI17-21.

Cingolani, G., L. Capaccio, D. D'Elia and G. Gadaleta (1997). "In organello footprinting analysis of rat mitochondrial DNA: protein interaction upstream of the Ori-L." Biochem Biophys Res Commun **231**(3): 856-860.

Clay Montier, L. L., J. J. Deng and Y. Bai (2009). "Number matters: control of mammalian mitochondrial DNA copy number." J Genet Genomics **36**(3): 125-131.

Clayton, D. A. (1982). "Replication of animal mitochondrial DNA." Cell **28**(4): 693-705.

Corral-Debrinski, M., J. M. Shoffner, M. T. Lott and D. C. Wallace (1992). "Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease." Mutat Res **275**(3-6): 169-180.

Coskun, P., J. Wyrembak, S. E. Schriner, H. W. Chen, C. Marciniack, F. Laferla and D. C. Wallace (2012). "A mitochondrial etiology of Alzheimer and Parkinson disease." Biochim Biophys Acta **1820**(5): 553-564.

Coskun, P. E., M. F. Beal and D. C. Wallace (2004). "Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication." Proc Natl Acad Sci U S A **101**(29): 10726-10731.

Cotney, J., S. E. McKay and G. S. Shadel (2009). "Elucidation of separate, but collaborative functions of the rRNA methyltransferase-related human mitochondrial

transcription factors B1 and B2 in mitochondrial biogenesis reveals new insight into maternally inherited deafness." Hum Mol Genet **18**(14): 2670-2682.

Cotney, J. and G. S. Shadel (2006). "Evidence for an early gene duplication event in the evolution of the mitochondrial transcription factor B family and maintenance of rRNA methyltransferase activity in human mtTFB1 and mtTFB2." J Mol Evol **63**(5): 707-717.

Cotney, J., Z. Wang and G. S. Shadel (2007). "Relative abundance of the human mitochondrial transcription system and distinct roles for h-mtTFB1 and h-mtTFB2 in mitochondrial biogenesis and gene expression." Nucleic Acids Res **35**(12): 4042-4054.

Crews, S., D. Ojala, J. Posakony, J. Nishiguchi and G. Attardi (1979). "Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication." Nature **277**(5693): 192-198.

Dairaghi, D. J., G. S. Shadel and D. A. Clayton (1995). "Human mitochondrial transcription factor A and promoter spacing integrity are required for transcription initiation." Biochim Biophys Acta **1271**(1): 127-134.

Dame, R. T., O. Espeli, D. C. Grainger and P. A. Wiggins (2012). "Multidisciplinary perspectives on bacterial genome organization and dynamics." Mol Microbiol **86**(5): 1023-1030.

Davis, A. F. and D. A. Clayton (1996). "In situ localization of mitochondrial DNA replication in intact mammalian cells." J Cell Biol **135**(4): 883-893.

De Rasmo, D., A. Signorile, E. Roca and S. Papa (2009). "cAMP response element-binding protein (CREB) is imported into mitochondria and promotes protein synthesis." FEBS J 276(16): 4325-4333.

Delius, H. and A. Worcel (1974). "Electron microscopic studies on the folded chromosome of Escherichia coli." Cold Spring Harb Symp Quant Biol 38: 53-58.

Demonacos, C., N. C. Tsawdaroglou, R. Djordjevic-Markovic, M. Papalopoulou, V. Galanopoulos, S. Papadogeorgaki and C. E. Sekeris (1993). "Import of the glucocorticoid receptor into rat liver mitochondria in vivo and in vitro." J Steroid Biochem Mol Biol 46(3): 401-413.

Dephoure, N., C. Zhou, J. Villen, S. A. Beausoleil, C. E. Bakalarski, S. J. Elledge and S. P. Gygi (2008). "A quantitative atlas of mitotic phosphorylation." Proc Natl Acad Sci U S A 105(31): 10762-10767.

Diffley, J. F. and B. Stillman (1991). "A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria." Proc Natl Acad Sci U S A 88(17): 7864-7868.

Diffley, J. F. and B. Stillman (1992). "DNA binding properties of an HMG1-related protein from yeast mitochondria." J Biol Chem 267(5): 3368-3374.

Dillon, S. C. and C. J. Dorman (2010). "Bacterial nucleoid-associated proteins, nucleoid structure and gene expression." Nat Rev Microbiol 8(3): 185-195.

DiMauro, S. and E. A. Schon (2003). "Mitochondrial respiratory-chain diseases." <u>N Engl J Med</u> **348**(26): 2656-2668.

Dinardo, M. M., C. Musicco, F. Fracasso, F. Milella, M. N. Gadaleta, G. Gadaleta and P. Cantatore (2003). "Acetylation and level of mitochondrial transcription factor A in several organs of young and old rats." <u>Biochem Biophys Res Commun</u> **301**(1): 187-191.

Ekstrand, M. I., M. Falkenberg, A. Rantanen, C. B. Park, M. Gaspari, K. Hultenby, P. Rustin, C. M. Gustafsson and N. G. Larsson (2004). "Mitochondrial transcription factor A regulates mtDNA copy number in mammals." <u>Hum Mol Genet</u> **13**(9): 935-944.

Enriquez, J. A., P. Fernandez-Silva, N. Garrido-Perez, M. J. Lopez-Perez, A. Perez-Martos and J. Montoya (1999). "Direct regulation of mitochondrial RNA synthesis by thyroid hormone." <u>Mol Cell Biol</u> **19**(1): 657-670.

Falkenberg, M., M. Gaspari, A. Rantanen, A. Trifunovic, N. G. Larsson and C. M. Gustafsson (2002). "Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA." <u>Nat Genet</u> **31**(3): 289-294.

Falkenberg, M. and N. G. Larsson (2009). "Structure casts light on mtDNA replication." <u>Cell</u> **139**(2): 231-233.

Falkenberg, M., N. G. Larsson and C. M. Gustafsson (2007). "DNA replication and transcription in mammalian mitochondria." <u>Annu Rev Biochem</u> **76**: 679-699.

Farge, G., N. Laurens, O. D. Broekmans, S. M. van den Wildenberg, L. C. Dekker, M. Gaspari, C. M. Gustafsson, E. J. Peterman, M. Falkenberg and G. J. Wuite (2012). "Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A." Nat Commun **3**: 1013.

Fernandez-Vizarra, E., J. A. Enriquez, A. Perez-Martos, J. Montoya and P. Fernandez-Silva (2008). "Mitochondrial gene expression is regulated at multiple levels and differentially in the heart and liver by thyroid hormones." Curr Genet **54**(1): 13-22.

Fisher, R. P. and D. A. Clayton (1985). "A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro." J Biol Chem **260**(20): 11330-11338.

Fisher, R. P. and D. A. Clayton (1988). "Purification and characterization of human mitochondrial transcription factor 1." Mol Cell Biol **8**(8): 3496-3509.

Fisher, R. P., T. Lisowsky, M. A. Parisi and D. A. Clayton (1992). "DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein." J Biol Chem **267**(5): 3358-3367.

Fisher, R. P., M. A. Parisi and D. A. Clayton (1989). "Flexible recognition of rapidly evolving promoter sequences by mitochondrial transcription factor 1." Genes Dev **3**(12B): 2202-2217.

Freyer, C., C. B. Park, M. I. Ekstrand, Y. Shi, J. Khvorostova, R. Wibom, M. Falkenberg, C. M. Gustafsson and N. G. Larsson (2010). "Maintenance of respiratory chain function in mouse hearts with severely impaired mtDNA transcription." Nucleic Acids Res **38**(19): 6577-6588.

Fu, G. K. and D. M. Markovitz (1998). "The human LON protease binds to mitochondrial promoters in a single-stranded, site-specific, strand-specific manner." Biochemistry **37**(7): 1905-1909.

Gadaleta, G., D. D'Elia, L. Capaccio, C. Saccone and G. Pepe (1996). "Isolation of a 25-kDa protein binding to a curved DNA upstream the origin of the L strand replication in the rat mitochondrial genome." J Biol Chem **271**(23): 13537-13541.

Gangelhoff, T. A., P. S. Mungalachetty, J. C. Nix and M. E. Churchill (2009). "Structural analysis and DNA binding of the HMG domains of the human mitochondrial transcription factor A." Nucleic Acids Res **37**(10): 3153-3164.

Garrido, N., L. Griparic, E. Jokitalo, J. Wartiovaara, A. M. van der Bliek and J. N. Spelbrink (2003). "Composition and dynamics of human mitochondrial nucleoids." Mol Biol Cell **14**(4): 1583-1596.

Garstka, H. L., W. E. Schmitt, J. Schultz, B. Sogl, B. Silakowski, A. Perez-Martos, J. Montoya and R. J. Wiesner (2003). "Import of mitochondrial transcription factor A (TFAM) into rat liver mitochondria stimulates transcription of mitochondrial DNA." Nucleic Acids Res **31**(17): 5039-5047.

Gaspari, M., N. G. Larsson and C. M. Gustafsson (2004). "The transcription machinery in mammalian mitochondria." Biochim Biophys Acta **1659**(2-3): 148-152.

Gauthier, B. R., A. Wiederkehr, M. Baquie, C. Dai, A. C. Powers, J. Kerr-Conte, F. Pattou, R. J. MacDonald, J. Ferrer and C. B. Wollheim (2009). "PDX1 deficiency causes mitochondrial dysfunction and defective insulin secretion through TFAM suppression." Cell Metab **10**(2): 110-118.

Gensler, S., K. Weber, W. E. Schmitt, A. Perez-Martos, J. A. Enriquez, J. Montoya and R. J. Wiesner (2001). "Mechanism of mammalian mitochondrial DNA replication: import of mitochondrial transcription factor A into isolated mitochondria stimulates 7S DNA synthesis." Nucleic Acids Res **29**(17): 3657-3663.

Ghivizzani, S. C., C. S. Madsen, M. R. Nelen, C. V. Ammini and W. W. Hauswirth (1994). "In organello footprint analysis of human mitochondrial DNA: human mitochondrial transcription factor A interactions at the origin of replication." Mol Cell Biol **14**(12): 7717-7730.

Gilkerson, R. W., E. A. Schon, E. Hernandez and M. M. Davidson (2008). "Mitochondrial nucleoids maintain genetic autonomy but allow for functional complementation." J Cell Biol **181**(7): 1117-1128.

Goto, Y., I. Nonaka and S. Horai (1990). "A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies." Nature **348**(6302): 651-653.

Hayashi, Y., M. Yoshida, M. Yamato, T. Ide, Z. Wu, M. Ochi-Shindou, T. Kanki, D. Kang, K. Sunagawa, H. Tsutsui and H. Nakanishi (2008). "Reverse of age-dependent memory impairment and mitochondrial DNA damage in microglia by an overexpression of human mitochondrial transcription factor a in mice." J Neurosci **28**(34): 8624-8634.

He, J., C. C. Mao, A. Reyes, H. Sembongi, M. Di Re, C. Granycome, A. B. Clippingdale, I. M. Fearnley, M. Harbour, A. J. Robinson, S. Reichelt, J. N. Spelbrink, J. E. Walker and I. J. Holt (2007). "The AAA+ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization." J Cell Biol **176**(2): 141-146.

Heyne, K., S. Mannebach, E. Wuertz, K. X. Knaup, M. Mahyar-Roemer and K. Roemer (2004). "Identification of a putative p53 binding sequence within the human mitochondrial genome." FEBS Lett **578**(1-2): 198-202.

Holt, I. J., J. He, C. C. Mao, J. D. Boyd-Kirkup, P. Martinsson, H. Sembongi, A. Reyes and J. N. Spelbrink (2007). "Mammalian mitochondrial nucleoids: organizing an independently minded genome." Mitochondrion **7**(5): 311-321.

Holt, I. J., H. E. Lorimer and H. T. Jacobs (2000). "Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA." Cell **100**(5): 515-524.

Iborra, F. J., H. Kimura and P. R. Cook (2004). "The functional organization of mitochondrial genomes in human cells." BMC Biol **2**: 9.

Ide, T., H. Tsutsui, S. Hayashidani, D. Kang, N. Suematsu, K. Nakamura, H. Utsumi, N. Hamasaki and A. Takeshita (2001). "Mitochondrial DNA damage and dysfunction

associated with oxidative stress in failing hearts after myocardial infarction." Circ Res **88**(5): 529-535.

Ikeuchi, M., H. Matsusaka, D. Kang, S. Matsushima, T. Ide, T. Kubota, T. Fujiwara, N. Hamasaki, A. Takeshita, K. Sunagawa and H. Tsutsui (2005). "Overexpression of mitochondrial transcription factor a ameliorates mitochondrial deficiencies and cardiac failure after myocardial infarction." Circulation **112**(5): 683-690.

Jazbutyte, V., F. Kehl, L. Neyses and T. Pelzer (2009). "Estrogen receptor alpha interacts with 17beta-hydroxysteroid dehydrogenase type 10 in mitochondria." Biochem Biophys Res Commun **384**(4): 450-454.

Kanki, T., K. Ohgaki, M. Gaspari, C. M. Gustafsson, A. Fukuoh, N. Sasaki, N. Hamasaki and D. Kang (2004). "Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA." Mol Cell Biol **24**(22): 9823-9834.

Karamanlidis, G., V. Bautista-Hernandez, F. Fynn-Thompson, P. Del Nido and R. Tian (2011). "Impaired mitochondrial biogenesis precedes heart failure in right ventricular hypertrophy in congenital heart disease." Circ Heart Fail **4**(6): 707-713.

Karamanlidis, G., L. Nascimben, G. S. Couper, P. S. Shekar, F. del Monte and R. Tian (2010). "Defective DNA replication impairs mitochondrial biogenesis in human failing hearts." Circ Res **106**(9): 1541-1548.

Kasamatsu, H. and J. Vinograd (1973). "Unidirectionality of replication in mouse mitochondrial DNA." Nat New Biol **241**(108): 103-105.

Kasashima, K., M. Sumitani, M. Satoh and H. Endo (2008). "Human prohibitin 1 maintains the organization and stability of the mitochondrial nucleoids." Exp Cell Res **314**(5): 988-996.

Kaufman, B. A., N. Durisic, J. M. Mativetsky, S. Costantino, M. A. Hancock, P. Grutter and E. A. Shoubridge (2007). "The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures." Mol Biol Cell **18**(9): 3225-3236.

Kaufman, B. A., J. E. Kolesar, P. S. Perlman and R. A. Butow (2003). "A function for the mitochondrial chaperonin Hsp60 in the structure and transmission of mitochondrial DNA nucleoids in Saccharomyces cerevisiae." J Cell Biol **163**(3): 457-461.

Keeney, P. M., C. K. Quigley, L. D. Dunham, C. M. Papageorge, S. Iyer, R. R. Thomas, K. M. Schwarz, P. A. Trimmer, S. M. Khan, F. R. Portell, K. E. Bergquist and J. P. Bennett, Jr. (2009). "Mitochondrial gene therapy augments mitochondrial physiology in a Parkinson's disease cell model." Hum Gene Ther **20**(8): 897-907.

King, M. P. and G. Attardi (1996). "Isolation of human cell lines lacking mitochondrial DNA." Methods Enzymol **264**: 304-313.

Kucej, M., B. Kucejova, R. Subramanian, X. J. Chen and R. A. Butow (2008). "Mitochondrial nucleoids undergo remodeling in response to metabolic cues." J Cell Sci **121**(Pt 11): 1861-1868.

Kukat, C., C. A. Wurm, H. Spahr, M. Falkenberg, N. G. Larsson and S. Jakobs (2011). "Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA." Proc Natl Acad Sci U S A **108**(33): 13534-13539.

Larsson, N. G., J. Wang, H. Wilhelmsson, A. Oldfors, P. Rustin, M. Lewandoski, G. S. Barsh and D. A. Clayton (1998). "Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice." Nat Genet **18**(3): 231-236.

Lee, J., C. H. Kim, D. K. Simon, L. R. Aminova, A. Y. Andreyev, Y. E. Kushnareva, A. N. Murphy, B. E. Lonze, K. S. Kim, D. D. Ginty, R. J. Ferrante, H. Ryu and R. R. Ratan (2005). "Mitochondrial cyclic AMP response element-binding protein (CREB) mediates mitochondrial gene expression and neuronal survival." J Biol Chem **280**(49): 40398-40401.

Lee, S., S. Kim, X. Sun, J. H. Lee and H. Cho (2007). "Cell cycle-dependent mitochondrial biogenesis and dynamics in mammalian cells." Biochem Biophys Res Commun **357**(1): 111-117.

Lee, Y. S., W. D. Kennedy and Y. W. Yin (2009). "Structural insight into processive human mitochondrial DNA synthesis and disease-related polymerase mutations." Cell **139**(2): 312-324.

Legros, F., F. Malka, P. Frachon, A. Lombes and M. Rojo (2004). "Organization and dynamics of human mitochondrial DNA." J Cell Sci **117**(Pt 13): 2653-2662.

Leonard, J. V. and A. H. Schapira (2000). "Mitochondrial respiratory chain disorders I: mitochondrial DNA defects." Lancet **355**(9200): 299-304.

Litonin, D., M. Sologub, Y. Shi, M. Savkina, M. Anikin, M. Falkenberg, C. M. Gustafsson and D. Temiakov (2010). "Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro." J Biol Chem **285**(24): 18129-18133.

Lodeiro, M. F., A. Uchida, M. Bestwick, I. M. Moustafa, J. J. Arnold, G. S. Shadel and C. E. Cameron (2012). "Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factor A in vitro." Proc Natl Acad Sci U S A **109**(17): 6513-6518.

Lombes, A., K. Aure, C. Bellanne-Chantelot, M. Gilleron and C. Jardel (2014). "Unsolved issues related to human mitochondrial diseases." Biochimie **100C**: 171-176.

Lu, B., J. Lee, X. Nie, M. Li, Y. I. Morozov, S. Venkatesh, D. F. Bogenhagen, D. Temiakov and C. K. Suzuki (2013). "Phosphorylation of human TFAM in mitochondria impairs DNA binding and promotes degradation by the AAA+ Lon protease." Mol Cell **49**(1): 121-132.

Luck, D. J. and E. Reich (1964). "DNA in Mitochondria of Neurospora Crassa." Proc Natl Acad Sci U S A **52**: 931-938.

Maassen, J. A., T. H. LM, E. Van Essen, R. J. Heine, G. Nijpels, R. S. Jahangir Tafrechi, A. K. Raap, G. M. Janssen and H. H. Lemkes (2004). "Mitochondrial diabetes: molecular mechanisms and clinical presentation." Diabetes **53 Suppl 1**: S103-109.

Magnusson, J., M. Orth, P. Lestienne and J. W. Taanman (2003). "Replication of mitochondrial DNA occurs throughout the mitochondria of cultured human cells." Exp Cell Res **289**(1): 133-142.

Malarkey, C. S., M. Bestwick, J. E. Kuhlwilm, G. S. Shadel and M. E. Churchill (2012). "Transcriptional activation by mitochondrial transcription factor A involves preferential distortion of promoter DNA." Nucleic Acids Res **40**(2): 614-624.

Margineantu, D. H., W. Gregory Cox, L. Sundell, S. W. Sherwood, J. M. Beechem and R. A. Capaldi (2002). "Cell cycle dependent morphology changes and associated mitochondrial DNA redistribution in mitochondria of human cell lines." Mitochondrion **1**(5): 425-435.

Martin, M., J. Cho, A. J. Cesare, J. D. Griffith and G. Attardi (2005). "Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis." Cell **123**(7): 1227-1240.

Matsushima, Y., C. Adan, R. Garesse and L. S. Kaguni (2005). "Drosophila mitochondrial transcription factor B1 modulates mitochondrial translation but not transcription or DNA copy number in Schneider cells." J Biol Chem **280**(17): 16815-16820.

Matsushima, Y., R. Garesse and L. S. Kaguni (2004). "Drosophila mitochondrial transcription factor B2 regulates mitochondrial DNA copy number and transcription in schneider cells." J Biol Chem **279**(26): 26900-26905.

Matsushima, Y., Y. Goto and L. S. Kaguni (2010). "Mitochondrial Lon protease regulates mitochondrial DNA copy number and transcription by selective degradation of mitochondrial transcription factor A (TFAM)." Proc Natl Acad Sci U S A **107**(43): 18410-18415.

McCulloch, V. and G. S. Shadel (2003). "Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity." Mol Cell Biol **23**(16): 5816-5824.

Metodiev, M. D., N. Lesko, C. B. Park, Y. Camara, Y. Shi, R. Wibom, K. Hultenby, C. M. Gustafsson and N. G. Larsson (2009). "Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome." Cell Metab **9**(4): 386-397.

Metzler, R. and T. Ambjornsson (2005). "Dynamic approach to DNA breathing." J Biol Phys **31**(3-4): 339-350.

Micol, V., P. Fernandez-Silva and G. Attardi (1997). "Functional analysis of in vivo and in organello footprinting of HeLa cell mitochondrial DNA in relationship to ATP and ethidium bromide effects on transcription." J Biol Chem **272**(30): 18896-18904.

Milanesi, L., A. Vasconsuelo, A. R. de Boland and R. Boland (2009). "Expression and subcellular distribution of native estrogen receptor beta in murine C2C12 cells and skeletal muscle tissue." Steroids **74**(6): 489-497.

Milenkovic, D., S. Matic, I. Kuhl, B. Ruzzenente, C. Freyer, E. Jemt, C. B. Park, M. Falkenberg and N. G. Larsson (2013). "TWINKLE is an essential mitochondrial helicase required for synthesis of nascent D-loop strands and complete mtDNA replication." Hum Mol Genet **22**(10): 1983-1993.

Monje, P. and R. Boland (2001). "Subcellular distribution of native estrogen receptor alpha and beta isoforms in rabbit uterus and ovary." J Cell Biochem **82**(3): 467-479.

Montoya, J., T. Christianson, D. Levens, M. Rabinowitz and G. Attardi (1982). "Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA." Proc Natl Acad Sci U S A **79**(23): 7195-7199.

Montoya, J., G. L. Gaines and G. Attardi (1983). "The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units." Cell **34**(1): 151-159.

Moslemi, A. R., M. Tulinius, E. Holme and A. Oldfors (1998). "Threshold expression of the tRNA(Lys) A8344G mutation in single muscle fibres." Neuromuscul Disord **8**(5): 345-349.

Mouw, K. W. and P. A. Rice (2007). "Shaping the Borrelia burgdorferi genome: crystal structure and binding properties of the DNA-bending protein Hbb." Mol Microbiol **63**(5): 1319-1330.

Moyes, C. D., B. J. Battersby and S. C. Leary (1998). "Regulation of muscle mitochondrial design." J Exp Biol **201**(Pt 3): 299-307.

Nass, M. M. (1969). "Mitochondrial DNA: Advances, Problems, and Goals." Science **165**(3888): 25-35.

Nass, S., M. M. Nass and U. Hennix (1965). "Deoxyribonucleic Acid in Isolated Rat-Liver Mitochondria." Biochim Biophys Acta **95**: 426-435.

Ngo, H. B., J. T. Kaiser and D. C. Chan (2011). "The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA." Nat Struct Mol Biol **18**(11): 1290-1296.

Ngo, H. B., G. A. Lovely, R. Phillips and D. C. Chan (2014). "Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation." Nat Commun **5**: 3077.

Ohgaki, K., T. Kanki, A. Fukuoh, H. Kurisaki, Y. Aoki, M. Ikeuchi, S. H. Kim, N. Hamasaki and D. Kang (2007). "The C-terminal tail of mitochondrial transcription factor a markedly strengthens its general binding to DNA." J Biochem **141**(2): 201-211.

Parisi, M. A. and D. A. Clayton (1991). "Similarity of human mitochondrial transcription factor 1 to high mobility group proteins." Science **252**(5008): 965-969.

Park, C. B., J. Asin-Cayuela, Y. Camara, Y. Shi, M. Pellegrini, M. Gaspari, R. Wibom, K. Hultenby, H. Erdjument-Bromage, P. Tempst, M. Falkenberg, C. M. Gustafsson and N. G. Larsson (2007). "MTERF3 is a negative regulator of mammalian mtDNA transcription." Cell **130**(2): 273-285.

Parvin, J. D. and P. A. Sharp (1993). "DNA topology and a minimal set of basal factors for transcription by RNA polymerase II." Cell **73**(3): 533-540.

Pellegrini, M., J. Asin-Cayuela, H. Erdjument-Bromage, P. Tempst, N. G. Larsson and C. M. Gustafsson (2009). "MTERF2 is a nucleoid component in mammalian mitochondria." Biochim Biophys Acta **1787**(5): 296-302.

Piao, Y., H. G. Kim, M. S. Oh and Y. K. Pak (2012). "Overexpression of TFAM, NRF-1 and myr-AKT protects the MPP(+)-induced mitochondrial dysfunctions in neuronal cells." Biochim Biophys Acta **1820**(5): 577-585.

Picca, A., F. Fracasso, V. Pesce, P. Cantatore, A. M. Joseph, C. Leeuwenburgh, M. N. Gadaleta and A. M. Lezza (2013). "Age- and calorie restriction-related changes in rat brain mitochondrial DNA and TFAM binding." Age (Dordr) **35**(5): 1607-1620.

Picca, A., V. Pesce, F. Fracasso, A. M. Joseph, C. Leeuwenburgh and A. M. Lezza (2013). "Aging and calorie restriction oppositely affect mitochondrial biogenesis through TFAM

binding at both origins of mitochondrial DNA replication in rat liver." PLoS One **8**(9): e74644.

Picca, A., V. Pesce, F. Fracasso, A. M. Joseph, C. Leeuwenburgh and A. M. Lezza (2014). "A comparison among the tissue-specific effects of aging and calorie restriction on TFAM amount and TFAM-binding activity to mtDNA in rat." Biochim Biophys Acta.

Pierro, P., L. Capaccio and G. Gadaleta (1999). "The 25 kDa protein recognizing the rat curved region upstream of the origin of the L-strand replication is the rat homologue of the human mitochondrial transcription factor A." FEBS Lett **457**(3): 307-310.

Pohjoismaki, J. L., S. Wanrooij, A. K. Hyvarinen, S. Goffart, I. J. Holt, J. N. Spelbrink and H. T. Jacobs (2006). "Alterations to the expression level of mitochondrial transcription factor A, TFAM, modify the mode of mitochondrial DNA replication in cultured human cells." Nucleic Acids Res **34**(20): 5815-5828.

Reyes, A., J. He, C. C. Mao, L. J. Bailey, M. Di Re, H. Sembongi, L. Kazak, K. Dzionek, J. B. Holmes, T. J. Cluett, M. E. Harbour, I. M. Fearnley, R. J. Crouch, M. A. Conti, R. S. Adelstein, J. E. Walker and I. J. Holt (2011). "Actin and myosin contribute to mammalian mitochondrial DNA maintenance." Nucleic Acids Res **39**(12): 5098-5108.

Rice, A. C., P. M. Keeney, N. K. Algarzae, A. C. Ladd, R. R. Thomas and J. P. Bennett Jr (2014). "Mitochondrial DNA Copy Numbers in Pyramidal Neurons are Decreased and Mitochondrial Biogenesis Transcriptome Signaling is Disrupted in Alzheimer's Disease Hippocampi." J Alzheimers Dis.

Rice, P. A., S. Yang, K. Mizuuchi and H. A. Nash (1996). "Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn." Cell **87**(7): 1295-1306.

Rorbach, J., R. Richter, H. J. Wessels, M. Wydro, M. Pekalski, M. Farhoud, I. Kuhl, M. Gaisne, N. Bonnefoy, J. A. Smeitink, R. N. Lightowlers and Z. M. Chrzanowska-Lightowlers (2008). "The human mitochondrial ribosome recycling factor is essential for cell viability." Nucleic Acids Res **36**(18): 5787-5799.

Rossignol, R., B. Faustin, C. Rocher, M. Malgat, J. P. Mazat and T. Letellier (2003). "Mitochondrial threshold effects." Biochem J **370**(Pt 3): 751-762.

Rubio-Cosials, A., J. F. Sidow, N. Jimenez-Menendez, P. Fernandez-Millan, J. Montoya, H. T. Jacobs, M. Coll, P. Bernado and M. Sola (2011). "Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter." Nat Struct Mol Biol **18**(11): 1281-1289.

Sagan, L. (1967). "On the origin of mitosing cells." J Theor Biol **14**(3): 255-274.

Satoh, M. and T. Kuroiwa (1991). "Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell." Exp Cell Res **196**(1): 137-140.

Schatz, G., E. Haslbrunner and H. Tuppy (1964). "Deoxyribonucleic acid associated with yeast mitochondria." Biochemical and Biophysical Research Communications **15**(2): 127-132.

Seidel-Rogol, B. L., V. McCulloch and G. S. Shadel (2003). "Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop." Nat Genet **33**(1): 23-24.

Shi, Y., D. W. Chan, S. Y. Jung, A. Malovannaya, Y. Wang and J. Qin (2011). "A data set of human endogenous protein ubiquitination sites." Mol Cell Proteomics **10**(5): M110 002089.

Shi, Y., A. Dierckx, P. H. Wanrooij, S. Wanrooij, N. G. Larsson, L. M. Wilhelmsson, M. Falkenberg and C. M. Gustafsson (2012). "Mammalian transcription factor A is a core component of the mitochondrial transcription machinery." Proc Natl Acad Sci U S A **109**(41): 16510-16515.

Shmookler Reis, R. J. and S. Goldstein (1983). "Mitochondrial DNA in mortal and immortal human cells. Genome number, integrity, and methylation." J Biol Chem **258**(15): 9078-9085.

Shoffner, J. M., M. T. Lott, A. M. Lezza, P. Seibel, S. W. Ballinger and D. C. Wallace (1990). "Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation." Cell **61**(6): 931-937.

Shutt, T. E., M. Bestwick and G. S. Shadel (2011). "The core human mitochondrial transcription initiation complex: It only takes two to tango." Transcription **2**(2): 55-59.

Shutt, T. E., M. F. Lodeiro, J. Cotney, C. E. Cameron and G. S. Shadel (2010). "Core human mitochondrial transcription apparatus is a regulated two-component system in vitro." Proc Natl Acad Sci U S A **107**(27): 12133-12138.

Sickmann, A., J. Reinders, Y. Wagner, C. Joppich, R. Zahedi, H. E. Meyer, B. Schonfisch, I. Perschil, A. Chacinska, B. Guiard, P. Rehling, N. Pfanner and C. Meisinger (2003). "The proteome of Saccharomyces cerevisiae mitochondria." Proc Natl Acad Sci U S A **100**(23): 13207-13212.

Simmons, R. A., I. Suponitsky-Kroyter and M. A. Selak (2005). "Progressive accumulation of mitochondrial DNA mutations and decline in mitochondrial function lead to beta-cell failure." J Biol Chem **280**(31): 28785-28791.

Sondheimer, N., J. K. Fang, E. Polyak, M. J. Falk and N. G. Avadhani (2010). "Leucine-rich pentatricopeptide-repeat containing protein regulates mitochondrial transcription." Biochemistry **49**(35): 7467-7473.

Spelbrink, J. N., F. Y. Li, V. Tiranti, K. Nikali, Q. P. Yuan, M. Tariq, S. Wanrooij, N. Garrido, G. Comi, L. Morandi, L. Santoro, A. Toscano, G. M. Fabrizi, H. Somer, R. Croxen, D. Beeson, J. Poulton, A. Suomalainen, H. T. Jacobs, M. Zeviani and C. Larsson (2001). "Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria." Nat Genet **28**(3): 223-231.

Stumpf, J. D., R. P. Saneto and W. C. Copeland (2013). "Clinical and molecular features of POLG-related mitochondrial disease." Cold Spring Harb Perspect Biol **5**(4): a011395.

Suarez, J., Y. Hu, A. Makino, E. Fricovsky, H. Wang and W. H. Dillmann (2008). "Alterations in mitochondrial function and cytosolic calcium induced by hyperglycemia are restored by mitochondrial transcription factor A in cardiomyocytes." Am J Physiol Cell Physiol **295**(6): C1561-1568.

Sugiyama, S., K. Hattori, M. Hayakawa and T. Ozawa (1991). "Quantitative analysis of age-associated accumulation of mitochondrial DNA with deletion in human hearts." Biochem Biophys Res Commun **180**(2): 894-899.

Suomalainen, A. and P. Isohanni (2010). "Mitochondrial DNA depletion syndromes--many genes, common mechanisms." Neuromuscul Disord **20**(7): 429-437.

Szczesny, R. J., L. S. Borowski, L. K. Brzezniak, A. Dmochowska, K. Gewartowski, E. Bartnik and P. P. Stepien (2010). "Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance." Nucleic Acids Res **38**(1): 279-298.

Szklarczyk, R. and M. A. Huynen (2010). "Mosaic origin of the mitochondrial proteome." Proteomics **10**(22): 4012-4024.

Takamatsu, C., S. Umeda, T. Ohsato, T. Ohno, Y. Abe, A. Fukuoh, H. Shinagawa, N. Hamasaki and D. Kang (2002). "Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein." EMBO Rep **3**(5): 451-456.

Tang, Y., E. A. Schon, E. Wilichowski, M. E. Vazquez-Memije, E. Davidson and M. P. King (2000). "Rearrangements of human mitochondrial DNA (mtDNA): new insights into the regulation of mtDNA copy number and gene expression." Mol Biol Cell **11**(4): 1471-1485.

Trifunovic, A. and N. G. Larsson (2008). "Mitochondrial dysfunction as a cause of ageing." J Intern Med **263**(2): 167-178.

van Bruggen, E. F., P. Borst, G. J. Ruttenberg, M. Gruber and A. M. Kroon (1966). "Circular mitochondrial DNA." Biochim Biophys Acta **119**(2): 437-439.

Wagner, S. A., P. Beli, B. T. Weinert, M. L. Nielsen, J. Cox, M. Mann and C. Choudhary (2011). "A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles." Mol Cell Proteomics **10**(10): M111 013284.

Wai, T., A. Ao, X. Zhang, D. Cyr, D. Dufort and E. A. Shoubridge (2010). "The role of mitochondrial DNA copy number in mammalian fertility." Biol Reprod **83**(1): 52-62.

Wai, T., D. Teoli and E. A. Shoubridge (2008). "The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes." Nat Genet **40**(12): 1484-1488.

Wang, D. D., Z. Shu, S. A. Lieser, P. L. Chen and W. H. Lee (2009). "Human mitochondrial SUV3 and polynucleotide phosphorylase form a 330-kDa heteropentamer to cooperatively degrade double-stranded RNA with a 3'-to-5' directionality." J Biol Chem **284**(31): 20812-20821.

Wang, W., G. W. Li, C. Chen, X. S. Xie and X. Zhuang (2011). "Chromosome organization by a nucleoid-associated protein in live bacteria." <u>Science</u> **333**(6048): 1445-1449.

Wang, Y. and D. F. Bogenhagen (2006). "Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane." <u>J Biol Chem</u> **281**(35): 25791-25802.

Wegrzyn, J., R. Potla, Y. J. Chwae, N. B. Sepuri, Q. Zhang, T. Koeck, M. Derecka, K. Szczepanek, M. Szelag, A. Gornicka, A. Moh, S. Moghaddas, Q. Chen, S. Bobbili, J. Cichy, J. Dulak, D. P. Baker, A. Wolfman, D. Stuehr, M. O. Hassan, X. Y. Fu, N. Avadhani, J. I. Drake, P. Fawcett, E. J. Lesnefsky and A. C. Larner (2009). "Function of mitochondrial Stat3 in cellular respiration." <u>Science</u> **323**(5915): 793-797.

Williams, R. S. (1986). "Mitochondrial gene expression in mammalian striated muscle. Evidence that variation in gene dosage is the major regulatory event." <u>J Biol Chem</u> **261**(26): 12390-12394.

Wong, T. S., S. Rajagopalan, S. M. Freund, T. J. Rutherford, A. Andreeva, F. M. Townsley, M. Petrovich and A. R. Fersht (2009). "Biophysical characterizations of human mitochondrial transcription factor A and its binding to tumor suppressor p53." <u>Nucleic Acids Res</u> **37**(20): 6765-6783.

Wrutniak, C., I. Cassar-Malek, S. Marchal, A. Rascle, S. Heusser, J. M. Keller, J. Flechon, M. Dauca, J. Samarut, J. Ghysdael and et al. (1995). "A 43-kDa protein related to c-Erb A

alpha 1 is located in the mitochondrial matrix of rat liver." <u>J Biol Chem</u> **270**(27): 16347-16354.

Xu, S., M. Zhong, L. Zhang, Y. Wang, Z. Zhou, Y. Hao, W. Zhang, X. Yang, A. Wei, L. Pei and Z. Yu (2009). "Overexpression of Tfam protects mitochondria against beta-amyloid-induced oxidative damage in SH-SY5Y cells." <u>FEBS J</u> **276**(14): 3800-3809.

Yakubovskaya, E., K. E. Guja, E. T. Eng, W. S. Choi, E. Mejia, D. Beglov, M. Lukin, D. Kozakov and M. Garcia-Diaz (2014). "Organization of the human mitochondrial transcription initiation complex." <u>Nucleic Acids Res</u> **42**(6): 4100-4112.

Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen and C. R. Woese (1985). "Mitochondrial origins." <u>Proc Natl Acad Sci U S A</u> **82**(13): 4443-4447.

Yasukawa, T., A. Reyes, T. J. Cluett, M. Y. Yang, M. Bowmaker, H. T. Jacobs and I. J. Holt (2006). "Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand." <u>EMBO J</u> **25**(22): 5358-5371.

Yasukawa, T., M. Y. Yang, H. T. Jacobs and I. J. Holt (2005). "A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA." <u>Mol Cell</u> **18**(6): 651-662.

Zeviani, M. and S. Di Donato (2004). "Mitochondrial disorders." <u>Brain</u> **127**(Pt 10): 2153-2172.

Zhou, L., A. Chomyn, G. Attardi and C. A. Miller (1997). "Myoclonic epilepsy and ragged red fibers (MERRF) syndrome: selective vulnerability of CNS neurons does not correlate with the level of mitochondrial tRNAlys mutation in individual neuronal isolates." J Neurosci **17**(20): 7746-7753.

Zollo, O., V. Tiranti and N. Sondheimer (2012). "Transcriptional requirements of the distal heavy-strand promoter of mtDNA." Proc Natl Acad Sci U S A **109**(17): 6508-6512.

Chapter 2

MITOCHONDRIAL DYNAMICS AND MTDNA MAINTENANCE

Yun E Wang[1], David C Chan[1,2]

[1]Division of Biology, California Institute of Technology, Pasadena, CA

[2]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA

This chapter was published as part of an article in *Cell* (Chen, Vermulst et al. 2010).

I.   <u>Introduction</u>

<u>Proteins involved in fission and fusion</u>

The transmembrane protein Fzo was first found to mediate mitochondrial membrane fusion in spermatozoa of *D. melanogaster* (Hales and Fuller 1997). In mammals, there are two orthologs, Mfn1 and Mfn2, which can form both homo-oligomeric and hetero-oligomeric complexes to effect fusion between two mitochondria (Koshiba, Detmer et al. 2004, Meeusen, McCaffery et al. 2004). Membrane fusion in mammals is also mediated by the GTPase Opa1, which localizes to the intermembrane space and inner membrane (Cipolat, Martins de Brito et al. 2004, Chen, Chomyn et al. 2005). Loss of Mfn1, Mfn2, or Opa1 leads to mitochondrial fragmentation and changes in morphology, and to loss of membrane potential, indicating impaired mitochondrial function (Chen, Detmer et al. 2003, Chen, Chomyn et al. 2005).

Mitochondrial fission is well-characterized in yeast, where it is mediated by several proteins: Fis1, an integral outer membrane protein, binds indirectly to the dynamin-related protein Dnm1 (Drp1 in mammals) through the adapter proteins Mdv1 or Caf4. Dnm1, in turn, is thought to form mitochondrial tubules and to effect fission through constriction (Smirnova, Griparic et al. 2001). While Fis1 and Dnm1 are each required for mitochondrial fission in yeast, either of the adaptor proteins is sufficient for a wild-type phenotype. No mammalian homologues have yet been found for Mdv1 or Caf4.

Pathology of dysregulation in mitochondrial dynamics

Defects in mitochondrial dynamics have been implicated in several human neurodegenerative diseases. ADOA, the most common inherited optic neuropathy, is directly caused by mutations in the Opa1 gene, and results in degeneration of retinal ganglion cells (Alexander, Votruba et al. 2000, Delettre, Lenaers et al. 2000). Loss of Mfn2 function is linked to CMT disease type 2a, in which the symptoms of distal motor and sensory impairments starting at the extremities are caused by a neuronal defect causing degeneration of long axons (Zuchner and Vance 2006). Moreover, there has been one account of a patient with a dominant negative allele of Drp1, presenting with reduced head growth and optic atrophy, amongst other defects, with mortality at 1 month of age (Waterham, Koster et al. 2007). Importantly, mice deficient in any of the fusion proteins die midgestation (Chen, Detmer et al. 2003, Davies, Hollins et al. 2007), indicating that mitochondrial fusion is essential for viability. Fission has also been shown to be essential in cellular processes, as downregulation results in a highly interconnected network of tubular mitochondria and to defects in cell division (Smirnova, Shurland et al. 1998, Sesaki and Jensen 1999, Smirnova, Griparic et al. 2001) and in apoptosis (Frank, Gaume et al. 2001, Lee, Jeong et al. 2004). Finally, PD (Van Laar and Berman 2009), AD (Bossy-Wetzel, Barsoum et al. 2003), HD (Bossy-Wetzel, Petrilli et al. 2008), ALS (Knott and Bossy-Wetzel 2008), and several other neurodegenerative diseases have been correlated with dysfunction in dynamics, emphasizing the importance of mitochondrial fission and fusion in maintenance of cellular homeostasis.

There are many hypotheses addressing the cellular function of mitochondrial dynamics. First, dynamics may be involved in maintenance of proper mitochondrial function. Loss of fusion leads to defects in respiratory capacity, and morphological and membrane potential heterogeneity (Chen, Chomyn et al. 2005). Fusion and fission cycles allow for exchange of mitochondrial contents and could thus allow for rescue of respiratory capacity of deficient mitochondria. Such exchange of material would also allow for re-distribution of mtDNA-containing nucleoids, which provide the genomic material for transcription of essential components of the electron transport chain. In cells deficient in fusion, in contrast to wild-type cells, a large fraction of mitochondria lack nucleoids and have a lowered membrane potential (Chen, Chomyn et al. 2005), indicating that proper nucleoid distribution is crucial for correct mitochondrial function. Furthermore, maintenance of proper mitochondrial function seems to be essential for development, as evidenced by the embryonic lethality of fusion-deficient mice (Chen, Detmer et al. 2003).

Second, fission and fusion are important for maintaining proper distribution and recruitment of mitochondria in neurons. Due to the distance that neuronal processes can extend from the cell body, mitochondria must be transported to areas of high energy demand, such as pre- and post-synaptic terminals (Li, Okamoto et al. 2004, Chang, Honick et al. 2006). Importantly, loss of Drp1 leads to a defect in transport and a concomitant loss of capacity for synaptic transmission (Stowers, Megeath et al. 2002, Guo, Macleod et al. 2005), a phenotype that, along with loss of respiratory capacity in fusion-deficient cells and loss of exchange of mitochondrial material, could explain the

pathology of neurodegenerative diseases associated with perturbation of dynamics.

Regulation of mtDNA copy number

Depletion of mtDNA copy number is well known to cause cellular defects in bioenergetics (Baron, Kudin et al. 2007). Unsurprisingly, many proteins which regulate mtDNA copy number exist in pathways related to the mitochondrial dNTP salvage pathway and to cytoplasmic de novo dNTP synthesis (Tyynismaa and Suomalainen 2009). Interestingly, many neurodegenerative diseases, including ALS and FRDA (Baron, Kudin et al. 2007), as well as many mitochondrial myopathies (Alberio, Mineri et al. 2007), have also been associated with mtDNA copy number deficiencies. This suggests that loss of mtDNA has a severely negative impact on neurons and muscles, which have the highest energetic demand of all cells in the body. This also suggests that even in sporadic forms of neurodegenerative disease, dysregulation of copy number may lead to disease pathology.

Because mitochondrial dynamics is hypothesized to play an important role in maintenance of mitochondrial viability and in proper distribution of mitochondrial nucleoids, we hypothesized that it may also have a role in maintenance of mtDNA. We used qPCR to elucidate the role of the fusion proteins in modulating mitochondrial genome copy number.

II.   <u>Results</u>

<u>Using qPCR to characterize mtDNA copy number</u>

It was unknown what effect loss of mitochondrial fusion had on mtDNA copy number. This could be characterized by determining the average mtDNA copy number per cell in different cell lines. I thus chose to utilize singleplex qPCR using SYBR Green (Brilliant SYBR Green 2X Master Mix, Agilent #929548) as the fluorescent marker for amplification, taking advantage of the technique's high level of accuracy and sensitivity in measuring DNA copy number over a wide dynamic range. Whole DNA prep of cultured MEFs to maximally capture both nuclear and mitochondrial DNA was performed largely according to established protocols, with some modifications (Ausubel 2002) (see Protocols section for details). Four sets of primers with no predicted homodimers, heterodimers, or hairpins (as predicted by the Eurofins MWG Operon oligo analysis tool) were designed, in the 0kb, 4kb, 5kb, and 6kb stretches of the mitochondrial genome and yielding 220-260bp products. Four primer sets found in a validated qPCR primer database (realtimeprimers.org) were tested for genomic qPCR capability, with primers within the PECAM1, two PCR'ing between SCN4a and mGH on chromosome 11; and one set priming in ActB; each product was 230-235bp long. Optimization of qPCR conditions (Bio-Rad CFX96) using prepared WT MEF whole cell DNA resulted in the selection of the 4kbmt primer set for mitochondrial copy number determination and the PECAM1 primer set for nuclear because of their high levels of consistency ($R^2 = 1.00$ and 0.98, respectively) and PCR efficiency approaching 100% (100.5% and 95.7%,

respectively) over a wide dynamic range of a 1000-fold change in copy number concentration **(Figure 2.1A-B)**.

Loss of fusion leads to depleted mtDNA

Separate whole DNA preps of Mfn1-null, Mfn2-null, Opa1-null, and Mfn-double null MEFs ($n = 6$ for each) were run using these primers, and the relative copy number per cell was determined using either the analysis of the difference in threshold amplification between mtDNA and nuclear genome, or the $\Delta\Delta C(t)$ method using the known dilution standard curve as a standard; both yielded similar results. The average mtDNA to nuclear genome (Shen-Li, O'Hagan et al.) ratio in MEFs was found to be 440 mtDNA:nDNA. Strikingly, while absence of either Mfn1 or Mfn2 alone did not impact mtDNA copy number, absence of both Mfn's, which leads to loss of all outer membrane fusion, or absence of Opa1, which leads to loss of all inner membrane fusion, results in dramatically depleted mtDNA **(Figure 2.2A).**

To verify that this finding was due to the loss of the mitofusins in Mfn-null cells, and not due to other conditions which would cause different mtDNA:nDNA copy number in these cell lines, ds-Red (mock), Mfn1-myc, or Mfn2-myc was reintroduced into Mfn-null cells by retroviral infection, and protein levels were monitored via Western at 2 and 4 weeks post-infection; mtDNA:nDNA was also simultaneously tracked. Successful restoration of either mitofusin to mitochondria **(Figure 2.2B-C)** resulted in partial rescue

of the fragmented mitochondrial morphology as well an increase in mtDNA:nDNA levels to that of WT MEFs **(Figure 2.2D)**. This observation is corroborated by data from mouse studies by Hsiuchen Chen in our lab which demonstrate that mice homozygous mutant for both Mfn1 and Mfn2 in skeletal muscle (MLC-Cre/dm) also exhibit a severe decrease in mtDNA:nDNA, a phenotype not seen in Mfn1-/-,Mfn2+/- or Mfn1+/-,Mfn2-/-mice.

III.   <u>Discussion</u>

Our lab has found (Hsiuchen Chen, Chen et al., 2011) that Mfn-null mice present with a phenotype strikingly similar to that found in human mitochondrial myopathies, unsurprising given that mtDNA levels per nuclear genome in skeletal muscle tissue were only 7% of that found in wild-type mice, a phenotype absent in mice heterozygous for either Mfn1 or Mfn2. My analysis of MEF lines confirms that mtDNA depletion occurs in Mfn-double null and Opa1-null cells and not in Mfn1-null and Mfn2-null MEFs; complete ablation of either inner mitochondrial fusion or outer membrane fusion is required for the phenotype. Overexpression of Mfn1 or Mfn2 is readily able to rescue the depletion in Mfn-double null cells, leading to the conclusion that it is loss of the fusion proteins that leads to the depletion phenotype. This illustrates the key role of mitofusins in mtDNA maintenance, and suggests the possibility that dysregulation of dynamics may manifest itself in the form of perturbation of mtDNA copy number, leading to cellular dysfunction.

Mitochondrial membrane fission and fusion are crucial for cellular and organismal viability (Chan 2006, Chan 2006). The implications of dysfunction are severe, as mutations in the proteins implicated in fusion and fission lead to neurodegenerative disease (Alexander, Votruba et al. 2000, Delettre, Lenaers et al. 2000, Zuchner and Vance 2006). Beyond mutations in the machinery that directly affects the machinery, phenotypic perturbations in mitochondrial dynamics are correlated with many of the most common human neurodegenerative diseases and myopathies, such as AD and PD (Knott and Bossy-Wetzel 2008). This implies that mitochondrial function is crucial for neuronal and muscular viability, and that loss of function leads to atrophy and to a devastating phenotype. Moreover, mtDNA depletion, which causes further mitochondrial dysfunction, has been implicated in several of these neurodegenerative diseases (Baron, Kudin et al. 2007), suggesting that depletion could contribute to the severity of the phenotype.. It is certainly possible that mtDNA depletion due to ablation of fusion is a result of loss of exchange of mitochondrial content, thus leading to heterogenous distribution of proteins required for proliferation of the genome and a consistent depletion phenotype. However, a more compelling story would be the direct or indirect regulation of mtDNA copy number by the fusion proteins. Further studies addressing the interplay between the proteins directly involved with mtDNA copy number, such as POLG, Twinkle, mtSSB, TFAM, POLRMT, and TFB2M, and the fusion proteins, will shed light on this intriguing effect.

IV. <u>Methods</u>

Quantitative PCR of mouse mtDNA and nuclear genome

One set of primers, 4kbmtF/R, and one set of nuclear DNA primers, PECAMF/R, was optimized for mouse mtDNA:

Primer ID	Sequence	Chromosome	Base pairs
4kbmtF	CCTATCACCCTTGCCATCAT	mt	3920 – 3939
4kbmtR	GAGGCTGTTGCTTGTGTGAC	mt	4094 – 4113
PECAMF	ATGGAAAGCCTGCCATCATG	11	55425 – 55406
PECAMR	TCCTTGTTGTTCAGCATCAC	11	55190 – 55209

qPCR was performed on Bio-Rad CFX-96 machine with using Brilliant SYBR Green 2X Master Mix, Agilent #929548 and the following cycling conditions: 10min@95C, [30sec@95C, 1min@58C, 1min@72C]x40.

Expression of Mfn1 and Mfn2 in MEFs

Viral supernatant with virus made from retroviral vector containing either Mfn1 or Mfn2 was applied to MEFs and incubated for 12 hours prior to replacement with normal growth media [DMEM (Invitrogen #11995) containing 10% fetal bovine serum, penicillin and streptomycin, and additional L-glutamine (2mM)].

Western Blotting of Mfn1 and Mfn2 expression

Cells were harvested at 0, 2, and 4 weeks post-infection and run on 10% SDS-PAGE gels. Primary antibodies were anti-Mfn1 (1:1000) and anti-Mfn2 (1:1000). Secondary antibodies were anti-chicken (1:10,000) and anti-rabbit (1:10,000, Jackson #111-035-003), respectively. Anti-actin (1:500, Imgenex #IMG-5142A) was used for loading control.

Immunocytochemistry of Mfn1 and Mfn2 expression

Cells were plated onto poly-lysine coated coverslips 24 hours prior to fixation with 10% formalin and permeabilization with acetone at -20C. Following 1hr block in PBS + 5% FBS, incubations were as follows: 1hr with anti-myc antibody (1:10, 9E10 clone) in PBS + 5% FBS; 3x 10min wash in PBS; 1hr with anti-mouse AF488 (1:500, Invitrogen #A10680); 3x 10min wash in PBS; mounting onto glass slides. Cells were visualized using Zeiss LSM 710 confocal microscope with PlanApochromat 63X/14 oil objective.

V.   Figure Legends

Figure 2.1. Primers for quantification of relative mtDNA copy number per cell. Primers used for quantification of mtDNA copy number (4kbmtF and 4kbmtR) and nuclear genome copy number (PECAMF and PECAMR). Standard dilution curve, as assayed on whole cell MEF DNA prep, for 4kbmt (A) and PECAM (B) primer sets shows linearity through 1000-fold dilution range.

Figure 2.2. Complete loss of fusion leads to mtDNA depletion. (A) Analysis of mtDNA:nDNA in WT MEFs and cells lacking Mfn1, Mfn2, both mitofusins, or Opa1 demonstrate that severe loss of mtDNA only occurs with complete ablation of fusion, as in Mfn-null and Opa1-null cells. Infection of Mfn-null cells with either Mfn1 or Mfn2 (C) leads to partial rescue on mitochondrial morphology and to a rapid, sustained increase in mtDNA copy number to that of WT MEFs. In (C), red immunostaining is HSP60 for labeling of mitochondria, and green is for myc, to which Mfn1 and Mfn2 are attached. I. WT; II. Mfn-null; III. Mfn-null + Mfn1-myc; IV. Mfn-null + Mfn2-myc.

Figure 2.1.

A

4kbmt Primer Standard Curve

Ct Value vs Dilution (1:10, 1:100, 1:1000, 1:10000)

B

PECAM Primer Standard Curve

Ct Value vs Dilution (1:10, 1:100, 1:1000, 1:10000)

Figure 2.2.

A

B

C
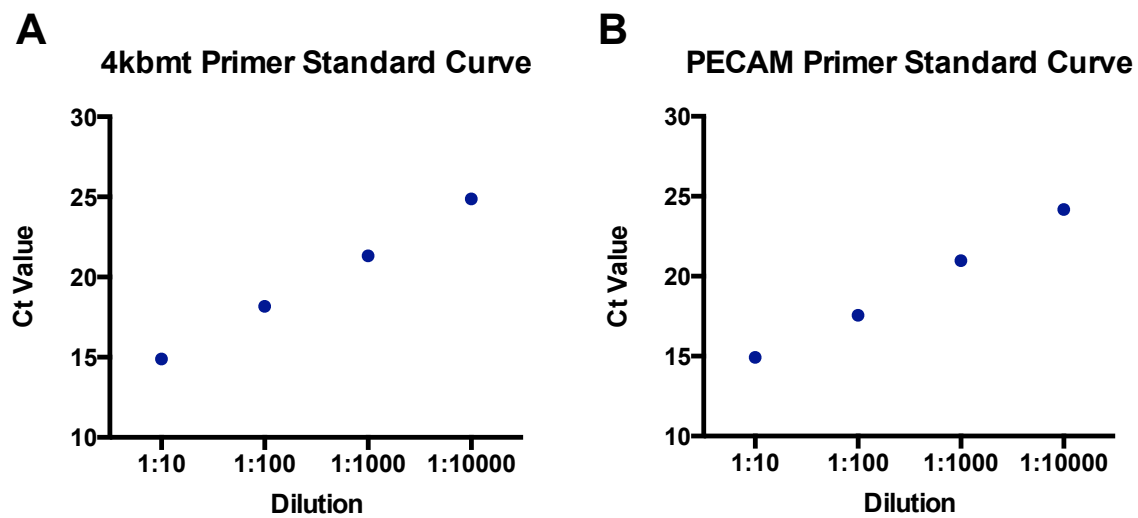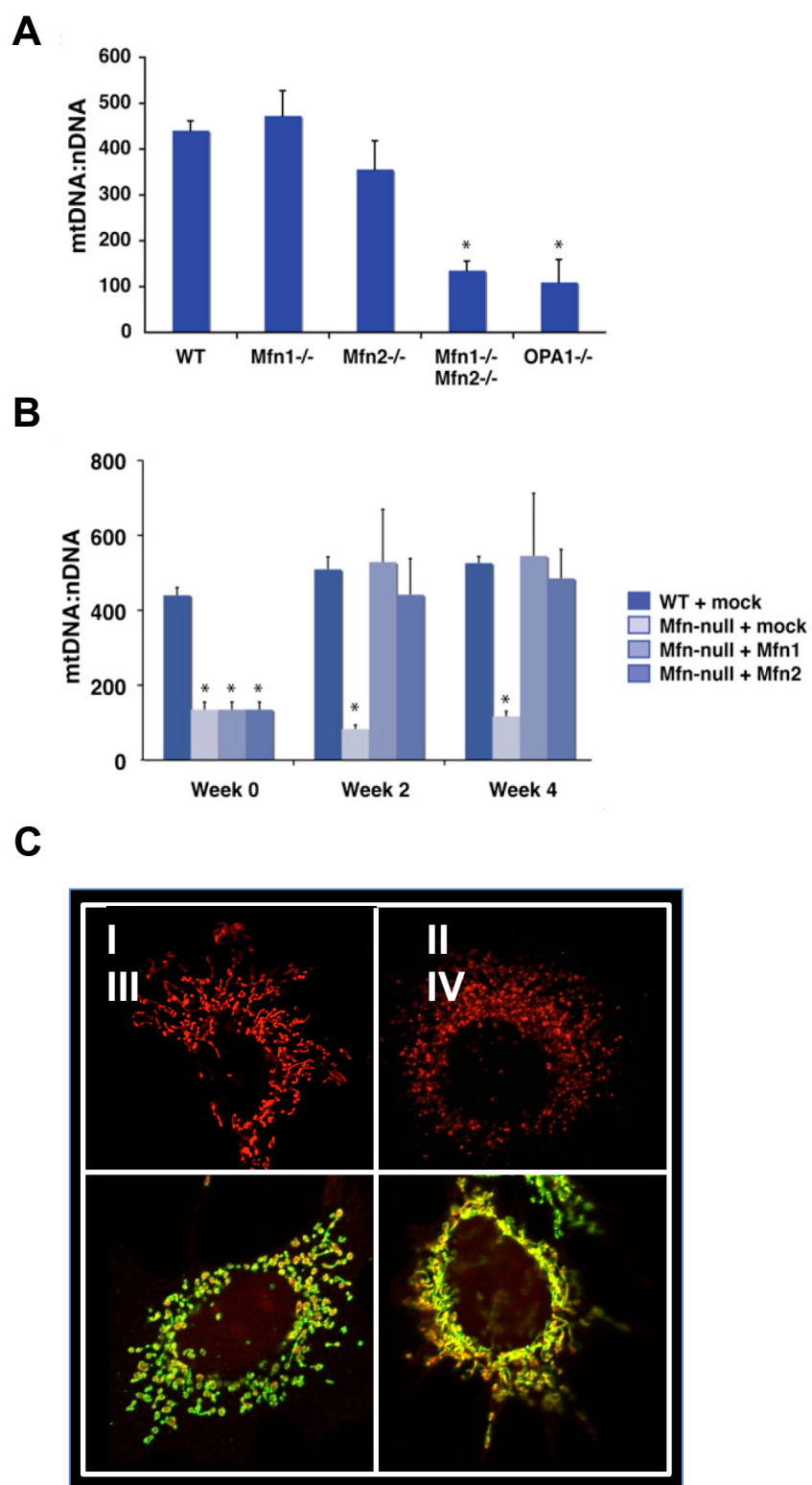
BIBLIOGRAPHY

Alberio, S., R. Mineri, V. Tiranti and M. Zeviani (2007). "Depletion of mtDNA: syndromes and genes." <u>Mitochondrion</u> **7**(1-2): 6-12.

Alexander, C., M. Votruba, U. E. Pesch, D. L. Thiselton, S. Mayer, A. Moore, M. Rodriguez, U. Kellner, B. Leo-Kottler, G. Auburger, S. S. Bhattacharya and B. Wissinger (2000). "OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28." <u>Nat Genet</u> **26**(2): 211-215.

Ausubel, F. M. (2002). <u>Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology</u>. New York, Wiley.

Baron, M., A. P. Kudin and W. S. Kunz (2007). "Mitochondrial dysfunction in neurodegenerative disorders." <u>Biochem Soc Trans</u> **35**(Pt 5): 1228-1231.

Bossy-Wetzel, E., M. J. Barsoum, A. Godzik, R. Schwarzenbacher and S. A. Lipton (2003). "Mitochondrial fission in apoptosis, neurodegeneration and aging." <u>Curr Opin Cell Biol</u> **15**(6): 706-716.

Bossy-Wetzel, E., A. Petrilli and A. B. Knott (2008). "Mutant huntingtin and mitochondrial dysfunction." <u>Trends Neurosci</u> **31**(12): 609-616.

Chan, D. C. (2006). "Mitochondria: dynamic organelles in disease, aging, and development." <u>Cell</u> **125**(7): 1241-1252.

Chan, D. C. (2006). "Mitochondrial fusion and fission in mammals." Annu Rev Cell Dev Biol **22**: 79-99.

Chang, D. T., A. S. Honick and I. J. Reynolds (2006). "Mitochondrial trafficking to synapses in cultured primary cortical neurons." J Neurosci **26**(26): 7035-7045.

Chen, H., A. Chomyn and D. C. Chan (2005). "Disruption of fusion results in mitochondrial heterogeneity and dysfunction." J Biol Chem **280**(28): 26185-26192.

Chen, H., S. A. Detmer, A. J. Ewald, E. E. Griffin, S. E. Fraser and D. C. Chan (2003). "Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development." J Cell Biol **160**(2): 189-200.

Chen, H., M. Vermulst, Y. E. Wang, A. Chomyn, T. A. Prolla, J. M. McCaffery and D. C. Chan (2010). "Mitochondrial fusion is required for mtDNA stability in skeletal muscle and tolerance of mtDNA mutations." Cell **141**(2): 280-289.

Cipolat, S., O. Martins de Brito, B. Dal Zilio and L. Scorrano (2004). "OPA1 requires mitofusin 1 to promote mitochondrial fusion." Proc Natl Acad Sci U S A **101**(45): 15927-15932.

Davies, V. J., A. J. Hollins, M. J. Piechota, W. Yip, J. R. Davies, K. E. White, P. P. Nicols, M. E. Boulton and M. Votruba (2007). "Opa1 deficiency in a mouse model of autosomal dominant optic atrophy impairs mitochondrial morphology, optic nerve structure and visual function." Hum Mol Genet **16**(11): 1307-1318.

Delettre, C., G. Lenaers, J. M. Griffoin, N. Gigarel, C. Lorenzo, P. Belenguer, L. Pelloquin, J. Grosgeorge, C. Turc-Carel, E. Perret, C. Astarie-Dequeker, L. Lasquellec, B. Arnaud, B. Ducommun, J. Kaplan and C. P. Hamel (2000). "Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy." <u>Nat Genet</u> **26**(2): 207-210.

Frank, S., B. Gaume, E. S. Bergmann-Leitner, W. W. Leitner, E. G. Robert, F. Catez, C. L. Smith and R. J. Youle (2001). "The role of dynamin-related protein 1, a mediator of mitochondrial fission, in apoptosis." <u>Dev Cell</u> **1**(4): 515-525.

Guo, X., G. T. Macleod, A. Wellington, F. Hu, S. Panchumarthi, M. Schoenfield, L. Marin, M. P. Charlton, H. L. Atwood and K. E. Zinsmaier (2005). "The GTPase dMiro is required for axonal transport of mitochondria to Drosophila synapses." <u>Neuron</u> **47**(3): 379-393.

Hales, K. G. and M. T. Fuller (1997). "Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase." <u>Cell</u> **90**(1): 121-129.

Knott, A. B. and E. Bossy-Wetzel (2008). "Impairing the mitochondrial fission and fusion balance: a new mechanism of neurodegeneration." <u>Ann N Y Acad Sci</u> **1147**: 283-292.

Koshiba, T., S. A. Detmer, J. T. Kaiser, H. Chen, J. M. McCaffery and D. C. Chan (2004). "Structural basis of mitochondrial tethering by mitofusin complexes." <u>Science</u> **305**(5685): 858-862.

Lee, Y. J., S. Y. Jeong, M. Karbowski, C. L. Smith and R. J. Youle (2004). "Roles of the mammalian mitochondrial fission and fusion mediators Fis1, Drp1, and Opa1 in apoptosis." Mol Biol Cell **15**(11): 5001-5011.

Li, Z., K. Okamoto, Y. Hayashi and M. Sheng (2004). "The importance of dendritic mitochondria in the morphogenesis and plasticity of spines and synapses." Cell **119**(6): 873-887.

Meeusen, S., J. M. McCaffery and J. Nunnari (2004). "Mitochondrial fusion intermediates revealed in vitro." Science **305**(5691): 1747-1752.

Sesaki, H. and R. E. Jensen (1999). "Division versus fusion: Dnm1p and Fzo1p antagonistically regulate mitochondrial shape." J Cell Biol **147**(4): 699-706.

Shen-Li, H., R. C. O'Hagan, H. Hou, Jr., J. W. Horner, 2nd, H. W. Lee and R. A. DePinho (2000). "Essential role for Max in early embryonic growth and development." Genes Dev **14**(1): 17-22.

Smirnova, E., L. Griparic, D. L. Shurland and A. M. van der Bliek (2001). "Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells." Mol Biol Cell **12**(8): 2245-2256.

Smirnova, E., D. L. Shurland, S. N. Ryazantsev and A. M. van der Bliek (1998). "A human dynamin-related protein controls the distribution of mitochondria." J Cell Biol **143**(2): 351-358.

Stowers, R. S., L. J. Megeath, J. Gorska-Andrzejak, I. A. Meinertzhagen and T. L. Schwarz (2002). "Axonal transport of mitochondria to synapses depends on milton, a novel Drosophila protein." <u>Neuron</u> **36**(6): 1063-1077.

Tyynismaa, H. and A. Suomalainen (2009). "Mouse models of mitochondrial DNA defects and their relevance for human disease." <u>EMBO Rep</u> **10**(2): 137-143.

Van Laar, V. S. and S. B. Berman (2009). "Mitochondrial dynamics in Parkinson's disease." <u>Exp Neurol</u>.

Waterham, H. R., J. Koster, C. W. van Roermund, P. A. Mooyer, R. J. Wanders and J. V. Leonard (2007). "A lethal defect of mitochondrial and peroxisomal fission." <u>N Engl J Med</u> **356**(17): 1736-1741.

Zuchner, S. and J. M. Vance (2006). "Mechanisms of disease: a molecular genetic update on hereditary axonal neuropathies." <u>Nat Clin Pract Neurol</u> **2**(1): 45-53.

Chapter 3

GENOME-WIDE ANALYSIS REVEALS COATING OF THE MITOCHONDRIAL GENOME BY TFAM

Yun E Wang[1], Georgi K Marinov[1], Barbara J Wold[1], David C Chan[1,2]

[1]Division of Biology, California Institute of Technology, Pasadena, CA

[2]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA

I. <u>Abstract</u>

Mitochondria contain a 16.6 kb circular genome encoding 13 proteins as well as mitochondrial tRNAs and rRNAs. Copies of the genome are organized into nucleoids containing both DNA and proteins, including the machinery required for mtDNA replication and transcription. The transcription factor TFAM is critical for initiation of transcription and replication of the genome, and is also thought to perform a packaging function. Although specific binding sites required for initiation of transcription have been identified in the D-loop, little is known about the characteristics of TFAM binding in its nonspecific packaging state. In addition, it is unclear whether TFAM also plays a role in the regulation of nuclear gene expression. Here we investigate these questions by using ChIP-seq to directly localize TFAM binding to DNA in human cells. Our results demonstrate that TFAM uniformly coats the whole mitochondrial genome, with no evidence of robust TFAM binding to the nuclear genome. Our study represents the first direct assessment of TFAM binding on a genome-wide scale in human cells.

II. <u>Introduction</u>

Mitochondria are essential eukaryotic organelles, serving as the epicenter of ATP production in the cell through oxidative phosphorylation. To perform this bioenergetic function, mitochondria utilize gene products encoded by the mitochondrial genome, a circular DNA that is 16.6 kb long. This genome is organized into DNA/protein structures

termed nucleoids (Bogenhagen, Rousseau et al. 2008). Mitochondrial DNA (mtDNA) encodes thirteen components of the electron transport chain, as well as 22 tRNAs and two ribosomal RNA genes. These gene products are essential for the proper function of the respiratory chain, and therefore maintenance of mtDNA levels and sequence fidelity is essential for cellular bioenergetics. In a human cell, there are hundreds to thousands of copies of the mtDNA genome (Bogenhagen Bogenhagen and Clayton 1974, Satoh Satoh and Kuroiwa 1991). Damage or depletion of mtDNA causes numerous inherited disorders, including Alpers' Disease, ataxia neuropathy spectrum, and progressive external ophthalmoplegia (Suomalainen and Isohanni 2010, Stumpf, Saneto et al. 2013). Furthermore, loss and damage to mtDNA has been implicated in cardiovascular disease (Sugiyama, Hattori et al. 1991, Ide, Tsutsui et al. 2001, Karamanlidis, Nascimben et al. 2010, Karamanlidis, Bautista-Hernandez et al. 2011), diabetes (Maassen, LM et al. 2004, Simmons, Suponitsky-Kroyter et al. 2005, Gauthier, Wiederkehr et al. 2009), neurodegenerative disorders such as Alzheimer's (Coskun, Beal et al. 2004, Coskun, Wyrembak et al. 2012), and aging (Corral-Debrinski, Shoffner et al. 1992, Trifunovic and Larsson 2008). Strikingly, increasing mtDNA copy number promotes cell survival or function in many models of disease associated with decreased mtDNA abundance, such as diabetes (Suarez, Hu et al. 2008, Gauthier, Wiederkehr et al. 2009), aging (Hayashi, Yoshida et al. 2008), Alzheimer's (Xu, Zhong et al. 2009), and Parkinson's (Keeney, Quigley et al. 2009, Piao, Kim et al. 2012). Thus, it is critical to understand how mtDNA copy number and integrity are maintained.

Mitochondrial transcription factor A (TFAM) is a DNA binding protein that plays multiple roles in regulating mtDNA function. As a sequence-specific transcription factor, it binds upstream of the light strand promoter (LSP) and heavy strand promoter 1 (HSP1) to activate initiation of transcription. At these sites, the footprint of TFAM binding is ~22 bp long (Fisher and Clayton 1988, Ngo, Kaiser et al. 2011). As a result, TFAM is essential for production of gene products from the mitochondrial genome. In addition, TFAM is required for normal mtDNA copy number, because RNA primers generated from LSP are used to prime mtDNA replication (Chang Chang and Clayton 1984, Chang Chang and Clayton 1985). Mice heterozygous for a knockout of TFAM exhibit not only an expected reduction (22%) in mitochondrial transcript levels in the heart and kidney, but also a universal 34% reduction in mtDNA copy number across all assayed tissues. Furthermore, homozygous knockout mice have no detectable levels of mtDNA and die during embryogenesis (Larsson, Wang et al. 1998), highlighting the importance of TFAM in maintenance of mtDNA levels and in cellular and organismal viability.

Apart from its sequence-specific functions, TFAM is thought to organize the mtDNA genome by coating it in a nonspecific manner. Although how TFAM packages mtDNA is not well-understood, it is known to bind nonspecifically to DNA (Fisher, Parisi et al. 1989) and is estimated to be sufficiently abundant to coat the genome completely (Alam, Kanki et al. 2003, Ekstrand, Falkenberg et al. 2004, Kaufman, Durisic et al. 2007). One model suggests that nonspecific binding radiates from the TFAM LSP binding site, which acts as a nucleation site for subsequent cooperative binding in a phased pattern to yield an inter-genome homogeneous pattern of binding (Fisher,

Lisowsky et al. 1992, Ghivizzani, Madsen et al. 1994). The packaging function of TFAM appears to have important consequences for maintenance of the mtDNA genome. A TFAM variant that is deficient in transcriptional activation but competent in DNA binding is capable of preventing mtDNA depletion (Kanki, Ohgaki et al. 2004). Therefore, as a prominent component of mtDNA nucleoids, TFAM appears to coat the mitochondrial genome, perhaps protecting it from turnover or deleterious damage.

Despite the importance of the associations of TFAM with mtDNA in the maintenance of mtDNA integrity and in cellular viability, these interactions have not been characterized in vivo. Therefore, to capture a high-resolution profile of TFAM-mtDNA interactions across the entire mitochondrial genome, we performed chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-seq) for TFAM in human HeLa cells.

Adaptation of ChIP-Seq to mtDNA

Canonical ChIP-seq was developed (Johnson, Mortazavi et al. 2007) in order to capture the in vivo interactions of transcription factors binding to the nuclear genome. This involves binding protein to the DNA with which it associates, shearing of DNA into smaller pieces in order to decrease the length of the DNA associated with the protein, capture of the protein via immunoprecipitation, and sequencing of the attached protein. Because DNA that is bound to protein is protected from shearing, this yields a series of

short sequences corresponding to the DNA directly adjacent to the protein of interest. Various computational techniques are then able to identify the regions of interest across the genome. As such, the general protocol for ChIP-seq is (per the Myers laboratory, http://myers.hudsonalpha.org/documents/Myers%20Lab%20ChIP-seq%20Protocol%20v041610.pdf):

1) Fixation of cells with formaldehyde to crosslink protein to DNA

2) Resuspension of cells in Farnham buffer to selective lyse cells while maintaining intact nuclei

3) Selective centrifugation of nuclei

4) Resuspension of pellet in RIPA buffer to lyse nuclei

5) Sonication to shear DNA to desired size

6) Immunoprecipitation using antibody against transcription factor of interest

7) Reversal of crosslinks and isolation of DNA

Even when performing ChIP-seq against a new transcription factor requires optimization of the protocol for optimized DNA recovery. Specifically, in the case of the mtDNA-targeted ChIP-seq that TFAM requires, such optimization is even more of a requirement. Fortunately, HeLa cells have been utilized in ChIP-seq experiments previously and work well with the standard protocol. However, the following steps remained to be optimized:

- Recovery of mtDNA in the supernatant used for IP

- Optimization of TFAM-specific antibodies used

-   Optimization of DNA sonication

Each of these steps was independently optimized before the full ChIP-seq protocol was performed, analyzed for mtDNA content, and sent for sequencing.

III. <u>Results</u>

HeLa Cell mtDNA ChIP-seq Optimization

i. Recovery of mitochondrial fraction in cellular lysate

The classic protocol for ChIP-seq involves resuspension of cross-linked DNA-protein complexes with Farnham buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40, Roche Protease Inhibitor Cocktail) first to solubilize the cellular membrane, followed by a 2000RPM centrifugation to pellet nuclei, removal of the supernatant, and resuspension of the pellet in RIPA buffer (1XPBS, 1% NP-40, 0.5% sodium deoxycholate, 1% SDS) to solubilize the nuclear membrane. However, the soft spin is not sufficient to pellet mitochondria, and therefore the pellet would most likely not contain mtDNA.

In order to determine the best method for segregation of mitochondria during this procedure, Westerns were performed on three different lysates in noncrosslinked HeLa cells: 1) mitochondria isolated from cells using nitrogen-bomb protocol ("Mito"); 2) the Farnham-based supernatant of the 2000RPM spin in the classic protocol ("Farnham"); and

3) the resuspended pellet after removal of the Farnham supernatant, ("RIPA"). The lysates were assayed for H2B (nuclear), Mfn2 (outer mitochondrial membrane), Opa1 (inner mitochondrial membrane), TFAM (mitochondrial matrix), and GAPDH (cytoplasmic/mitochondrial), and β-Actin (loading control) **(Figure 3.1A)**. β-Actin levels were consistent across each sample. Unsurprisingly, GAPDH was not found in the RIPA sample, remaining in the Farnham fraction; GAPDH was also found in the Mito fraction, but this is unsurprising given that GAPDH is known to be able to localize to mitochondria (Tarze, Deniaud et al. 2007). Also, as expected, a large amount of H2B was recovered in the RIPA sample, but not in the Farnham. Where there was recovery of H2B in the mitochondrial prep, this may be attributed to some nuclear contamination in the preparation, and is seen consistently across preps.

In terms of the mitochondrial proteins, neither Mfn2 nor Opa1 were seen in the RIPA sample, although this may be attributed to the low relative amounts of these proteins in cells. It is unclear why Opa1 was found in the Farnham sample, as neither TFAM nor Mfn1 were found. In the RIPA sample, a significant amount of TFAM was recovered, comparable to concentrations in the Mito sample, and TFAM levels were low in the Farnham sample. This is possibly because whole cells, unlysed by the Farnham buffer, were centrifuged into the pellet used for the RIPA sample. Indeed, in some cell lines, douncing is required to fully break the cell membrane and release the nucleus. Because we sought to recover TFAM and its interactions with mtDNA, we concluded that the majority of TFAM can be isolated using either a mitochondrial isolation, or the RIPA pellet. However, due to the ambiguity of the localization of other mitochondrial proteins, and due

to a desire to capture a maximal number of TFAM-DNA interactions, a whole-cell RIPA lysate would allow for full recovery of the protein. Although mitochondrial isolation would also be able to do so, due to the difficulty of isolating mitochondria from crosslinked cells, we chose to proceed with the whole-cell RIPA buffer lysis protocol. This method of lysis also had another benefit: as it has been suggested that TFAM localizes to the nucleus and regulates nuclear expression, recovery of nuclear DNA would allow for analysis of nuclear binding patterns as well.

ii.   Optimization of antibodies for TFAM IP

A crucial, rate-limiting step of ChIP-seq is identifying a high-quality antibody that is capable of faithfully immunoprecipitating the protein of interest in cross-linked cells. Our lab had used full-length TFAM in order to custom-generate antibodies in mouse (Susan Ou, Caltech Monoclonal Antibody Facility), and screened for ELISA effectiveness (Elizabeth L. Nelson). Three top candidates were identified: H101A6, 20G2C12, and 20F8A9.

Furthermore, two different systems of Dynabeads for immunoprecipitation were tested for effectiveness. The first utilized Dynabeads MyOne Streptavidin T1 (Invitrogen #65601), in conjunction with protein L as adapter between streptavidin and the antibody. Because protein L can associate with all antibodies containing a kappa chain, it can pull down many antibodies from many classes, such as IgG, IgM, IgA, IgE, and IgD. The

second utilized Dynabeads M-280 sheep anti-mouse IgG (Invitrogen #11202D), which is only capable of pulling down mouse IgG antibodies. In the first method, streptavidin, which is covalently linked onto the surface of the magnetic Dynabeads, is incubated with protein L, which binds to streptavidin with near-covalent strength. After washing to remove excess protein L, conjugated beads were then incubated with each antibody. The second method is the classic method of antibody conjugation to beads, where the Dynabeads, to which sheep anti-mouse antibodies have been covalently linked, are incubated with the antibody of choice.

Testing of the two Dynabeads methods was performed on the same two fractions of noncrosslinked cells as previously: the Farnham supernatant post-centrifugation, and the pellet resuspended in RIPA buffer. Because the 20F8A9 antibody was not available in sufficient quantities for experiments, only the H101A6 and 20G2C12 antibodies were utilized to test the two Dynabeads methods. Cell lysate was incubated with each Dynabeads type, already conjugated to each respective TFAM antibody, overnight with nutation. Beads were then washed and conjugated protein was recovered using sample buffer containing 2% SDS. Immunoblotting with a custom-generated anti-TFAM antibody optimized for Westerns, 18G102B2E11, with protein A/G as secondary, showed no recovery of TFAM using the streptavidin Dynabeads or using the H101A6 antibody **(Figure 3.1B)**. On the other hand, immunoprecipitation using sheep anti-mouse Dynabeads and the 20G2C12 antibody pulled down significant amounts of TFAM. TFAM was even recovered, albeit in smaller quantities, in the Farnham lysate sample. Thus, antibody H101A6 failed to immunoprecipitate TFAM, either because it was not captured by the

sheep anti-mouse Dynabeads, or because it cannot bind to TFAM under the immunoprecipitation conditions utilized.

When sufficient quantities of the 20F8A9 antibody were available, immunoprecipitation using the sheep anti-mouse Dynabeads was performed on whole cell RIPA lysate and found to immunoprecipitate TFAM, although at lower levels than of the 20G2C12 antibody **(Figure 3.1C)**. TFAM was not pulled down using Dynabeads conjugated to an anti-Myc antibody, demonstrating the specificity of the antibody and of the protocol. Finally, to verify that TFAM could still be captured when cells were crosslinked, whole cell lysate was incubated with 20G2C12 antibody conjugated beads, and the decrosslinked protein positively blotted for TFAM **(Figure 3.1D)**. Although both the 20G2C12 and 20F8A9 antibodies were able to pull down TFAM, subsequent experiments were performed with only the 20G2C12 antibody due to its stronger immunoprecipitation capability.

iii.   Further characterization of TFAM IP antibodies

To further validate the two antibodies that were capable of immunoprecipitating TFAM, 20G2C12 and 20F8A9, they were utilized as detection antibodies for Western blot and for immunocytochemistry. When utilized for Western blot, the antibodies should be able to uniquely detect a ~23kDa band for TFAM, and immunocytochemistry should demonstrate binding to protein localized in a generally punctate manner to the

mitochondria. However, the absence of either of these qualities does not denote that the antibody is incapable of binding to TFAM, as some antibodies do not detect the corresponding protein in all uses.

Western blot of HeLa cell whole cell lysate with the 20G2C12 antibody (1:200) yielded a single, ~23kDa band, confirming its specificity **(Figure 3.1E)**. The 20F8A9 antibody did not detect any bands, but this is not unexpected, as discussed above. Immunocytochemistry using these two antibodies as primary antibody in conjunction with anti-PPIF for mitochondrial matrix labeling and anti-mouse AF488 and anti-rabbit AF546 was performed, and imaging was using a Zeiss LSM 710 confocal microscope with PlanApochromat 63X/14 oil objective. Cells were HeLa cells plated onto poly-lysine coated glass coverslips 48 hours prior to fixation in formaldehyde and permeabilization with 0.1% Triton X-100. Cells were also stained with DAPI to visualize nuclei. Representative images of the immunostained cells can be seen in **(Figure 3.2A)**. Thus, using both antibodies, we visualize mitochondria-specific staining which is somewhat punctate in morphology, as has been previously described as typical of TFAM staining.

We also wanted to determine whether the signal detected by the antibodies colocalized with mtDNA nucleoid puncta. Immunocytochemistry with the TFAM antibodies and an anti-DNA primary and corresponding secondary antibody was utilized. This particular immunostaining was performed sequentially, because both the TFAM and DNA antibodies are raised in mouse; no DNA antibodies raised in other species are currently available for purchase. Fortunately, we took advantage of the fortuitous fact that

our DNA antibody is an IgM, which could be of a different class than our TFAM antibodies, as the majority of antibodies raised against an immunogen are IgG. A protocol was optimized to eliminate cross-reactivity, where an IgM-specific secondary antibody was utilized to detect the DNA antibody. The order of antibody application was thus as follows: anti-TFAM antibody; anti-mouse IgG antibody; anti-DNA antibody; and anti-mouse IgM antibody. This lead to no detectable background **(Figure 3.2C)**. Visualization of nucleoids in relation to the 20G2C12 and 20F8A9 antibody staining revealed perfect colocalization of signals, with the TFAM signal enveloping that of the mtDNA, as expected **(Figure 3.2B)**.

iv.   *Optimization of DNA sonication*

The precision of ChIP-seq is highly dependent on the size of DNA after it has been sonicated. Because ChIP-seq relies on sequencing of just the ends of DNA in a sample, smaller DNA fragments leads to more precise localization of interactions; size is even more crucial for mapping of TFAM, due to the small size of the mitochondrial genome. Furthermore, efficiency of sonication can vary widely between machines, even those that are the same model. Thus, optimization of sonication using a QSonica Sonicator 4000 was performed on crosslinked whole cell lysate without the immunoprecipitation step. The Myer lab, using a Sonics VibraCell sonicator, recommends sonicating at 30s on/30s off intervals, at 5W power, for a total of 6 30s on intervals as a starting point. To test multiple conditions, the following conditions were used: 10amp (arbitrary unit of power, device can be set between 0 and 100amp), for 2, 4, 6, 8, or 30 sets; and 20amp, for 2, 4, 6, 8, or 30

sets. After sonication, samples were decrosslinked and run on a 2.5% agarose gel to determine the size of the sheared DNA **(Figure 3.3A)**. Three lanes were loaded for each sample, each successive lane with double the relative amount of sample; the first, second, and third lanes of each sample were of equal relative quantity, respectively. Unfortunately, the conditions were largely unsuccessful, with the exception of the conditions for each amplitude with 30 sets, suggesting that a higher number of sets was required to shear DNA to an acceptable length.

This hypothesis was verified, and conditions for sonication were optimized for an average DNA length of 150-200bp. The best conditions were assayed to be 40amp for 45 sets with lysates of lower concentration **(Figure 3.3B)**, and 40amp for 90 sets with lysates of concentration of cells required for successful ChIP-seq **(Figure 3.3C)**. Due to the extremely long time-frame of sonication, however, it was imperative to find a method to cool the sample during this step. Sonication at such a high power also generates additional heat, raising concerns of damaging or denaturing the protein, inadvisable for a protocol reliant on immunoprecipitation. Therefore, an experimental setup was devised where the sample was held in a 3:1 isopropanol:water bath kept at approximately -35°C with agitation from a stirbar beneath the sample. This succeeded in keeping the sample below room temperature, although the precise temperature is not known due to the difficulty of obtaining an accurate measurement in such a small sample.

v.   *Verifying capture of mtDNA in ChIP-seq samples*

After optimization of the aforementioned steps, the full protocol performed on 150 x 10^6 HeLa cells using the 20G2C12 antibody yielded 139ng of DNA for the IP (via Qubit Fluorometer, Invitrogen), and 2.06ug of control DNA, derived from removing a 100ul aliquot of the 1ml sonicated whole cell lysate before incubation with Dynabeads. However, it was important to verify that mtDNA had been captured before sending the sample was sent for sequencing. Therefore, qPCR was utilized to analyze whether was an increase in mtDNA copy number relative to nuclear genome.

Six primers sets to the LSP promoter, and four to the HSP promoter, where TFAM has previously been DNase footprinted, were designed. Primers in the iB5s were utilized for nuclear genome quantification, having previously been validated by other labs. Two primer sets in the LSP promoter (LSP1 and LSP2), and two in the HSP promoter (HSP3 and HSP4) were found to have high consistency and PCR efficiency approaching over a wide dynamic range. These primer pairs were used to quantify the amount of enrichment of mtDNA sequences with TFAM immunoprecipitation. Fortunately, while there were 190-300 copies of each of the mtDNA templates per nuclear genome in the nonimmunoprecipitated, sonicated sample, there were between 65,700 and 105,000 copies per genome in the TFAM IP DNA **(Figure 3.4)**, an enrichment of up to ~400-fold. Of note, it is not surprising that the exact copy number determined by each mtDNA primer set is not exactly the same, as qPCR was conducted on sonicated DNA, which is subject to breakage bias.

Detection of TFAM-DNA interactions using ChIP-seq

Given the high efficiency of 20G2C12 in immunoprecipitating TFAM, as well as its high specificity, we used it to capture TFAM-associated DNA fragments for ChIP-seq analysis. HeLa cells were cultured in DMEM (Invitrogen #11995) containing 10% bovine serum (Invitrogen #16170), penicillin and streptomycin, and additional L-glutamine (2mM). It was imperative that cells be fed 24 hours before harvest for ChIP-seq, and the harvest performed at 80-90% confluency, in order to capture the cells while they were actively growing. mtDNA copy number increases (and thus presumably the dynamics of protein binding to the genome changes as well) when cells reach confluency and no longer proliferate normally. After formaldehyde fixation and quenching, cells were resuspended in RIPA buffer. Lysate was sonicated prior to immunoenrichment and size-selected prior to library building so that the average fragment length of the final library was centered around 200 bp **(Figure 3.5)**, a fragment distribution allowing for high-resolution deconvolution of binding events. Sequencing was performed using Illumina GAIIx and Illumina HiSeq 2000 machines. In total, we generated 3 replicates and matching controls. The sequencing depth of all samples was between 18 million and 48 million mappable reads, which is generally sufficient for comprehensive identification of transcription factor binding sites (Landt, Marinov et al. 2012).

A common concern with ChIP-seq datasets is the variability of enrichment for true binding events as compared to background. In a typical ChIP-seq experiment, a

minority of sequencing reads originates from binding events, with the majority representing random genomic DNA. Even for the same DNA binding factor, large variations in the strength of enrichment can be observed, and therefore it is critical to assess the degree of enrichment before downstream analysis. A number of ChIP-seq quality control metrics have been developed (Landt, Marinov et al. 2012) for nuclear transcription factors. However, TFAM is expected to bind to the mitochondrial genome, which has very different characteristics from the nuclear genome. In addition, it is predicted to bind both in the classical localized manner (Kharchenko, Tolstorukov et al. 2008) as well as broadly across the mitochondrial genome. As a result, metrics for evaluating nuclear transcription factors are not well-suited for analysis of TFAM binding data. We therefore examined the fraction of sequencing reads in our libraries mapping to the mitochondria as a proxy for the enrichment of TFAM binding events. Strikingly, between 30% and 75% of TFAM ChIP-seq reads mapped to the mitochondrial genome, while 0.4 – 1.9% of reads mapped to the mitochondrial genome in the input samples, indicating that our TFAM ChIP-seq datasets are indeed highly enriched for TFAM binding events **(Figure 3.6B)**. We note that 75% ChIP enrichment is extremely high (in fact, practically unprecedented) for any transcription factor dataset (Landt, Marinov et al. 2012), thus underscoring the high experimental quality of our datasets.

Because partial copies of the mitochondrial genome are also present in the nuclear genome, not all reads originating from mtDNA can be mapped uniquely. Therefore, we characterized TFAM binding to mtDNA and to the nuclear genome separately; computational work was performed by Georgi K. Marinov. Sequencing reads were

trimmed from 38 to 36 bp and then mapped to either the female set of human chromosomes, excluding the Y chromosome and all random chromosomes and haplotypes, or, separately, to the mitochondrial genome alone **(Figure 3.6A)**. For a standard nuclear transcription factor, this approach may cause some reads originating from the nuclear genome to artificially map to the mitochondrial genome. However, given that TFAM is known to bind to the mitochondrial genome and the extremely high enrichment for TFAM binding to mtDNA in our TFAM ChIP-seq libraries, this should not be a significant confounding factor. The hg19 version of the human genome was used as the reference template and Bowtie 0.12.7 (Langmead, Trapnell et al. 2009) was used for read alignment. We restricted our analysis to reads mapping perfectly without any mismatches to further increase mapping accuracy.

i.   *TFAM coats the mitochondrial genome*

As discussed previously, TFAM has not only been proposed to bind specifically to well-defined binding sites in the D-loop, but has also been suggested to play a nonspecific packaging role in the nucleoid that is essential for mtDNA integrity. However, little is known about the pattern of non-specific binding of TFAM to the mitochondrial genome. Localized binding at the D-loop and diffuse binding across the rest of the genome are expected to result in distinct ChIP-seq signal profiles. Localized, "point-source" binding to DNA results in an asymmetric distribution of reads mapping to the forward and reverse strand around the binding site of the protein (Kharchenko,

Tolstorukov et al. 2008, Pepke, Wold et al. 2009), while diffuse binding does not produce such strand asymmetry.

To characterize TFAM binding to mtDNA, we examined the forward and reverse strand read distribution after mapping TFAM ChIP-seq and input library reads to the mitochondrial genome. Strikingly, we did not observe regions of obvious enrichment and strand asymmetry in the D-loop; in particular, we did not see specific binding at the predicted HSP1 and LSP sites. On the whole, the TFAM ChIP-seq signal was broadly distributed over the whole mitochondrial chromosome, and while coverage was not perfectly uniform, the amplitude of the non-uniformity was not significant and the signal profile closely tracked that of the input sample **(Figure 3.7)**. The low level of non-uniformity likely results from sequencing biases, which has been documented to skew coverage (Dohm, Lottaz et al. 2008, Ross, Russ et al. 2013). Because our libraries were carefully size-selected for fragments in the 200 bp range, discrete TFAM binding sites would be expected to yield discrete signal localizations. Therefore, we interpret these results as evidence for the uniform coating of the whole mitochondrial genome by TFAM. We observed one region of apparent localized enrichment exhibiting strand asymmetry in the ND2 ORF near the O_L, which is discussed below.

To further verify our results, we carried out ChIP-seq against TFAM with a second TFAM monoclonal antibody, 20F8A9. We obtained similar results **(Figure 3.8A)** and found significant correlation between the 20F8A9 dataset and the three datasets obtained from the 20G2C12 antibody datasets (**Figure 3.8B**, $p < 0.0001$). Interestingly,

the correlation coefficient between ChIPs and between inputs was generally higher than that between a ChIP and its corresponding input, suggesting that while there is high repudibility between like datasets, there is less uniformity when the data is compared against its control.

ii.   TFAM localizes specifically to a locus on the mitochondrial genome

Attempts at utilizing available peak-calling software such as MACS and ERANGE were unsuccessful. In the case of ERANGE, no peaks were called, and with MACS, while peaks were called, they tiled the genome and visual inspection revealed that they were nonsignificant **(Figure 3.9)**. We attribute this to the largely uniform coverage of the genome by TFAM, resulting in signal ubiquitously. Because this type of coverage is unprecedented in the nuclear genome, software is not designed to deconvolute in a situation such as this and peaks cannot be called computationally.

We did visually inspect the genome for evidence of strong, specific binding sites which were obvious above the general nonspecific binding observed across the genome. A single site, centered at 5175 bp, displayed the morphology typical of single binding events. Prediction of curvature is based on the presence and localization of $A(n)$ and $T(n)$ tracts, where $n \geq 3$. A curvature score is determined based on the size and location of these tracts relative to each other; in essence, the longer the tract and the closer to 10.5bp (1 DNA turn) away from the previous tract, the more likely to cause curvature. Previous

predictions and experiments have suggested that the last half of the ND2 ORF has exhibits sequence-dependent DNA curvature (Welter, Dooley et al. 1989, Gadaleta, D'Elia et al. 1996), a characteristic of many prokaryotic origins of replication (Mukherjee, Patel et al. 1985, Deb, DeLucia et al. 1986). It is thought that such curvature acts as a "landmark" for binding of transcription-related proteins. However, computational analysis of the region of specific TFAM binding predicted a curvature score of zero within 200bp of the 5175 bp site, as compared to the threshold value of 1.5 considered significant for intrinsic curvature (Plaskon and Wartell 1987). Although EMSA was utilized to characterize curvature of the region upstream of O_L in human mtDNA, the area surveyed stopped 50 bp short of this binding site (Welter, Dooley et al. 1989). Thus, there is no evidence for sequence-dependent DNA curvature in the vicinity of the peak.

On the other hand, TFAM has been localized just upstream of the O_L in rat mitochondria. In fact, the locus, discovered by affinity pull-down from mitochondrial protein lysate using the specific mtDNA sequence as probe, resides 520 bp upstream of the rat O_L. In comparison, our identified site is 546bp upstream, suggesting that they may be congruent.

iii.   *No evidence for TFAM binding to nuclear genome*

Previous studies have suggested that TFAM can be found in the nucleus and that it modulates the transcription of nuclear genes. In rat neonatal cardiac myocytes, TFAM was found to bind to the promoter of SERCA2, the homolog of human sarco-endoplasmic reticulum calcium-ATPase 2 (ATP2A2), and was implicated in regulating its transcription (Watanabe, Arai et al. 2011). Given the extremely high degree of TFAM binding enrichment in our datasets, any robust nuclear TFAM binding events should be readily detectable. To analyze nuclear binding, we excluded all sequencing reads mapping to the mitochondrial genome and used the resulting set of reads to identify putative TFAM binding sites. We first looked for significant global read clustering using cross-correlation between reads mapping to the forward and the reverse DNA strands (Kharchenko, Tolstorukov et al. 2008, Landt, Marinov et al. 2012). Cross-correlation plots for input samples and for TFAM ChIP-seq datasets were indistinguishable from each other **(Figure 3.10A-B)**. Next, we called putative TFAM binding sites using MACS2 (Zhang, Liu et al. 2008). Using default settings (corresponding to a q-value cut-off of 10^{-2}) we identified 72, 137 and 153 sites respectively for the three replicates generated with antibody 20G2C12, and a single site for the 20F8A9 antibody. However, manual inspection of each of the identified sites revealed that all were likely to represent artifacts, mostly associated with repetitive DNA sequences, as none had the expected strand asymmetry of read distribution around a binding site. Instead, the two strand profiles at each site were identical (summarized in **Figure 3.10D**, with the classic nuclear transcription factor NRSF shown for comparison in **Figure 3.10C**), and numerous

unmappable regions and repetitive elements were present in the immediate vicinity of many of the called sites. Inspection of the ATP2A2 gene revealed no TFAM enrichment neither in the promoter region nor anywhere else in the neighborhood of the gene **(Figure 3.10E)**. Furthermore, we do not detect nuclear localization of TFAM in our cells via immunocytochemistry (Fig. 1C). Therefore, in HeLa cells under normal growth conditions, we find no evidence for specific binding of TFAM to nuclear target genes.

IV.   <u>Discussion</u>

ChIP-seq was originally developed to capture the in vivo interactions of nuclear transcription factors with DNA. However, many interesting mitochondrial transcription factors exist which are vital for cellular and organismal viability. Study of these interactions required optimization of the standard ChIP-seq protocol, due to the localization of mtDNA in the mitochondrial and because of the physical differences – nuclear DNA is long and linear, while mtDNA is short and circular. We show here successful recovery of mtDNA in DNA prepared via our modified ChIP-seq protocol, and localization of 30-75% of reads to the mitochondrial genome, highlighting the high experimental quality of our methods.

Previous in vitro studies have suggested that TFAM binds specifically to LSP and HSP1, and that it may also bind nonspecifically in a phased manner. Furthermore, evidence has been presented for its nuclear localization and action as a canonical nuclear

transcription factor in rat neonatal cardiac myocytes. However, no direct genome-wide measurements of TFAM binding have been previously reported. Our TFAM ChIP-seq data reveal very high enrichment for reads mapping to the mitochondrial genome, but a binding pattern that largely mirrors the read distribution observed in the input DNA, suggesting broad, non-specific binding to mitochondrial genome. This pattern is highly reproducible, indicating that the average population-wide state of TFAM-mtDNA interactions is stable. We found no correlation between irregularities in TFAM signal distribution and characteristics of the mitochondrial genome such as GC content. Thus, we conclude that TFAM binds to the mitochondrial genome nonspecifically and without bias when cells are grown under typical culture conditions. Although we do not observe the synchronized phased binding seen in in-vitro studies, we cannot rule out a model where individual mtDNAs have such a pattern of binding initiating from a non-universal nucleation site.

Strikingly, we did not observe localized enrichment of binding at the known LSP and HSP1 TFAM binding sites. Peak patterns mirrored that of the input in these regions, and no ChIP-seq peaks displaying the canonical strand asymmetry in read distribution were observed. This finding can be explained by a model in which the interaction of TFAM with the LSP and HSP1 binding sites is relatively transient and infrequent compared to a more stable non-specific association with the genome in its packaging state.

We did detect one site in the genome exhibiting the characteristics of a specific, localized ChIP-seq peak, centered at 5175 bp in the ND2 ORF. The localized nature of the ChIP signal at this site suggests higher occupancy of TFAM. This peak localizes to 546 bp upstream of the O_L. Strikingly, TFAM has previously been localized 520 bp upstream of the O_L of rat mtDNA (Gadaleta, D'Elia et al. 1996, Cingolani, Capaccio et al. 1997, Pierro, Capaccio et al. 1999). We found no sequence similarity between the rat and human sites, and in general this region of the mtDNA genome shows low homology between the two species. Further work will be required to understand the significance of this putative TFAM binding site.

Finally, analysis of all datasets for TFAM binding to the nuclear genome yielded no hits distinguishable from common ChIP-seq artifacts. Although Watanabe et al. observed regulation of the SERCA2 gene in rat myocytes, we did not detect TFAM binding at the promoter of its ortholog in humans. Previous studies have shown nuclear localization of TFAM in rat hepatoma cells (Dong, Ghoshal et al. 2002), as well as an alternate isoform of TFAM in mouse testis nuclei (Larsson, Garman et al. 1996). We have thus far been unable to detect nuclear TFAM localization in HeLa cells **(Figure 3.2A)**, suggesting that nuclear localization and transcriptional regulation may be cell type or perhaps species-dependent. ChIP-seq in different cell lines may be able to detect such nuclear interactions.

We demonstrate here the first in vivo ChIP-seq analysis of TFAM binding to the mitochondrial genome. Aside from generalized, largely non-specific binding across the

mitochondrial genome, we detected a putative specific binding site upstream of the origin of light strand replication. We do not observe the expected binding at the known HSP1 and LSP sites, nor did we identify any nuclear binding sites. An area that remains to be explored is the dynamic nature of TFAM-DNA interactions with respect to both the nuclear and mitochondrial genomes. ChIP-chip on the yeast mitochondrial genome has shown that metabolic changes can lead to differential binding of the yeast TFAM homolog, Abf2p (Kucej, Kucejova et al. 2008). It is possible that such remodeling also occurs in the mammalian system, and further studies will provide insight into the dynamic nature of the mtDNA-protein interactions within the nucleoid that serve to protect its integrity.

V.   **Methods**

Cell growth and treatment

HeLaS3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Invitrogen #11995) containing 10% bovine serum (Invitrogen #16170), penicillin and streptomycin, and additional L-glutamine (2mM). Cells were fed 24 hours before harvest for ChIP-seq, which was performed at 80-90% confluency.

Antibody Production and characterization

Antibodies were produced by the Caltech Monoclonal Antibody Facility and raised against the full-length TFAM protein in mouse. Immunoprecipitation with 20G2C12 and 20F8A9 TFAM antibodies and Myc antibody (Santa Cruz #sc-40) was performed according to established protocols using M-280 sheep anti-mouse Dynabeads (Invitrogen #11201D). Immunoblotting of IP products was performed using a monoclonal TFAM 18G102B2E11 antibody, also custom generated, at 1:2000, with goat anti-mouse HRP antibody (1:10,000, Jackson ImmunoResearch #115-056-003). Immunoblotting of HeLa whole cell lysate with 20G2C12 was performed at a 1:200 dilution and with goat anti-mouse HRP antibody.

Additional immunoblotting antibody concentrations were as follows: b-actin (1:500, Imgenex #IMG-5142A); H2B (1:500, Imgenex #IMG-359); Mfn2 (1:1000, Sigma #128K4840), Opa1 (1:2000, in-house), GAPDH (1:2000, Imgenex #IMG5143A). Secondary antibodies, all at 1:10,000, included: goat anti-mouse (Jackson #115-056-003), and goat anti-rabbit (Jackson #111-035-003), and protein A/G.

Isolation of Mitochondria

Isolation of mitochondrial was performed using "nitrogen bomb" method, per Nickie Chan. Cells were harvested by 1000rpm centrifugation for 10min, then subsequently washed with ice-cold PBS and pelleted at 1000rpm again. Cell pellet was then resuspended in isolation buffer (220mM mannitol, 70mM sucrose, 10mM HEPES-KOH ph7.4, 1mM EGTA) and HALT protease inhibitor (Pierce #87786), and transferred to Parr 45ml nitrogen bomb (Parr #4639). After cell lysate was relased from bomb, it was centrifuged at 600g x 5min and the supernatant was collected. The pellet was dounced gently to break down remaining whole cells in the pellet, and recentrifuged. Supernatants were then pooled, and centrifuged at 650g x 5min repeatedly until no pellet was visibile in order to rid the lysate of any remaining whole cells and debris. The mitochondrondria were then pelleted at 10,000g x 10min and resuspended in RIPA buffer for future experiments.

Visualizing DNA sonication efficiency

Cells were formaldehyde crosslinked and quenched, and then sonicated using QSonica Sonicator 4000. Decrosslinking was performed by 65C incubation overnight, and resultant DNA was purified via phenol-chloroform extraction followed by PCR purification (Qiagen #28104), substituting buffer PM for buffer PB in the classic protocol. 4, 8, or 16ul of the 100ul lysate was then loaded onto a 2.5% agarose gel for visualization with EtBr incorporation.

Quantitative PCR for human DNA

Four sets of primers, LSP1, LSP2, HSP3, and HSP4, were utilized for mtDNA quantification. One set of primers, iB5s, was used for nuclear DNA quantificiation.

Primer ID	Sequence
LSP1F	GATCACAGGTCTATCACCCTATTAACCACT
LSP1R	CAGCGTCTCGCAATGCTATCGC
LSP2F	CGGGAGCTCTCCATGCATTTGGTATTTTC
LSP2R	CAAAGACAGATACTGCGACATAGGGTGC
HSP3F	CACTGAAAATGTTTAGACGGGCTCACATCA
HSP3R	CTGGAACGGGGATGCTTGCATG
HSP4F	CACATCACCCCATAAACAAATAGGTTTGGTCCTAGCC
HSP4R	TTTAGAGGGTGAACTCACTGGAACGG
iB5sF	AGACTCCGACCCTTCCTGAC
iB5sR	CCTCCGAGCTTGGGATAAAG

qPCR was performed on Bio-Rad CFX-96 machine with using Brilliant SYBR Green 2X Master Mix, Agilent #929548 and the following cycling conditions: 10min@95C,

[30sec@95C, 1min@60.2C, 1min@72C]x40.

Immunocytochemistry

HeLa cells cultured as described above were plated onto poly-lysine coated glass coverslips 48 hours prior to fixation in formaldehyde and permeabilization with 0.1% Triton X-100. For colocalization of TFAM to mitochondria, 20G2C12 or 20F8A9 antibodies were used at 1:10 in conjunction with PPIF at 1:200 (ProteinTech #18466-1-AP). Secondary antibodies were goat anti-mouse AF488 (1:500, Invitrogen #A11001) and donkey anti-rabbit AF546 (1:500, Invitrogen #A10040). Cells were also stained with DAPI to visualize nuclei. Immunocytochemistry to visualize colocalization of mitochondrial nucleoids and TFAM was performed sequentially due to both antibodies being raised in mouse. Sequential immunostaining yielded no background fluorescence due to cross-antibody reactivity (data not shown). Order was as follows: anti-TFAM antibody (1:10); goat anti-mouse AF488 (1:500, Invitrogen #A11001); anti-DNA antibody (1:25, Millipore #CBL186); goat anti-mouse AF555 (1:500, Invitrogen #A21426), DAPI. Images were acquired with a Zeiss LSM 710 confocal microscope with PlanApochromat 63X/1.4 oil objective. Z-stack acquisitions were converted to maximum z-projections using ImageJ software.

Chromatin immunoprecipitation and sequencing

ChIP experiments and preparation of DNA for sequencing were performed following standard procedures (Johnson, Mortazavi et al. 2007) with some modifications. Cells were fixed for 10min at RT in 1% formaldehyde, harvested using a cell scraper, washed once in ice-cold PBS, and resuspended in RIPA buffer with protease inhibitor. The sample was then sonicated using a 3.2mm microtip (QSonica Sonicator 4000) at 30s on/30s off intervals and 40% amplitude for 180min while in a -30°C 3:1 isopropanol and water bath containing dry ice. Subsequent steps were performed as per the standard protocol. DNA was size-selected during library building to an average fragment size of 200bp. Libraries were sequenced using Illumina GAIIx and Illumina HiSeq 2000. Sequencing data is available under GEO accession record GSE48176.

Sequencing data processing and analysis

Sequencing reads were trimmed down to 36 bp and then mapped against either the female set of human chromosomes (excluding the Y chromosome and all random chromosomes and haplotypes) or the mitochondrial genome alone, using the hg19 version of the human genome as a reference. Bowtie 0.12.7 (Langmead, Trapnell et al. 2009) was used for aligning reads, not allowing for any mismatches between the reads and the reference. ChIP-seq peaks were called using MACS2 (Zhang, Liu et al. 2008) with default settings except for the mfold parameter, which was lowered to (2,30). Circos plots

were generated using Circos version 0.60 (Krzywinski, Schein et al. 2009). Additional data processing was carried out using custom-written python scripts. ENCODE data was downloaded from the UCSC browser (http://hgdownload-test.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeHaibTfbs) and its use here complies with its terms of usage. Pearson correlation coefficient, t-test, and p values were calculated using embedded and custom Microsoft Excel functions.

VI. Figure Legends

Figure 3.1. Localizing mitochondria in cellular fractions. A) Mitochondrial prep [1], RIPA-buffer solubilized pellet [2], and Farnham buffer solubilized supernatant [3] were immunoblotted for actin, H2B, Mfn2, TFAM, Opa1, and GAPDH. TFAM localized to [1] and [2], but not [3]. B) Antibodies H101A6 and 20G2C12 were tested for ability to immunoprecipitate TFAM using either streptavidin- or sheep anti-mouse-conjugated Dynabeads. Western is anti-TFAM with protein A/G secondary. Two types of lysate were used: Farnham supernatant ("Farnham") and pellet resuspended in RIPA buffer. Conditions as follows: [1] 20G2C12 RIPA; [2] 20G2C12 Farnham; [3] H101A6 RIPA; [4] H101A6 RIPA; [5] Myc Farnham; [6] 20G2C12 RIPA; [7] 20G2C12 Farnham; [8] H101A6 RIPA; [9] H101A6 RIPA; [10] Myc RIPA; [11] Myc Farnham; [12] RIPA lysate alone. 1-5 are using streptavidin-conjugated beads, and 6-11 are using sheep-anti-mouse conjugated beads. Red arrow indicates TFAM band. No TFAM was capture by any condition with streptavidin beads. 20G2C12 antibodies captures the most TFAM in RIPA lysate, with some captured in Farnham lysate. H101A6 does not capture any TFAM, nor does Myc antibody. RIPA lysate shown as control. Note that ~26 kDa band is light chain, and ~50 kDa band is heavy chain. C) Immunoprecipitation of TFAM from noncrosslinked whole cell lysate. HeLa cell lysate was applied to sheep anti-mouse Dynabeads conjugated to anti-Myc [1], 20G2C12 TFAM antibody [2], 20F8A9 TFAM antibody [3], or a 50/50 mixture of 20G2C12 and 20F8A9 TFAM antibodies [4]. The labeled bands are: I. Antibody heavy chain; II. antibody light chain; III. TFAM. D) 20G2C12 antibody successfully immunoprecipitates TFAM that has been crosslinked to

DNA. Cells were fixed by formaldehyde, sonicated, and then incubated with sheep-anti-mouse antibodies conjugated to 20G2C12 antibody. After immunoprecipitation, decrosslinking of DNA and protein was performed and protein lysate was assayed by anti-TFAM immunoblotting. [1] Immunoprecipitated lysate; [2] immunoprecipitated lysate with 2x beads; [3] supernatant of immunoprecipitation (after removal of beads); [4] supernatant of immunoprecipitation, 2x beads (after removal of beads). Same % of each sample is loaded in each lane. Not all TFAM from lysate is captured; use of twice the concentration of beads yielded poorer results. E) Western blot of whole cell HeLa lysate using the 20G2C12 antibody detects only a ~23kDa band.

Figure 3.2. Characterization of TFAM monoclonal antibodies. A) Immunocytochemistry showing TFAM localization in HeLa cells. Mitochondria were identified by PPIF staining; B) mtDNA was identified by anti-DNA staining. There was no evidence for nuclear localization of TFAM using either antibody. C) sequential immunostaining of cells with TFAM and DNA antibodies yielded no cross-reactivity when either the DNA or TFAM antibody was not used; respective secondary antibodies were utilized.

Figure 3.3. Optimization of DNA sonication. HeLa whole cell lysate was sonicated using either 10amp or 20amp power, and either 2,4, 6, or 30 replicates of 30s on/30s off

sonication. Only 10/20amp x 30 was successful in yielding DNA near the size desired, indicating that a larger number of cycles was required. B) 30am x 45 and 40amp x 45 are able to successfully shear DNA to an average length of 150-200bp in small samples of cell lysate. In larger samples C), 45amp x 90 was sufficient to effect shearing. 50amp x 90 yielded no benefit over 40amp, so the lower power value was selected due to it being less damaging to sample.

Figure 3.4. ChIP-seq against TFAM yields high ratio of mtDNA:nDNA. DNA was collected either before immunoprecipitation (Sonicated DNA), or after immunoprecipitation (TFAM ChIP), and was assayed by qPCR using four different primer sets for mtDNA content. While mtDNA:nDNA ratio was 190-300 in pre-immunoprecipitated cells, it was 65,700-105000/genome after, indicating a very high enrichment for mtDNA, as expected with a successful ChIP.

Figure 3.5. Bioanalyzer of TFAM ChIP-seq shows majority of DNA centered at ~200bp. Analysis of ChIP-seq DNA sent for sequencing shows that the majority was sheared to ~ 222bp.

Figure 3.6. ChIP-seq analysis of genome-wide TFAM binding. (A) Overview of computational processing of data. Reads were trimmed to 36 bp and then either mapped against the mitochondrial genome (ChrM), or the complete hg19 version of the genome. After removing multireads and alignments to the mitochondrial genome, peaks in the nuclear genome were called using MACS2. (B) The proportion of sequencing reads mapping to chrM in ChIP and input datasets. All replicates of the ChIP-seq resulted in at least 30% of reads mapping to the mitochondrial genome, much greater than the 0.4-1.9% of reads mapping to mtDNA in the input datasets. Replicates 1-3 were performed using the 20G2C12 antibody, while Replicate 4 was performed using the 20F8A9 antibody.

Figure 3.7. Coating of the mitochondrial genome by TFAM in HeLa cells. Circos plot of plus strand and minus strand TFAM ChIP-seq and input read density signal over chrM. (A, E) Annotation of protein coding (green on forward/heavy strand, red on reverse/light strand), ribosomal RNA (blue) and tRNA (blue on forward/heavy strand, grey on reverse/light strand) transcripts. (B) D-loop (black), LSP promoter (large red tile), known LSP TFAM binding site (small red tile), HSP promoter (large blue tile), known HSP1 TFAM binding site (small blue tile), and origins of heavy strand replication (Ori-b, orange tile; O_H, yellow tile). (C) TFAM ChIP-seq signal on forward (red) and reverse (blue) strands. (D) Input signal on forward (red) and reverse (blue) strands. (F) Origin of light strand replication (yellow tile). Note that the input signal is exaggerated 60-fold relative to the ChIP-seq signal in order to visualize coverage irregularities. The signal

from the TFAM ChIP-seq largely follows that of the input, indicating generalized binding across the mitochondrial genome.

Figure 3.8: Comparison of profiles of TFAM binding to mitochondrial genome. (A) Circos plots of TFAM ChIP-seq experiments: (1) 20F8A9 antibody ChIP-Seq; (Shen-Li, O'Hagan et al.) 20G2C12 replicate 1; (Gerstein, Lu et al.) 20G2C12 replicate 2; (4) 20G2C12 replicate 3. Read profiles are very similar across replicates and antibodies. (B) R^2 coefficient of determination values for dataset pairs: rep to rep – TFAM ChIP pairs; rep to input – TFAM ChIP to corresponding input; input to input – input pairs.

Figure 3.9. MACS calls along the mitochondrial genome. MACS was utilized to call peaks on the mitochondrial genome (first replicate). 1) Called peaks tile the genome. 2) Plus-strand read distribution; 3) minus-strand read distribution.

Figure 3.10. Absence of TFAM binding to the nuclear genome. (A) Cross-correlation plot of input DNA computed over the nuclear genome. (B) Cross-correlation plot of TFAM ChIP-seq computed over the nuclear genome. (C) Distribution of ChIP-seq reads mapping to the plus and minus strand around called binding sites in a ChIP-seq dataset for the NRSF transcription factor (Schoenherr and Anderson 1995) [51] in HeLa cells,

generated by the ENCODE consortium. (D) Distribution of TFAM ChIP-seq reads mapping to the plus and minus strand around called binding sites indicates lack of real binding sites. (E) No ChIP-seq enrichment around the promoter of the SERCA2/ATP2A2 gene, previously suggested to be a TFAM target.
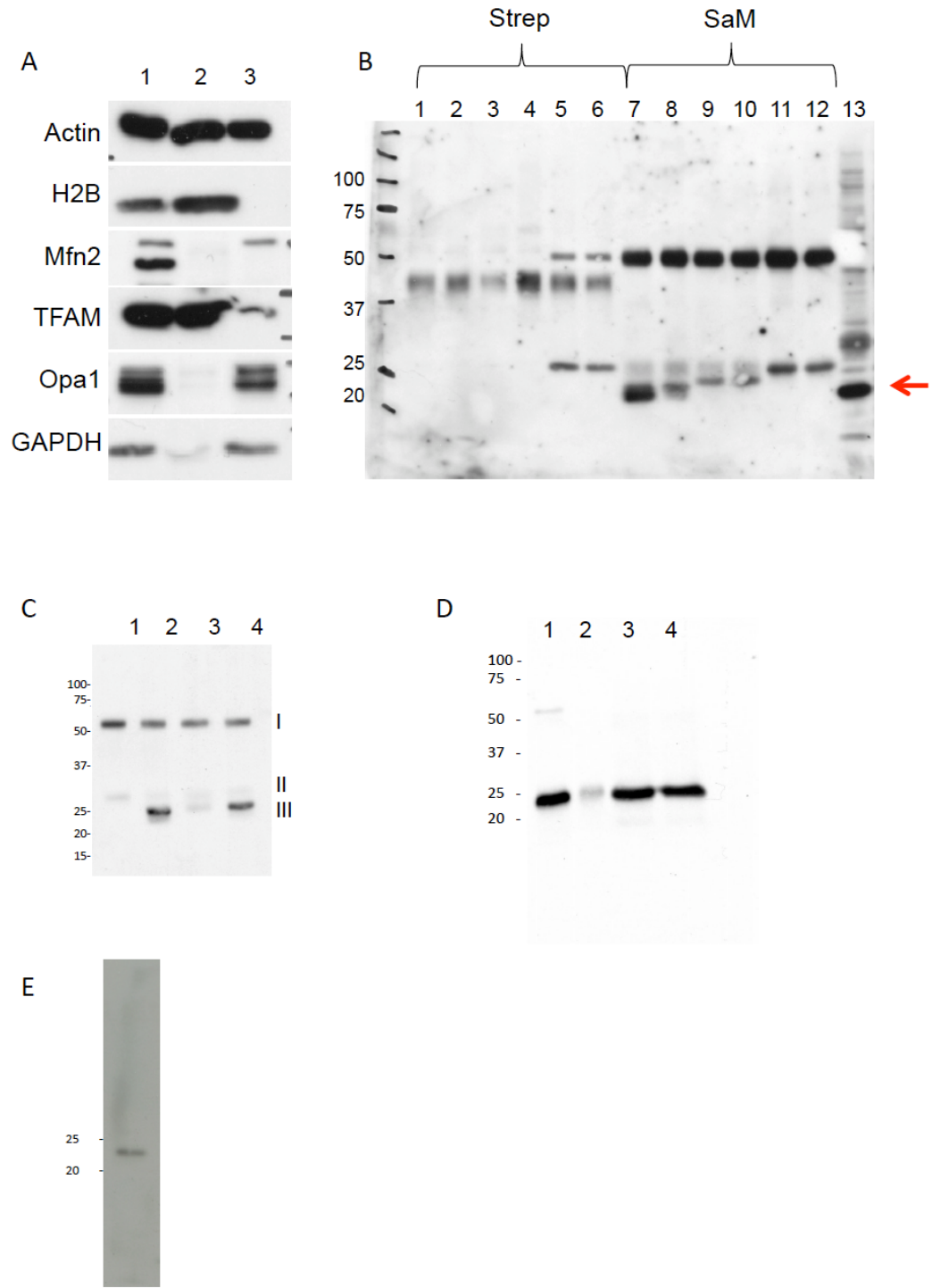
VII.   Figures
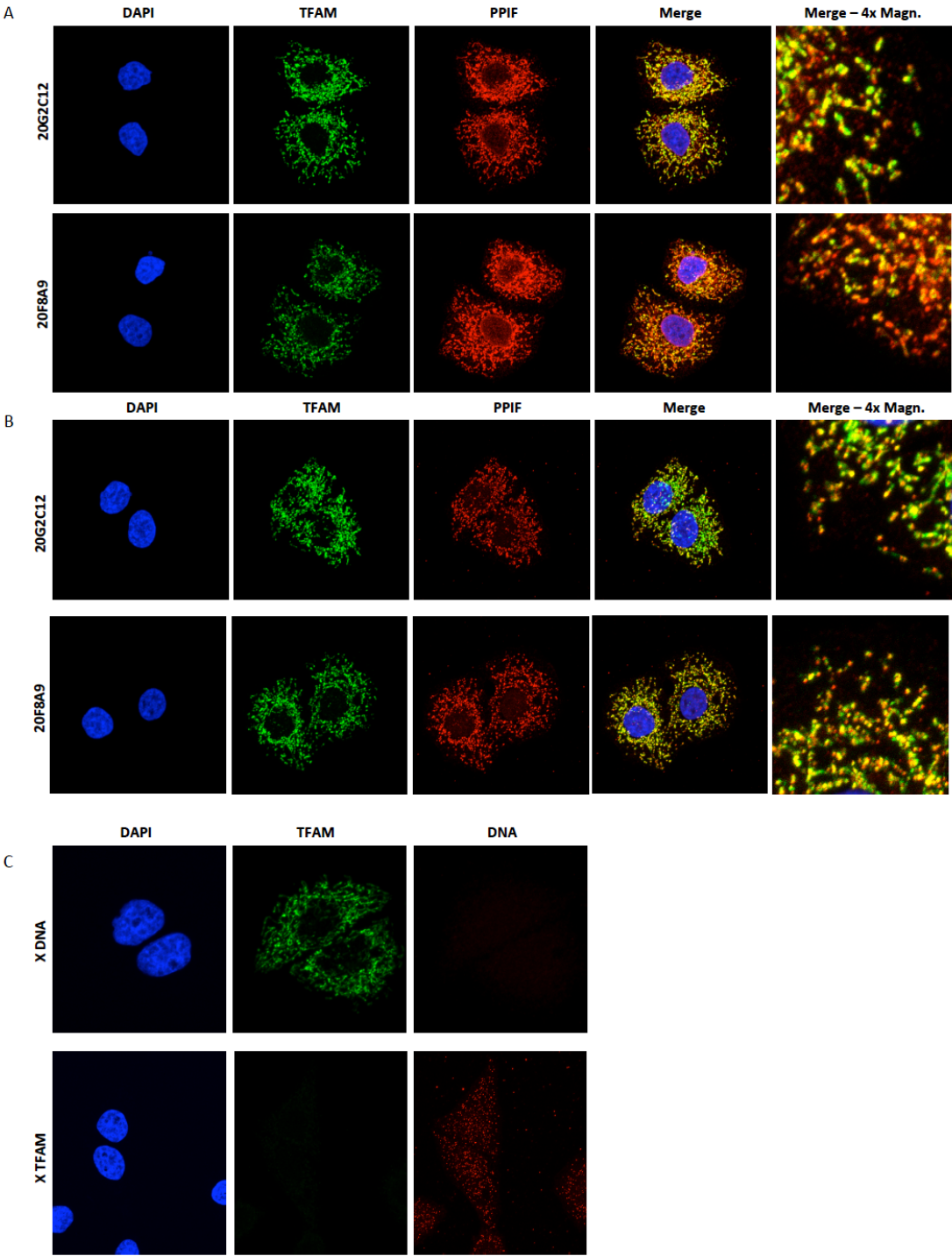
Figure 3.1.

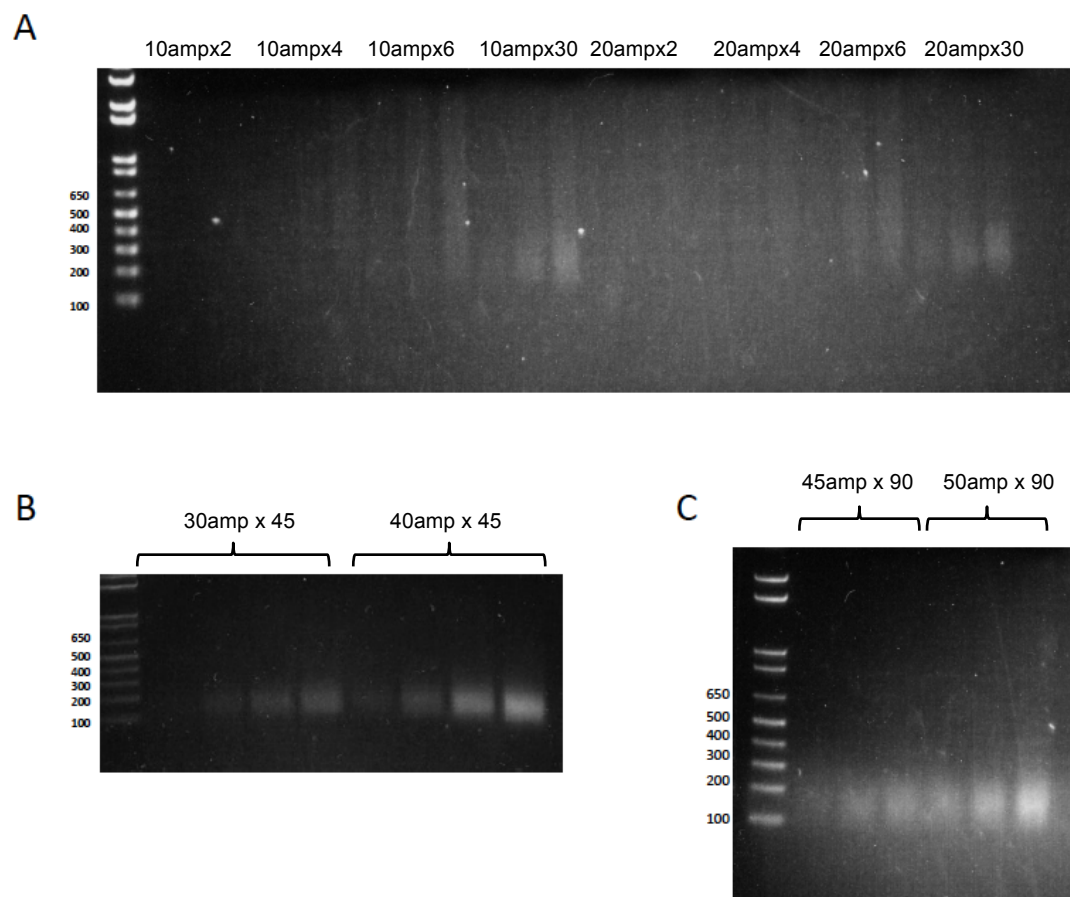**Figure 3.2.**

Figure 3.3.

A

10ampx2 10ampx4 10ampx6 10ampx30 20ampx2 20ampx4 20ampx6 20ampx30

650
500
400
300
200

100

B

30amp x 45   40amp x 45

650
500
400
300
200
100

C

45amp x 90   50amp x 90

650
500
400
300

200

100

Figure 3.4.

Sample	mtDNA:nDNA	Ratio
Sonicated DNA	LSP1:iB5s	200
Sonicated DNA	LSP2:iB5s	190
Sonicated DNA	HSP3:iB5s	250
Sonicated DNA	HSP4:iB5s	300
TFAM ChIP	LSP1:iB5s	83500
TFAM ChIP	LSP2:iB5s	65700
TFAM ChIP	HSP3:iB5s	75900
TFAM ChIP	HSP4:iB5s	105000

Figure 3.5.

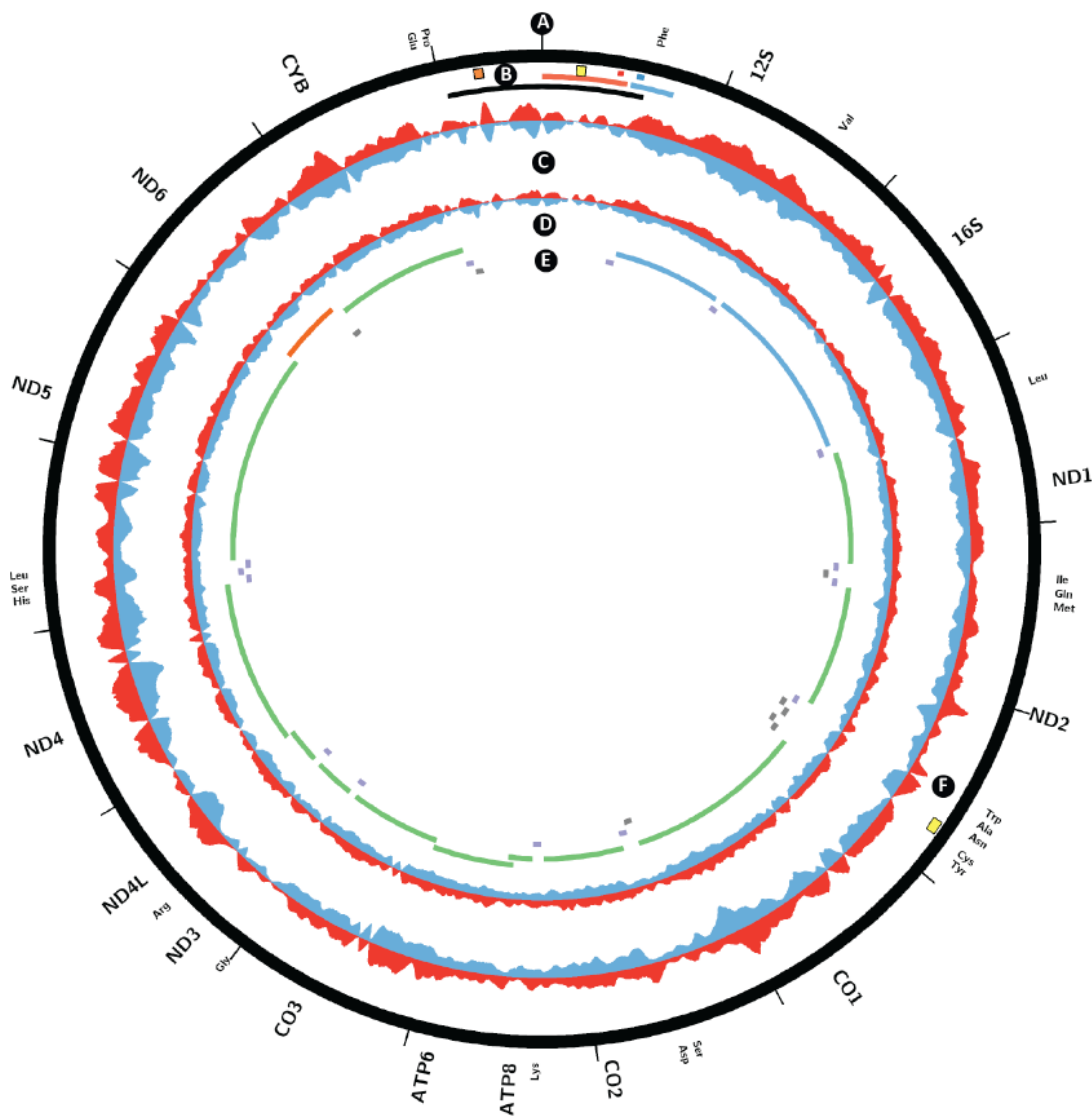**Figure 3.6.**

A

B

130
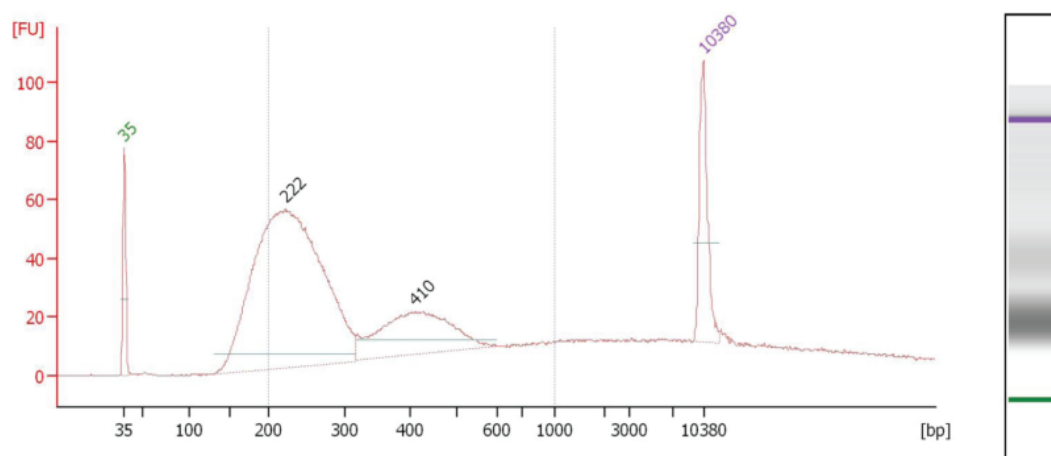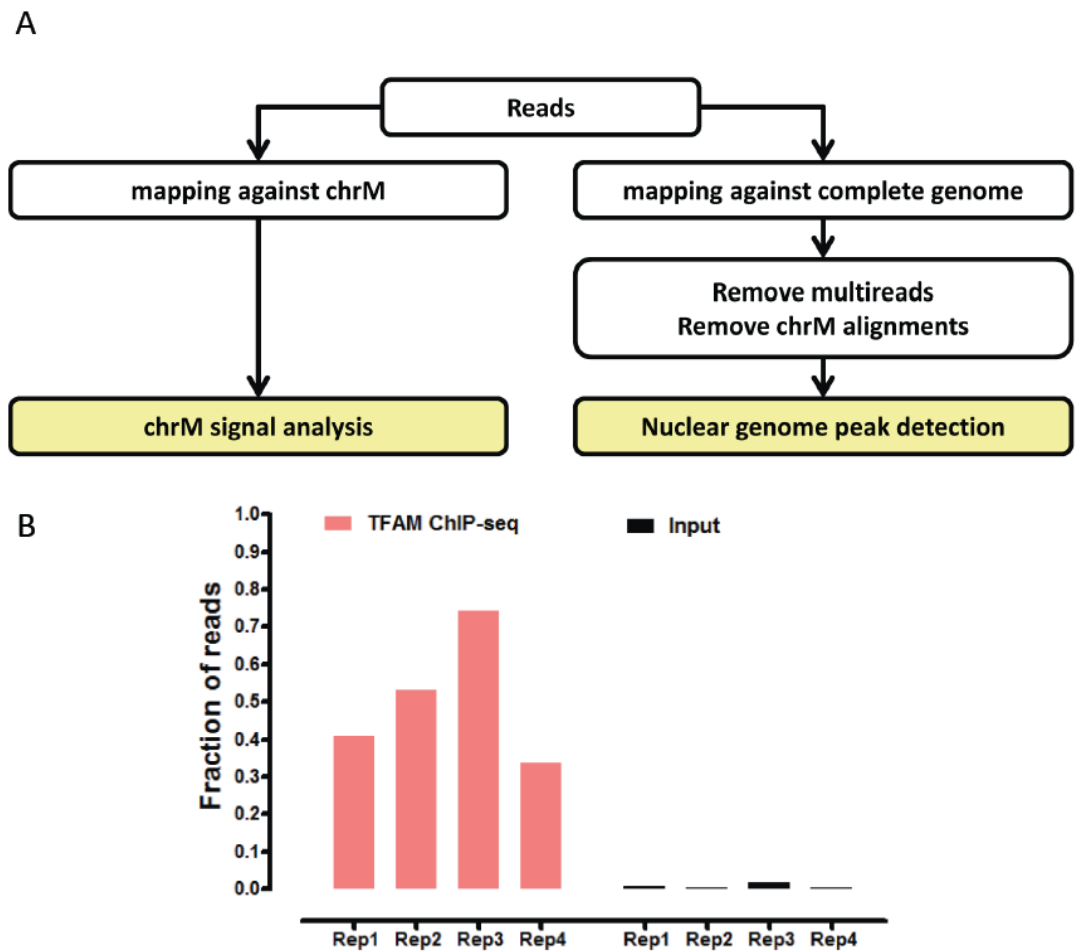
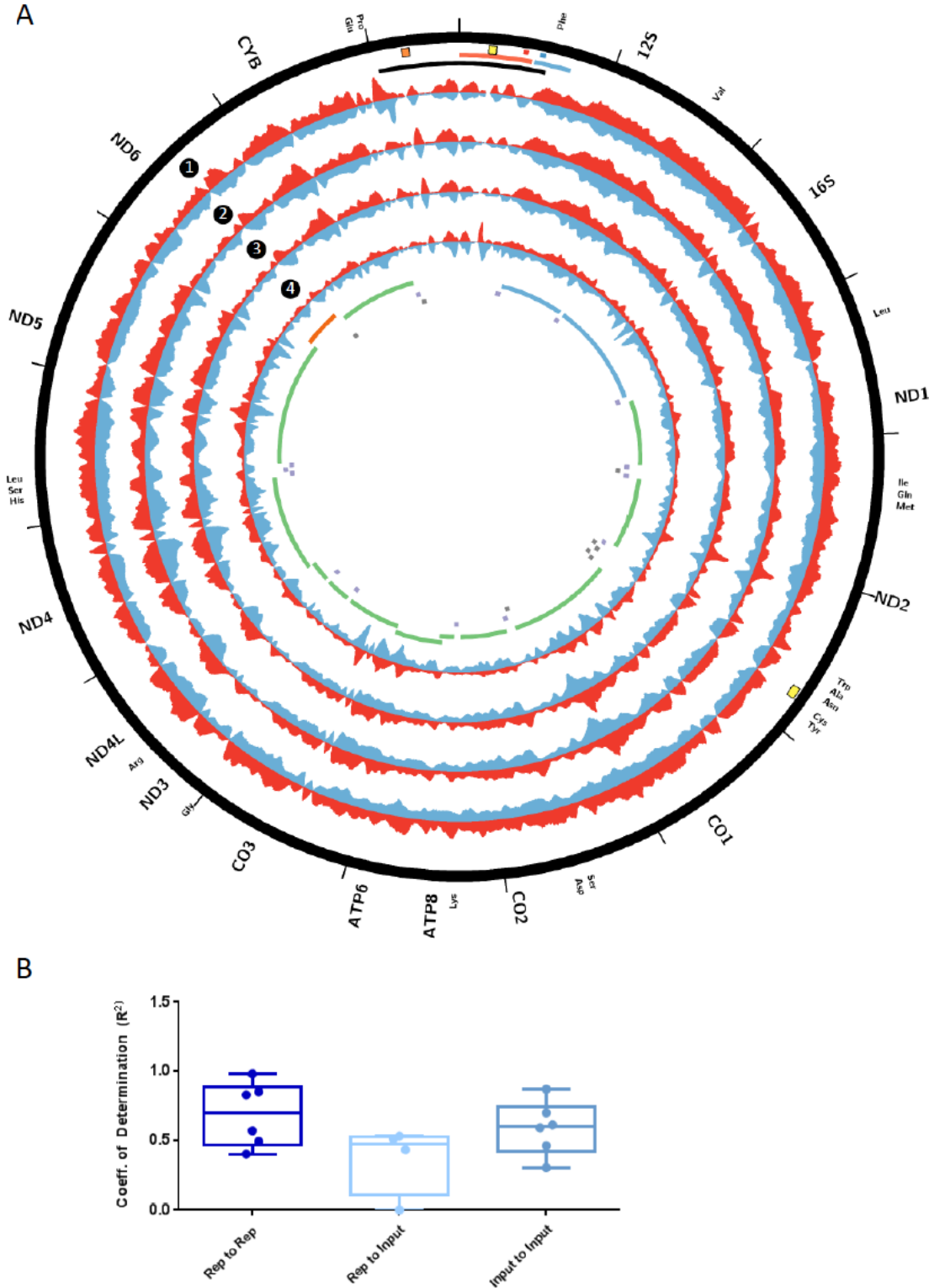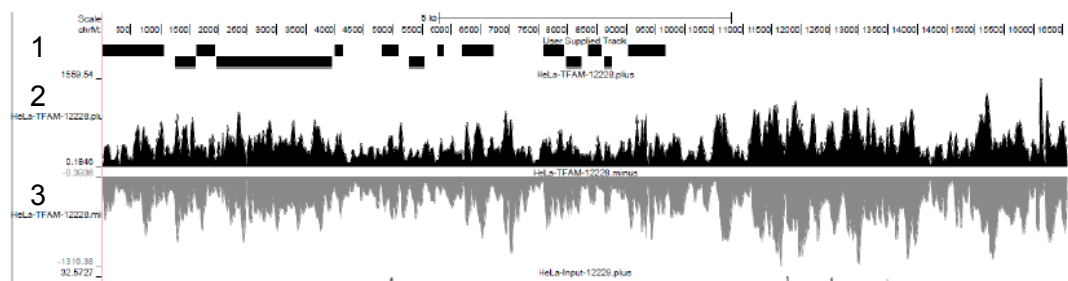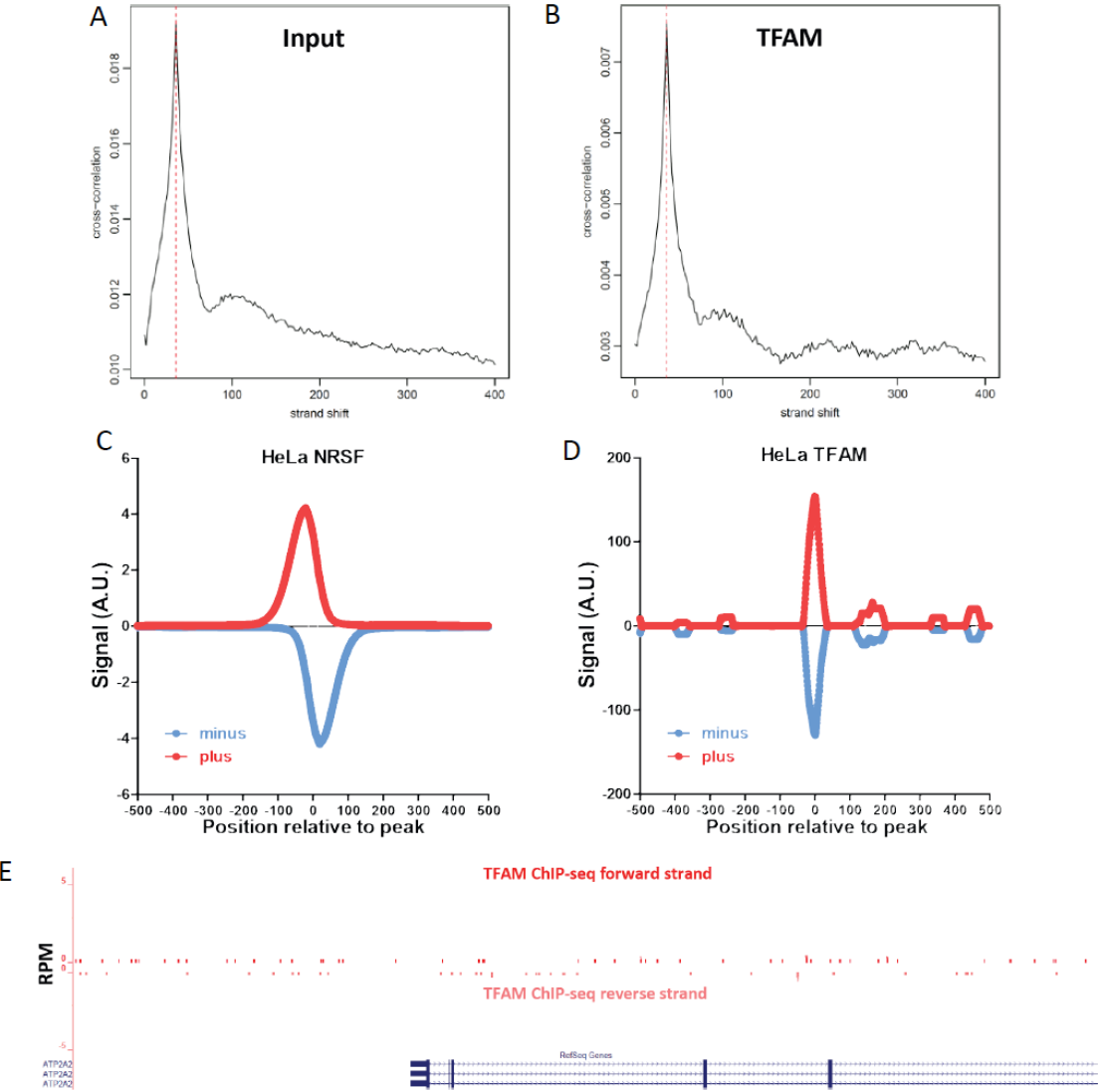**Figure 3.7.**

Figure 3.8.

Figure 3.9.

Figure 3.10.

BIBLIOGRAPHY

Alam, T. I., T. Kanki, T. Muta, K. Ukaji, Y. Abe, H. Nakayama, K. Takio, N. Hamasaki and D. Kang (2003). "Human mitochondrial DNA is packaged with TFAM." <u>Nucleic Acids Res</u> **31**(6): 1640-1645.

Bogenhagen, D. and D. A. Clayton (1974). "The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid." <u>J Biol Chem</u> **249**(24): 7991-7995.

Bogenhagen, D. F., D. Rousseau and S. Burke (2008). "The layered structure of human mitochondrial DNA nucleoids." <u>J Biol Chem</u> **283**(6): 3665-3675.

Chang, D. D. and D. A. Clayton (1984). "Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA." <u>Cell</u> **36**(3): 635-643.

Chang, D. D. and D. A. Clayton (1985). "Priming of human mitochondrial DNA replication occurs at the light-strand promoter." <u>Proc Natl Acad Sci U S A</u> **82**(2): 351-355.

Cingolani, G., L. Capaccio, D. D'Elia and G. Gadaleta (1997). "In organello footprinting analysis of rat mitochondrial DNA: protein interaction upstream of the Ori-L." <u>Biochem Biophys Res Commun</u> **231**(3): 856-860.

Corral-Debrinski, M., J. M. Shoffner, M. T. Lott and D. C. Wallace (1992). "Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease." <u>Mutat Res</u> **275**(3-6): 169-180.

Coskun, P., J. Wyrembak, S. E. Schriner, H. W. Chen, C. Marciniack, F. Laferla and D. C. Wallace (2012). "A mitochondrial etiology of Alzheimer and Parkinson disease." Biochim Biophys Acta **1820**(5): 553-564.

Coskun, P. E., M. F. Beal and D. C. Wallace (2004). "Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication." Proc Natl Acad Sci U S A **101**(29): 10726-10731.

Deb, S., A. L. DeLucia, A. Koff, S. Tsui and P. Tegtmeyer (1986). "The adenine-thymine domain of the simian virus 40 core origin directs DNA bending and coordinately regulates DNA replication." Mol Cell Biol **6**(12): 4578-4584.

Dohm, J. C., C. Lottaz, T. Borodina and H. Himmelbauer (2008). "Substantial biases in ultra-short read data sets from high-throughput DNA sequencing." Nucleic Acids Res **36**(16): e105.

Dong, X., K. Ghoshal, S. Majumder, S. P. Yadav and S. T. Jacob (2002). "Mitochondrial transcription factor A and its downstream targets are up-regulated in a rat hepatoma." J Biol Chem **277**(45): 43309-43318.

Ekstrand, M. I., M. Falkenberg, A. Rantanen, C. B. Park, M. Gaspari, K. Hultenby, P. Rustin, C. M. Gustafsson and N. G. Larsson (2004). "Mitochondrial transcription factor A regulates mtDNA copy number in mammals." Hum Mol Genet **13**(9): 935-944.

Fisher, R. P. and D. A. Clayton (1988). "Purification and characterization of human mitochondrial transcription factor 1." Mol Cell Biol **8**(8): 3496-3509.

Fisher, R. P., T. Lisowsky, M. A. Parisi and D. A. Clayton (1992). "DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein." J Biol Chem **267**(5): 3358-3367.

Fisher, R. P., M. A. Parisi and D. A. Clayton (1989). "Flexible recognition of rapidly evolving promoter sequences by mitochondrial transcription factor 1." Genes Dev **3**(12B): 2202-2217.

Gadaleta, G., D. D'Elia, L. Capaccio, C. Saccone and G. Pepe (1996). "Isolation of a 25-kDa protein binding to a curved DNA upstream the origin of the L strand replication in the rat mitochondrial genome." J Biol Chem **271**(23): 13537-13541.

Gauthier, B. R., A. Wiederkehr, M. Baquie, C. Dai, A. C. Powers, J. Kerr-Conte, F. Pattou, R. J. MacDonald, J. Ferrer and C. B. Wollheim (2009). "PDX1 deficiency causes mitochondrial dysfunction and defective insulin secretion through TFAM suppression." Cell Metab **10**(2): 110-118.

Gerstein, M. B., Z. J. Lu, E. L. Van Nostrand, C. Cheng, B. I. Arshinoff, T. Liu, K. Y. Yip, R. Robilotto, A. Rechtsteiner, K. Ikegami, P. Alves, A. Chateigner, M. Perry, M. Morris, R. K. Auerbach, X. Feng, J. Leng, A. Vielle, W. Niu, K. Rhrissorrakrai, A. Agarwal, R. P. Alexander, G. Barber, C. M. Brdlik, J. Brennan, J. J. Brouillet, A. Carr, M. S. Cheung, H. Clawson, S. Contrino, L. O. Dannenberg, A. F. Dernburg, A. Desai, L. Dick, A. C. Dose, J. Du, T. Egelhofer, S. Ercan, G. Euskirchen, B. Ewing, E. A. Feingold, R. Gassmann, P. J. Good, P. Green, F. Gullier, M. Gutwein, M. S. Guyer, L. Habegger, T. Han, J. G. Henikoff, S. R. Henz, A. Hinrichs, H. Holster, T. Hyman, A. L. Iniguez, J. Janette, M. Jensen, M.

Kato, W. J. Kent, E. Kephart, V. Khivansara, E. Khurana, J. K. Kim, P. Kolasinska-Zwierz, E. C. Lai, I. Latorre, A. Leahey, S. Lewis, P. Lloyd, L. Lochovsky, R. F. Lowdon, Y. Lubling, R. Lyne, M. MacCoss, S. D. Mackowiak, M. Mangone, S. McKay, D. Mecenas, G. Merrihew, D. M. Miller, 3rd, A. Muroyama, J. I. Murray, S. L. Ooi, H. Pham, T. Phippen, E. A. Preston, N. Rajewsky, G. Ratsch, H. Rosenbaum, J. Rozowsky, K. Rutherford, P. Ruzanov, M. Sarov, R. Sasidharan, A. Sboner, P. Scheid, E. Segal, H. Shin, C. Shou, F. J. Slack, C. Slightam, R. Smith, W. C. Spencer, E. O. Stinson, S. Taing, T. Takasaki, D. Vafeados, K. Voronina, G. Wang, N. L. Washington, C. M. Whittle, B. Wu, K. K. Yan, G. Zeller, Z. Zha, M. Zhong, X. Zhou, E. C. mod, J. Ahringer, S. Strome, K. C. Gunsalus, G. Micklem, X. S. Liu, V. Reinke, S. K. Kim, L. W. Hillier, S. Henikoff, F. Piano, M. Snyder, L. Stein, J. D. Lieb and R. H. Waterston (2010). "Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project." Science **330**(6012): 1775-1787.

Ghivizzani, S. C., C. S. Madsen, M. R. Nelen, C. V. Ammini and W. W. Hauswirth (1994). "In organello footprint analysis of human mitochondrial DNA: human mitochondrial transcription factor A interactions at the origin of replication." Mol Cell Biol **14**(12): 7717-7730.

Hayashi, Y., M. Yoshida, M. Yamato, T. Ide, Z. Wu, M. Ochi-Shindou, T. Kanki, D. Kang, K. Sunagawa, H. Tsutsui and H. Nakanishi (2008). "Reverse of age-dependent memory impairment and mitochondrial DNA damage in microglia by an overexpression of human mitochondrial transcription factor a in mice." J Neurosci **28**(34): 8624-8634.

Ide, T., H. Tsutsui, S. Hayashidani, D. Kang, N. Suematsu, K. Nakamura, H. Utsumi, N. Hamasaki and A. Takeshita (2001). "Mitochondrial DNA damage and dysfunction associated with oxidative stress in failing hearts after myocardial infarction." Circ Res **88**(5): 529-535.

Johnson, D. S., A. Mortazavi, R. M. Myers and B. Wold (2007). "Genome-wide mapping of in vivo protein-DNA interactions." Science **316**(5830): 1497-1502.

Kanki, T., K. Ohgaki, M. Gaspari, C. M. Gustafsson, A. Fukuoh, N. Sasaki, N. Hamasaki and D. Kang (2004). "Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA." Mol Cell Biol **24**(22): 9823-9834.

Karamanlidis, G., V. Bautista-Hernandez, F. Fynn-Thompson, P. Del Nido and R. Tian (2011). "Impaired mitochondrial biogenesis precedes heart failure in right ventricular hypertrophy in congenital heart disease." Circ Heart Fail **4**(6): 707-713.

Karamanlidis, G., L. Nascimben, G. S. Couper, P. S. Shekar, F. del Monte and R. Tian (2010). "Defective DNA replication impairs mitochondrial biogenesis in human failing hearts." Circ Res **106**(9): 1541-1548.

Kaufman, B. A., N. Durisic, J. M. Mativetsky, S. Costantino, M. A. Hancock, P. Grutter and E. A. Shoubridge (2007). "The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures." Mol Biol Cell **18**(9): 3225-3236.

Keeney, P. M., C. K. Quigley, L. D. Dunham, C. M. Papageorge, S. Iyer, R. R. Thomas, K. M. Schwarz, P. A. Trimmer, S. M. Khan, F. R. Portell, K. E. Bergquist and J. P. Bennett, Jr. (2009). "Mitochondrial gene therapy augments mitochondrial physiology in a Parkinson's disease cell model." Hum Gene Ther **20**(8): 897-907.

Kharchenko, P. V., M. Y. Tolstorukov and P. J. Park (2008). "Design and analysis of ChIP-seq experiments for DNA-binding proteins." Nat Biotechnol **26**(12): 1351-1359.

Krzywinski, M., J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones and M. A. Marra (2009). "Circos: an information aesthetic for comparative genomics." Genome Res **19**(9): 1639-1645.

Kucej, M., B. Kucejova, R. Subramanian, X. J. Chen and R. A. Butow (2008). "Mitochondrial nucleoids undergo remodeling in response to metabolic cues." J Cell Sci **121**(Pt 11): 1861-1868.

Landt, S. G., G. K. Marinov, A. Kundaje, P. Kheradpour, F. Pauli, S. Batzoglou, B. E. Bernstein, P. Bickel, J. B. Brown, P. Cayting, Y. Chen, G. DeSalvo, C. Epstein, K. I. Fisher-Aylor, G. Euskirchen, M. Gerstein, J. Gertz, A. J. Hartemink, M. M. Hoffman, V. R. Iyer, Y. L. Jung, S. Karmakar, M. Kellis, P. V. Kharchenko, Q. Li, T. Liu, X. S. Liu, L. Ma, A. Milosavljevic, R. M. Myers, P. J. Park, M. J. Pazin, M. D. Perry, D. Raha, T. E. Reddy, J. Rozowsky, N. Shoresh, A. Sidow, M. Slattery, J. A. Stamatoyannopoulos, M. Y. Tolstorukov, K. P. White, S. Xi, P. J. Farnham, J. D. Lieb, B. J. Wold and M. Snyder (2012). "ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia." Genome Res **22**(9): 1813-1831.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome Biol **10**(3): R25.

Larsson, N. G., J. D. Garman, A. Oldfors, G. S. Barsh and D. A. Clayton (1996). "A single mouse gene encodes the mitochondrial transcription factor A and a testis-specific nuclear HMG-box protein." Nat Genet **13**(3): 296-302.

Larsson, N. G., J. Wang, H. Wilhelmsson, A. Oldfors, P. Rustin, M. Lewandoski, G. S. Barsh and D. A. Clayton (1998). "Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice." Nat Genet **18**(3): 231-236.

Maassen, J. A., T. H. LM, E. Van Essen, R. J. Heine, G. Nijpels, R. S. Jahangir Tafrechi, A. K. Raap, G. M. Janssen and H. H. Lemkes (2004). "Mitochondrial diabetes: molecular mechanisms and clinical presentation." Diabetes **53 Suppl 1**: S103-109.

Mukherjee, S., I. Patel and D. Bastia (1985). "Conformational changes in a replication origin induced by an initiator protein." Cell **43**(1): 189-197.

Ngo, H. B., J. T. Kaiser and D. C. Chan (2011). "The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA." Nat Struct Mol Biol **18**(11): 1290-1296.

Pepke, S., B. Wold and A. Mortazavi (2009). "Computation for ChIP-seq and RNA-seq studies." Nat Methods **6**(11 Suppl): S22-32.

Piao, Y., H. G. Kim, M. S. Oh and Y. K. Pak (2012). "Overexpression of TFAM, NRF-1 and myr-AKT protects the MPP(+)-induced mitochondrial dysfunctions in neuronal cells." Biochim Biophys Acta **1820**(5): 577-585.

Pierro, P., L. Capaccio and G. Gadaleta (1999). "The 25 kDa protein recognizing the rat curved region upstream of the origin of the L-strand replication is the rat homologue of the human mitochondrial transcription factor A." FEBS Lett **457**(3): 307-310.

Plaskon, R. R. and R. M. Wartell (1987). "Sequence distributions associated with DNA curvature are found upstream of strong E. coli promoters." Nucleic Acids Res **15**(2): 785-796.

Ross, M. G., C. Russ, M. Costello, A. Hollinger, N. J. Lennon, R. Hegarty, C. Nusbaum and D. B. Jaffe (2013). "Characterizing and measuring bias in sequence data." Genome Biol **14**(5): R51.

Satoh, M. and T. Kuroiwa (1991). "Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell." Exp Cell Res **196**(1): 137-140.

Schoenherr, C. J. and D. J. Anderson (1995). "The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes." Science **267**(5202): 1360-1363.

Shen-Li, H., R. C. O'Hagan, H. Hou, Jr., J. W. Horner, 2nd, H. W. Lee and R. A. DePinho (2000). "Essential role for Max in early embryonic growth and development." Genes Dev **14**(1): 17-22.

Simmons, R. A., I. Suponitsky-Kroyter and M. A. Selak (2005). "Progressive accumulation of mitochondrial DNA mutations and decline in mitochondrial function lead to beta-cell failure." J Biol Chem **280**(31): 28785-28791.

Stumpf, J. D., R. P. Saneto and W. C. Copeland (2013). "Clinical and molecular features of POLG-related mitochondrial disease." Cold Spring Harb Perspect Biol **5**(4): a011395.

Suarez, J., Y. Hu, A. Makino, E. Fricovsky, H. Wang and W. H. Dillmann (2008). "Alterations in mitochondrial function and cytosolic calcium induced by hyperglycemia are restored by mitochondrial transcription factor A in cardiomyocytes." Am J Physiol Cell Physiol **295**(6): C1561-1568.

Sugiyama, S., K. Hattori, M. Hayakawa and T. Ozawa (1991). "Quantitative analysis of age-associated accumulation of mitochondrial DNA with deletion in human hearts." Biochem Biophys Res Commun **180**(2): 894-899.

Suomalainen, A. and P. Isohanni (2010). "Mitochondrial DNA depletion syndromes--many genes, common mechanisms." Neuromuscul Disord **20**(7): 429-437.

Tarze, A., A. Deniaud, M. Le Bras, E. Maillier, D. Molle, N. Larochette, N. Zamzami, G. Jan, G. Kroemer and C. Brenner (2007). "GAPDH, a novel regulator of the pro-apoptotic mitochondrial membrane permeabilization." Oncogene **26**(18): 2606-2620.

Trifunovic, A. and N. G. Larsson (2008). "Mitochondrial dysfunction as a cause of ageing." J Intern Med **263**(2): 167-178.

Wang, Y. E., G. K. Marinov, B. J. Wold and D. C. Chan (2013). "Genome-wide analysis reveals coating of the mitochondrial genome by TFAM." <u>PLoS One</u> **8**(8): e74513.

Watanabe, A., M. Arai, N. Koitabashi, K. Niwano, Y. Ohyama, Y. Yamada, N. Kato and M. Kurabayashi (2011). "Mitochondrial transcription factors TFAM and TFB2M regulate Serca2 gene transcription." <u>Cardiovasc Res</u> **90**(1): 57-67.

Welter, C., S. Dooley, K. D. Zang and N. Blin (1989). "DNA curvature in front of the human mitochondrial L-strand replication origin with specific protein binding." <u>Nucleic Acids Res</u> **17**(15): 6077-6086.

Xu, S., M. Zhong, L. Zhang, Y. Wang, Z. Zhou, Y. Hao, W. Zhang, X. Yang, A. Wei, L. Pei and Z. Yu (2009). "Overexpression of Tfam protects mitochondria against beta-amyloid-induced oxidative damage in SH-SY5Y cells." <u>FEBS J</u> **276**(14): 3800-3809.

Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li and X. S. Liu (2008). "Model-based analysis of ChIP-Seq (MACS)." <u>Genome Biol</u> **9**(9): R137.

C h a p t e r 4

EVIDENCE FOR SITE-SPECIFIC OCCUPANCY OF THE MITOCHONDRIAL GENOME BY NUCLEAR TRANSCRIPTION FACTORS

Georgi K Marinov[1] *, Yun E Wang[1] *, David C Chan[1,2], Barbara J Wold[1]

* These authors contributed equally to the publication.

[1]Division of Biology, California Institute of Technology, Pasadena, CA

[2]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA

I. **Abstract**

Mitochondria contain their own circular genome, with mitochondria-specific transcription and replication systems and corresponding regulatory proteins. All of these proteins are encoded in the nuclear genome and are post-translationally imported into mitochondria. While several classically nuclear transcription factors have been reported to act in mitochondria and to bind to the D-loop, there has been no comprehensive mapping of their occupancy patterns across the genome. Furthermore, it is not clear how many other nuclear TFs may also be found in mitochondria. By using ChIP-seq data from the ENCODE, mouseENCODE and modENCODE consortia for 151 human, 31 mouse and 35 *C. elegans* factors, we were able to identify 8 human and 3 mouse transcription factors with strong localized enrichment over the mitochondrial genome that was usually associated with the corresponding recognition sequence motif. Notably, these sites of occupancy are often the sites with highest ChIP-seq signal intensity within both the nuclear and mitochondrial genomes and are thus best explained as true binding events to mitochondrial DNA, which exist in high copy number in each cell. We corroborated these findings by immunocytochemical staining evidence for mitochondrial localization. However, we were unable to find clear evidence for mitochondrial binding in ENCODE and other publicly available ChIP-seq data for most factors previously reported to localize there. As the first global analysis of nuclear transcription factors binding in mitochondria, this work opens the door to future studies that probe the functional significance of the phenomenon.

II.   <u>Introduction</u>

Mitochondria are the primary site of ATP production through oxidative phosphorylation and are therefore critical to eukaryotic cells. It is widely accepted that they arose as the result of an endosymbiotic event (Sagan 1967) between the ancestor of modern eukaryotes and a member of the α-proteobacteria clade (Yang, Oyaizu et al. 1985). Reflective of the organelle's prokaryotic ancestry, mitochondria retain their own circular genome (Nass, Nass et al. 1965), although its size has been greatly reduced in many eukaryotes through transfer of genes to the eukaryotic nucleus. After transcription and translation of nuclear components of the separate mitochondrial transcription, replication and regulatory machineries, a number of which retain evidence of their prokaryotic origin (Szklarczyk and Huynen 2010), the protein products are then imported back into the mitochondria to modulate organellar function.

The mitochondrial genome in mammals encodes 13 proteins, all of which are components of the electron transport chain, as well as 22 tRNAs and two rRNAs (Anderson, Bankier et al. 1981, Bibb, Van Etten et al. 1981). Mitochondrial DNA (mtDNA) is organized in cells as macromolecular DNA-protein complexes called nucleoids. Mitochondrial genes are densely packed along the genome, with the notable exception of the D-loop regulatory region (Shadel and Clayton 1997), which is located within the NCR. Transcription initiates in the D-loop, is carried out by the mitochondrial-specific RNA polymerase POLRMT, and results in long polycistronic transcripts from each strand (called the Heavy- or H-strand and the Light- or L-strand), from the light

strand promoter (LSP) and the two Heavy strand promoters (HSP1 and HSP2) (Cantatore and Attardi 1980, Montoya, Christianson et al. 1982). In addition, the transcription factors TFAM (Fisher and Clayton 1985, Fisher and Clayton 1988, Fisher, Lisowsky et al. 1992), and TFB2M as well as the methyltransferase TFB1M (Falkenberg, Gaspari et al. 2002, Gaspari, Larsson et al. 2004, Metodiev, Lesko et al. 2009) are required for initiation and regulation of transcription (Shutt, Bestwick et al. 2011). Unlike many of the proteins involved in regulation of the mitochondrial genome, these transcription factors are generally accepted as not being of prokaryotic origin. Instead, they are genes of eukaryotic ancestry, appropriated for their function through co-evolution of the organellar and cellular genomes and imported into mitochondria to regulate mtDNA transcription.

In addition to these well-characterized regulators of mitochondrial transcription, previous studies on a few transcription factors have suggested that certain ones which normally effect regulation of the nuclear genome may have an indirect or even direct effect on mitochondrial transcription (Achanta, Sasaki et al. 2005, Ryu, Lee et al. 2005, Leigh-Brown, Enriquez et al. 2010). The glucocorticoid receptor was the first such factor reported to localize to mitochondria and to interact with mtDNA (Demonacos, Tsawdaroglou et al. 1993, Demonacos, Djordjevic-Markovic et al. 1995, Koufali, Moutsatsou et al. 2003, Psarra, Solakidi et al. 2006). A 43kDa isoform of the thyroid hormone T3R α receptor, p43, has been found to directly control mitochondrial transcription (Wrutniak, Cassar-Malek et al. 1995, Casas, Rochard et al. 1999, Enriquez, Fernandez-Silva et al. 1999, Enriquez, Fernandez-Silva et al. 1999). CREB has been shown to localize to mitochondria, and ChIP followed by SACO on select 21-bp CRE-

like sites on the genome suggests that it binds to the D-loop (Cammarota, Paratcha et al. 1999, Lee, Kim et al. 2005, Ryu, Lee et al. 2005). The tumor suppressor transcription factor p53 has been implicated in mtDNA repair and regulation of gene expression through interactions with TFAM (Marchenko, Zaika et al. 2000, Yoshida, Izumi et al. 2003, Heyne, Mannebach et al. 2004, Achanta, Sasaki et al. 2005). The mitochondrial localization of the estrogen receptor is also well established, for both its ERα and ERβ isoforms, and it too has been suggested to bind to the D-loop (Monje and Boland 2001, Chen, Delannoy et al. 2004). NFκB and IκBα have been found in mitochondria and have been proposed to regulate mitochondrial gene expression (Cogswell, Kashatus et al. 2003, Johnson, Witzel et al. 2011). The AP-1 and PPARγ2 transcription factors have been proposed to localize to mitochondria and bind to the genome (Casas, Domenjoud et al. 2000, Ogita, Okuda et al. 2002, Ogita, Fujinami et al. 2003), and the MEF2D transcription factor was found to regulate the expression of the ND6 gene by binding to a consensus sequence recognition motif within it (She, Yang et al. 2011). Finally, the presence of STAT3 in mitochondria has been found to be important for the function of the electron transport chain and also to be necessary for TNF-induced necroptosis (Gough, Corlett et al. 2009, Wegrzyn, Potla et al. 2009, Szczepanek, Chen et al. 2011, Shulga and Pastorino 2012, Szczepanek, Chen et al. 2012), although direct mtDNA binding has not been established. Mitochondrial localization has also been reported for STAT1 and STAT5 (Boengler, Hilfiker-Kleiner et al. 2010, Chueh, Leong et al. 2010).

Despite the evidence for nuclear transcription factors localizing to mtDNA, direct *in vivo* immunoprecipitation evidence for the binding of these factors to the genome

exists only for CREB (Lee, Kim et al. 2005), p53 (Achanta, Sasaki et al. 2005), and MEF2D (She, Yang et al. 2011), and with the exception of MEF2D characterization, is limited to the D-loop region. No prior studies have assayed transcription factor occupancy across the entire mitochondrial genome in vivo with modern high-resolution techniques such as ChIP-seq (Chromatin Immunoprecipitation coupled with deep sequencing (Johnson, Mortazavi et al. 2007)). As a result, the precise nature, and in many instances the existence, of the proposed binding events remains unknown. The limited sampling of transcription factors in previous studies also leaves uncertain how common or rare localization to mitochondria and binding to mtDNA is for nuclear transcription factors in general.

Here we survey the large compendium of ChIP-seq and other functional genomic data made publicly available by the ENCODE, mouseENCODE and modENCODE Consortia (Gerstein, Lu et al. 2010, mod, Roy et al. 2010, Consortium 2011, Consortium, Bernstein et al. 2012, Mouse, Stamatoyannopoulos et al. 2012) to identify transcription factors that associate directly with mtDNA and to characterize the nature of these interactions. We identify eight human and three mouse transcription factors for which robust evidence of site-specific occupancy in the mitochondrial genome exists. These sites exhibit the strand asymmetry typical of nuclear transcription factor binding sites, usually contain the recognition motifs for the factors in question, and are typically the strongest (as measured by ChIP-seq signal strength) binding sites found in both the nuclear and mitochondrial genome by a wide margin. Notably, these interactions are all found outside of the non-coding D-loop region. The D-loop region itself exhibits

widespread sequencing read enrichment for dozens of transcription factors. However, it does not show the aforementioned feature characteristics of true binding events. Though not observed in control datasets generated from sonicated input DNA, the high ChIP-seq signal over the D-loop is frequently seen in control datasets generated using mock immunoprecipitation, suggesting that it is likely to represent an experimental artifact. Examination of available ChIP-seq data for the transcription factors previously proposed to play a role in mitochondria (GR, ERα, CREB, STAT3, p53) revealed no robust binding sites except for enrichment in the D-loop. Resolving the functional significance of the identified occupancy sites in future studies should provide exciting insights into the biology of both mitochondrial and nuclear transcriptional regulation.

III.   <u>Results</u>

In the course of our study of TFAM occupancy in the mitochondrial and nuclear genomes (Wang, Marinov et al. 2013), we noticed that a number of nuclear transcription factors exhibit localized enrichment in certain areas of the mitochondrial genome in ChIP-seq data **(Figure 4.1)**. These events could be divided in two classes: high ChIP-seq signal over the NCR, and localized high read density over regions outside of it. Given prior reports suggesting that nuclear transcription factors might act in mitochondria, the potential of exploiting the power and resolution of existing ChIP-seq data to shed light on this phenomenon is significant. Thus, we surveyed available functional genomics data to

characterize the general prevalence of the phenomenon among transcription factors and investigate evidence of occupancy in detail. We took advantage of the wide compendium of human, mouse, fly and worm functional genomics data generated by the ENCODE, mouseENCODE, and modENCODE cortoria (Gerstein, Lu et al. 2010, mod, Roy et al. 2010, Consortium 2011, Consortium, Bernstein et al. 2012, Mouse, Stamatoyannopoulos et al. 2012) and analyzed it for mitochondrial binding events.

In collaboration with Georgi K. Marinov, we found:

1) The human, mouse, and *C. elegans* mitochondrial genomes are largely mappable

2) The signal intensity of mtDNA-mapping peaks is dependent on cell type

3) Eight human transcription factors and three mouse transcription factors have strong localized enrichment over the mitochondrial genome

4) Mitochondrial peaks rank in top three peaks in intensity across the genomes

5) Transcription factor cellular localization for identified TFs is mitochondrial despite the lack of a recognized MTS

Computation was performed by Georgi Marinov.

Identifying transcription factor binding events in the mitochondrial genome

Publicly available ENCODE and mouseENCODE ChIP-seq and control data from the UCSC Genome Browser and modENCODE data, were downloaded from

ftp://ftp.modencode.org, which included ChIP-seq data for 151 transcription factors in human cell lines (Wang, Zhuang et al. 2012), 31 in mouse, and 35 in *C.elegans*. We also downloaded DNase hypersensitvity (both DNase-seq (Thurman, Rynes et al. 2012) and DGF (Neph, Vierstra et al. 2012)), FAIRE-seq (Song, Zhang et al. 2011) and MNase-seq data as these datasets provide valuable orthogonal information about potentially artifactual patterns of read enrichment over the mitochondrial genome.

It is well known that the nuclear genome contains partial copies of the mitochondrial genome (NUMTs) (du Buy and Riley 1967, Hazkani-Covo, Zeller et al. 2010). Depending on their levels of divergence from the mitochondrial sequence, they can present an informatics challenge for distinguishing binding events to the true mitochondrial genome from binding events to NUMTs. For this reason, we aligned reads simultaneously against the nuclear and mitochondrial genomes, and retained only reads that map uniquely, and with no mismatches, relative to the reference. Although this removed some regions of the genome from analysis, we felt it was a necessary to use the most stringent set of putative mitochondrial binding events.

Fortunately, while the human mitochondrial genome contains numerous small islands of unmappable sequence, no large completely unmappable segments were present (**Figure 4.2A**). The mouse genome contains a large unmappable stretch between the CO1 and ND4 genes (**Figure 4.2B**). The *C. elegans* mitochondrial genome is almost completely uniquely mappable (**Figure 4.2C**). In contrast, the *D. melanogaster* genome is almost completely unmappable (**Figure 4.2D**), that some NUMTs were of very high

sequence similarity. We therefore excluded fly datasets from further analysis and focused on human, mouse and worm data.

Mitochondrial read recovery is proportional to cell type mtDNA copy number

Mammalian cells typically contain hundreds to thousands of copies of mtDNA, with the precise number varying depending on the metabolic needs of the particular cell type (Bogenhagen and Clayton 1974, Williams 1986, Satoh and Kuroiwa 1991). This variation is relevant to our analysis because the relative read density over the mitochondrial genome is expected to scale with the mtDNA:nuclear DNA ratio for a given cell. Thus, cell types with very high mtDNA copy number are expected to display correspondingly elevated background read density over the mitochondrial genome. Several types of ENCODE data provide a rough proxy for the relative mitochondrial genome copy number per cell. In particular, the fraction of reads originating from the mitochondrial genome in DNase hypersensitivity and ChIP control datasets is expected to scale accordingly. We examined the distribution of this fraction in ENCODE and mouseENCODE DGF datasets and observed very large differences between different cell lines and tissues **(Figure 4.3)**. For example, about half of reads in K562 DGF data originated from mitochondria, while the fraction was less than 2% in CD20+ B-cells **(Figure 4.3A)**. Notably, these differences are in many cases (though not always) consistent with what is known about the cell lines, with certain cancer cell lines (such as K562 and A549) and muscle cells (LHCN) showing the largest number of mitochondrial

reads, while primary cells with small volumes of cytoplasm such as B-cells showed the least.

Mouse DGF data was available mostly for tissues, and the fraction of mitochondrial reads in these was much smaller compared to both the human cell lines and the few mouse cell lines assayed **(Figure 4.3B)**. This is consistent with a significant proportion of cells in tissues being in a less active metabolic state than cell lines in culture. Still, we observed expected differences between tissues. For example, one of the tissues that was most enriched for reads mapping to the mitochondrial genome was the heart. This cell and tissue-dependent difference was observed in ChIP control datasets as well. We observed similarly large differences in ChIP control datasets **(Figure 4.3C-D),** although the absolute number of reads was much lower than it was in DGF data. Again, the mouse tissues with the highest number of mitochondrial reads were the more metabolically active ones, such as brown adipose tissue, cortex, and heart.

These large differences in background read coverage between different cells lines/tissues have two consequences for the analysis of putative transcription factor binding to the mitochondrial genome. First, peak calling algorithms usually used to identify transcription factor binding sites from ChIP-seq data may not work equally well in different cell lines due to the highly variable background read density. Second, these differences render comparing the strength of binding across cell lines difficult.

To address the technical challenges that the difference in fraction of reads posed, we devised a normalization procedure to convert read coverage to signal intensity *z-*

scores reflecting how strongly regions of enrichment stand out compared to the average background read density along the mitochondrial genome for each dataset. We then used the maximum z-scores for each dataset to identify datasets with very strong such enrichment, which we then examined manually in detail, as traditional peak-calling software does not adapt well to peak calling on the small, read-dense mitochondrial genome.

D-Loop ChIP-seq signal is likely artifactual

The distribution of read density z-scores for transcription factor ChIP-seq and control datasets in seven ENCODE human cell lines (GM1278, K562, HepG2, HeLa, H1-hESC, IMR90 and A549) is shown in **Figures 4.4, 4.5 and 4.6**. A wide range in the values of the maximum z-score is observed, from less than 5, to more than 100. Strikingly, most factors exhibit high read density in the NCR. One obvious explanation for this observation is that it represents an experimental artifact. This is likely, as the NCR contains the D-loop (Sebastian, Faralli et al. 2013), the unique triple-strand structure of which could conceivably either cause overrepresentation of DNA fragments originating from it in sequencing libraries or it could be non-specifically bound by antibodies during the immunoprecipitation process. To distinguish between these possibilities, we carried out the same analysis on DNase, FAIRE and MNase data. As these assays do not involve an immunoprecipitation step, they are a proper control for sequencing artifacts. We did not observe significant localized read enrichment in these

datasets **(Figure 4.7)**, suggesting that the observed read enrichment over the D-loop is not due to sequencing biases or overrepresentation of D-loop fragments in ChIP libraries. Similarly, we did not observe enrichment in the matched sonicated input ChIP-seq control datasets. However, a number of mock-immunoprecipitation [IgG] control datasets did exhibit high z-scores (up to >50 in K562 cells) and closely matched the signal profile over the D-loop of ChIP-seq datasets **(Figure 4.8B)**. We also examined the forward and reverse strand read distribution in the NCR **(Figure 4.8)**. Site-specific transcription factor binding events display a characteristic asymmetry in the distribution of reads mapping to the forward and reverse strands, with reads on the forward strand showing a peak to the left of the binding site and reads on the reverse strand showing a peak to the right of it **(Figure 4.8C)**. Such read asymmetry was not observed in the D-loop region (average profile shown in **Figure 4.8A**, individual dataset profile shown in **Figure 4.1**).

These results suggest that while immunoprecipitation is necessary for high enrichment over the D-loop, the enrichment might not be mediated by the proteins targeted by the primary antibody. This does not explain why a large number of factors show little enrichment over the D-loop **(Figures 4.4, 4.5 and 4.6)** and why some factors show enrichment that is much higher than that observed in K562 IgG controls, with z-scores of up to 300 (compared to a maximum of 50 for the most highly enriched IgG controls). Still, given the lack of clear hallmarks of site-specific occupancy, and the IgG control results, enrichment over the D-loop has to be provisionally considered to be primarily the result of an experimental artifact, even if it cannot be ruled that at least in

some cases it is the result of real biochemical association with nuclear transcriptional regulators.

Identification of mtDNA peaks in 8 human transcription factors

In contrast to the widespread, but likely artifactual, read enrichment over the D-loop, we observed strong enrichment, exhibiting the canonical characteristics of a ChIP-seq peak over a true transcription factor binding site, in other regions of the human mitochondrial genome for eight of the examined transcription factors using a minimum z-score threshold of 20: CEBPβ, c-Jun, JunD, MafF, MafK, Max, NFE2 and Rfx5. **Figures 4.9 and 4.10** show the forward and reverse strand read distribution for representative replicates of each factor in each assayed cell line, as well as the occurrences of the corresponding explanatory motifs (identified from the top 500 ChIP-seq peaks in the nuclear genome, see Methods for details). The putative binding sites outside of the D-loop are characterized by an asymmetric forward and reverse strand read distribution, and in most cases, the presence of the explanatory motif in a position consistent with binding by the factor. We identified multiple binding sites for CEBPβ: a strong site of enrichment around the 5′ end of the CYB gene, what seems to be two closely clustered sites in the ND4 gene, a weaker site in the ND4L gene, and two other regions of enrichment over CO2 and CO1 **(Figure 4.9D)**. A single very strong binding site over the ND3 gene was observed for c-Jun, as well as two weaker sites, one coinciding with the ND4 CEBPβ sites and one near the 5′ end of ATP6 **(Figure 4.9B)**; the strong ND3 site was

also observed for JunD in HepG2 cells. Max exhibited two putative binding sites: one in the middle of the 16S rRNA gene, containing a cluster of Max motifs, and another one around the 5′ end of CO3, which also contains a cluster of Max motifs but is in a region of poor mappability. Strikingly, a common and very strong MafK and MafF binding site is present near the 3′ end of ND5, though it does not contain the common explanatory motif for both factors **(Figure 4.10A-B)**. MafK and MafF act as a complex in conjunction with MafG, so colocalization in the ChIP-seq data argues for a true binding event. Several putative binding sites were identified for NFE2: one close to the CEBPβ site in the 5′end of CYB, one over the tRNA cluster between ND4 and ND5, one in the 5′ end of ATP6 and one in the 16S rRNA gene **(Figure 4.10C)**. Finally, two putative binding sites are observed for Rfx5, at the 5′ end of ND5 and in the middle of CO2 **(Figure 4.10D)**. Intriguingly, these binding events are not always present in all cell lines. For example, CEBPβ binding around CYB was absent in K562, A549 and H1-hESC cells, while the MafK ND5 binding site was absent in GM18278 and H1-hESC cells, but present in the other cell lines for which data is available.

Identification of mtDNA peaks in 3 mouse and 1 *C. elegans* transcription factor

The same analysis as described above was carried out for both mouse and *C. elegans* datasets. We saw that three transcription factors (Max, MafK, and USF2) also exhibiting strong enrichment elsewhere in the mouse mitochondrial genome **(Figure 4.12)**. We observe a single MafK binding site, containing the explanatory motif and

situated over the tRNA cluster between the ND2 and CO1 genes **(Figure 4.12A)**. Max displayed a strong set of binding sites in the ND4 gene, and a weaker binding site near the 5' end of ND5; both sites contained the explanatory motif **(Figure 4.12B)**. Finally, a single site, also containing the explanatory motif for the factor and situated near the ND5 Max site, was present in CH12 USF2 datasets (but not in MEL cells) **(Figure 4.12C)**. MafK and Max were also assayed in human cells, and, as discussed above, putative mitochondrial sites were identified for both, though not at obviously orthologous positions in the genome to those found in the mouse data. We also analyzed available ChIP-seq data for the mouse orthologs of c-Jun and JunD, which in human cells exhibited putative mitochondrial binding sites. In contrast to observation in human, we did not detect strong sites for either protein in mouse.

Most *C. elegans* ChIP-seq datasets did not show very strong enrichment over the mitochondrial genome **(Figure 4.13A),** with the exception of DPY-27 and W03F9.2. Of these, only W03F9.2 exhibited regions of enrichment with the characteristics of transcription factor binding sites **(Figure 4.13B).** However, very little is known about this protein and the significance of its binding to the mitochondrial genome is unclear.

Mitochondrial peaks rank in top three peaks in intensity

The occupancy observations reported above for human and mouse mitochondria do not formally rule out the possibility that there are unannotated NUMTs in the genomes

of the cell lines in which binding is detected in our analysis and the observed binding is in fact nuclear. Such an explanation is superficially likely, given that binding to the mitochondrial genome was observed in some cell lines and not in others. However, closer examination reveals that this hypothesis would require different NUMTs in different cell lines as the cell lines that lack binding are not the same for all factors. For example, MafF and MafK binding is very prominent in K562 cells but CEBPβ and c-Jun seem not to bind to mtDNA in those cells. While still possible, we consider the independent insertion of multiple partial NUMTs in different cell lines to be an unlikely explanation for the observed binding patterns.

Each chromosome in the nuclear genome exists as only two copies in diploid cells, as compared to the hundreds of mitochondria, each of which may contain multiple copies of the mitochondrial genome (Bogenhagen and Clayton 1974, Williams 1986, Satoh and Kuroiwa 1991). Thus, higher read density over mitochondrial transcription factor binding sites than over nuclear ones is expected, assuming similar occupancy rates. We compared the peak height (in RPM) of the top 10 nuclear peaks (peak calls generated by the ENCODE consortium were downloaded from the UCSC Genome Browser) with that of the putatively mitochondrial binding sites **(Figure 4.14)**. We found that the mitochondrial binding sites are usually the strongest binding sites by a wide margin, or at least within the top three of all peaks. For example, while the strongest nuclear MafK peak in mouse CH12 cells has a peak height of 14.5 RPM, the mitochondrial binding site has a peak height of 290 RPM. These observations are difficult to explain as being the result of binding to unannotated NUMTs in the nuclear genome, but are entirely

consistent with the hypothesis that these factors indeed bind to the large number of copies of the mitochondrial genome present in each cell.

No mitochondrial targeting sequence in identified transcription factors

Algorithms for prediction of mitochondrial targeting sequences on proteins have previously been developed (Claros and Vincens 1996, Emanuelsson, Brunak et al. 2007). To test whether these sequences existed in our identified transcription factors, we ran TargetP and MitoProt analysis for all known isoforms **(Table 1)**. Unfortunately, no MTS was identified in any isoform using a threshold value of 80% probability. However, it is not unusual that mitochondrial proteins have no identified MTS.

Evidence for localization of transcription factors to mitochondria

If the observed binding sites in ChIP-seq data are the result of actual association of nuclear transcription factors with mtDNA, then these transcription factors should exhibit mitochondrial localization. We tested this directly by performing immunocytochemistry for MafK in HepG2 cells **(Figure 4.15)**. Strikingly, even though we did not find a predicted MTS, we observe clear colocalization of MafK to mitochondria in 60% of cells (*n = 124*). These observations provide independent corroboration for the mtDNA binding events identified through ChIP-seq.

No robust mitochondrial occupancy in ChIP-seq data for previously reported mitochondrially targeted nuclear factors

We note that none of the factors previously reported to be localized to mitochondria and to bind to mtDNA was retrieved by our analysis, even though CREB, GR, ERα, IRF3, NFκB, STAT1, STAT5A and STAT3 were assayed by the ENCODE Consortium. This failure could be attributed to the use of too stringent a z-score threshold when selecting datasets with significant enrichment. We therefore examined available ChIP-seq data against these factors more carefully **(Figure 4.16, Figure 4.17)**. We also performed the same analysis on published mouse and human p53 ChIP-seq data (Aksoy, Chicas et al. 2012, Li, He et al. 2012, Kenzelmann Broz, Spano Mello et al. 2013) **(Figure 4.18)**. Again, we did not observe any major sites of enrichment outside of the D-loop. For these factors, the D-loop region exhibits the same putatively artifactual pattern discussed previously. And for STAT3 and p53, even the enrichment over the D-loop was low. The one factor for which binding to mtDNA is confirmed by ChIP-seq is MEF2D, data for two of the isoforms of which in mouse C2C12 myoblasts was recently published (Sebastian, Faralli et al. 2013) **(Figure 4.19).** It exhibits a very complex binding pattern over large portions of the mouse mitochondrial genome, which is not straightforward to interpret, but nevertheless a number of locations exhibit strand asymmetry and contain the MEF2 sequence recognition motif. Notably, most of these are outside the ND6 gene.

It is at present not clear how to interpret these discrepancies. It is not surprising that some of these factors do not exhibit binding to mtDNA, as they were reported to play

a role in mitochondrial biology through mechanisms other than regulating gene expression (for example, IRF3 and STAT3). However, this is not the case for all of them. One possibility is that many prior studies reporting physical association of transcription factors with the D-loop suffered from the same artifactual read enrichment over that region that we observe, but this would not have been noticeable using the methods of the time. This would not be surprising, as it is only apparent that D-loop enrichment is likely to be artifactual when the high spatial resolution of ChIP-seq is combined with the joint analysis of input and mock immunoprecipitation controls. However, the mitochondrial localization of these factors has been carefully documented in a number of cases (Cammarota, Paratcha et al. 1999, Casas, Domenjoud et al. 2000, De Rasmo, Signorile et al. 2009). Another possibility is that binding to mtDNA only occurs under certain physiological conditions and the factors were assayed using ChIP-seq only in cellular states not matching those. Further analysis of ChIP-seq data collected over a wide range of conditions should help resolve these issues.

IV.   <u>Discussion</u>

We report here the first large-scale characterization of the association of nuclear transcription factors along the entire mitochondrial genome by utilizing the vast ChIP-seq data resource made publicly available by the ENCODE and modENCODE consortia. We find two classes of signal enrichment events, neither of which is detected in high-

throughput sequencing datasets that do not involve immunoprecipitation and therefore they are not due to sequencing biases. First, the majority of factors for which we detect strong read enrichment over the mitochondrial genome display high ChIP-seq signal only over the D-loop non-coding region in both human and mouse datasets. However, these signals do not have the characteristics of sequence specific occupancy and are present in a number of mock-immunoprecipitation control datasets. They are thus best explained as experimental artifacts, although it remains possible that they represent real non-canonical association with the D-loop for some factors. Second, for a subset of factors, specific ChIP-seq peaks are observed outside of the D-loop, and these display the additional hallmark characteristics of sequence specific occupancy.

Nuclear transcription factors previously reported to localize to mitochondria either did not exhibit significant enrichment in the available ChIP-seq datasets or, when they did, it was over the D-loop region with similar non-specific read distribution shape as other factors. In contrast, applying conservative thresholds we found eight human and three mouse transcription factors (two in common between the two species) that strongly occupy sites outside of the D-loop. They display the strand asymmetry pattern around the putative binding site that typifies true nuclear ChIP-seq peaks. Even more convincing is the fact that the explanatory motif for the factor is usually found under the observed enrichment peaks, further suggesting that they correspond to true in vivo biochemical events.

There are three main explanations for our observations. First, it is possible that despite our considerable bioinformatic precautions the observed binding events are in fact nuclear, originating from NUMTs present in the genomes of the cell lines assayed, but absent from the reference genome sequence. We believe that this is very unlikely. An experimental argument against unknown NUMTs comes from the strength of the ChIP-seq signal we see in the mitochondrial genome. These signals are much higher than even the strongest peaks in the nuclear genome for the same factor in the same dataset. This is expected for true mitochondrial genome binding because of the presence of many copies of the mitochondrial genome per cell, in contrast to the presence of only two copies of the nuclear genome. Second, it is possible that mitochondria are sometimes lysed in vivo, with mitochondrial DNA spilling into the cytoplasm where transcription factors could then bind. This cannot be ruled out based on the ChIP-seq data alone, but we consider it unlikely, as this would need to happen with a sufficient frequency to explain the remarkable strength of mitochondrial occupancy sites. The third, and most plausible, interpretation is that these nuclear transcription factors indeed translocate to the mitochondria and interact with the genome, as has been observed for the D-loop in some previous studies for other factors. Indeed, immunocytochemistry experiments in our study confirm the presence of MafK in mitochondria in a majority of HepG2 cells.

Several major questions are raised by our results. First, it is not clear how these nuclear transcription factors are targeted to the mitochondria. Mitochondrial proteins are typically imported into the mitochondrial matrix through the TIM/TOM protein translocator complex, and are targeted to the organelle by a mitochondrial localization

sequence, which is cleaved upon import. We scanned both human and mouse versions of our factors for mitochondrial target sequences (MTS) with both Mitoprot (Claros and Vincens 1996) and TargetP (Emanuelsson, Brunak et al. 2007) (using default settings), but we were unable to identify significant matches using either. This seems to be a common feature of nuclear transcription factors previously found to localize to mitochondria, most of which lack import sequences and are instead imported through other means (Casas, Domenjoud et al. 2000, Szczepanek, Chen et al. 2012). Posttranslational modifications may be important for import, as has been demonstrated for STAT3 in TNF-induced necroptosis (Shulga and Pastorino 2012).

Second, it is unclear why the same factor binds detectably to the mitochondrial genome in some cell types but not in others. It is certainly possible that different splice isoforms or post-translationally modified proteins are present in different cell types, with only some capable of being imported into mitochondria, or that import into mitochondria only happens under certain physiological conditions only met in some cell lines.

Third, the question of the biochemical reality of transcription factor binding at the D-loop remains open. Previous studies understandably focused on the D-loop, given its well-appreciated importance in regulating mitochondrial transcription. As a consequence, the literature supporting a role for some nuclear factors in mitochondria suggests that they do so through binding to the D-loop. Our analysis of ChIP-seq data, which was carried out in an agnostic manner, revealed that dozens of transcription factors – many more than had been studied locally at the D-loop alone – also show high level of

enrichment over the D-loop. However, the observed enrichment has characteristics suggesting that these signals are mainly due to experimental artifacts. In support of this judgment, the explanatory motifs for most of these factors were generally not found under the area of strongest enrichment in the D-loop. Therefore a conservative interpretation is that enrichment over the D-loop is an artifact in most cases.

Finally, and most importantly, the functional significance of factor occupancy observed by ChIP-seq remains unknown. It is entirely possible that it represents biochemical noise, with transcription factors entering the mitochondria because they have the right biochemical properties necessary to be imported, then binding to mtDNA but with little functional consequence. Alternatively, nuclear transcription factors may in fact be playing a regulatory role in mtDNA. It is difficult to imagine the exact mechanisms through which they might be acting, aside from interactions with the regulatory D-loop. While we do observe pairs of related factor such as c-Jun and JunD, and MafK and MafF binding to the same sites, binding events are overall widely dispersed over the mitochondrial genome and are found outside of the known regulatory regions. Plausible regulatory relationships are therefore not obvious and our results suggest that biological noise should be the working null hypothesis explaining the data. The functional regulatory role of these nuclear transcription factors in mitochondria is a very exciting possibility but it will have to be demonstrated in subsequent studies. Direct functional tests are the golden standard for establishing regulatory relationships, using gain and loss of function experiments and genetic manipulation of putative regulatory sites. The latter is at present not possible for mitochondria while the former are difficult to interpret in the

case of the role of nuclear transcription factors in mitochondrial gene regulation, as it is not easy to separate the direct effects of binding to mtDNA from the indirect effects of transcriptional changes in the nucleus. Thus, it may be some time before definitive answers to these questions are obtained. In the meantime, larger compendia of transcription factor ChIP-seq data such as those expected to be generated by the next phase of the ENCODE project will be a primary source of further insight by providing binding data for additional nuclear transcription factors that will clarify allowed or preferred occupancy patterns across the mitochondrial genome.

This analysis is the large-scale characterization of the association of nuclear transcription factors along the entire mitochondrial genome by utilizing the vast ChIP-seq data resource made publicly available by the ENCODE and modENCODE consortia. We find two classes of signal enrichment events, neither of which is detected in high-throughput sequencing datasets that do not involve immunoprecipitation, and therefore they are not due to sequencing biases. While the majority of factors for which we detect strong read enrichment over the mitochondrial genome display high ChIP-seq signal only over the D-loop non-coding region in both human and mouse datasets, these signals do not have the characteristics of sequence specific occupancy and are present in a number of mock-immunoprecipitation control datasets. They are thus best explained as experimental artifacts, although it remains possible that they represent real non-canonical association with the D-loop for some factors. Second, for a subset of factors, specific ChIP-seq peaks are observed outside of the D-loop, and these display the additional hallmark characteristics of sequence specific occupancy.

Previously identified nuclear transcription factors were not detected in our analysis of non-D-loop binding. One possibility for this is that many prior studies reporting physical association of transcription factors with the D-loop suffered from the same artifactual read enrichment over that region. This would not be surprising, as it is only apparent that D-loop enrichment is likely to be artifactual when the high spatial resolution of ChIP-seq is combined with the joint analysis of input and mock immunoprecipitation controls. However, the mitochondrial localization of these factors has been carefully documented in a number of cases (Cammarota, Paratcha et al. 1999, Casas, Rochard et al. 1999). The specific binding sites of these transcription factors remains to be studied in greater depth.

V.   <u>Methods</u>

Except for where indicated otherwise, all analysis was carried out using custom-written python scripts by Georgi Marinov.

Sequencing read alignment

Raw sequencing reads were downloaded from the UCSC genome browser for ENCODE and mouseENCODE data (Mouse, Stamatoyannopoulos et al. 2012), and from ftp://ftp.modencode.org for modENCODE data (data current as of February 2012). ChIP-seq data for p53 was obtained from GEO series GSE26361 (Li, He et al. 2012), GSE46240 (Kenzelmann Broz, Spano Mello et al. 2013) and GSE42728 (Aksoy, Chicas et al. 2012). Reads were aligned using Bowtie (Langmead, Trapnell et al. 2009), version 0.12.7. Human data was mapped against either the female or the male set of human chromosomes (excluding the Y chromosome and/or all random chromosomes and haplotypes) depending on the sex of the cell line (where the sex was known, otherwise the Y chromosome was included), genome version hg19. Mouse data was mapped against the mm9 version of the mouse genome. modENCODE *D. melanogaster* data was mapped against the dm3 version of the fly genome. modENCODE data for *C. elegans* was mapped against the ce10 version of the worm genome. Reads were mapped with the following settings: "-v 2 -k 2 -m 1 -t --best –strata", which allow for two mismatches relative to the reference, however for all downstream analysis only reads mapping

uniquely and with zero mismatches were considered, to eliminate any possible mapping artifacts.

Mappability track generation

Mappability was assessed as follows. Sequences of length N bases were generated starting at each position in the mitochondrial genome. The resulting set of "reads" was then mapped against the same bowtie index used for mapping real data. Positions covered by N reads were considered fully mappable. In this case, $N = 36$ as this is the read length for most of the sequencing data analyzed in this study.

Signal normalization of ChIP-seq data over the mitochondrial genome

Because the number of mitochondria per cell varies from one cell line/tissue to another, direct comparisons between datasets based on the absolute magnitude of the signal in RPM are not entirely valid. For this reason, we normalized the signal as follows. For each dataset, we fit a Gamma distribution over the RPM coverage scores for the bottom F_b percentile of fully mappable position on the mitochondrial chromosome. The estimated parameters were then used to rescale the raw signal over all position to a z-score. This results in datasets with strong peaks receiving low z-scores over most of the mappable mitochondrial genome, and very high z-scores over the regions with highly localized

enrichment. We used F=0.8 for our analysis. As this procedure is sensitive to datasets with very low total read coverage over the mitochondrial genome, we restricted our analysis to datasets with at least 5000 uniquely mappable reads (and with no mismatches to the reference), or more than 10x coverage. We used a z-score cutoff of 20 to select datasets with high enrichment over the mitochondrial genome, as it was the highest z-score observed in sonicated input samples.

Motif analysis

The peak calls for human and mouse ENCODE data available from the USCS Genome Browser were used to find de novo motifs for transcription factors from ChIP-seq data. The sequence around the peak summit (using a 50bp radius) was retrieved for the top 500 called peaks for each factor in each cell line and motifs were called using the MEME program in the MEME SUITE, version 4.6.1 (Bailey, Boden et al. 2009). The MEME-defined position weight matrix was then used to scan the mitochondrial genome for motif matches following the approach described in (Mortazavi, Leeper Thompson et al. 2006).

Cell growth and immunocytochemistry

HepG2 cells were grown following the standard ENCODE protocol [DMEM media, 4mM L-glutamine, 4.5g/L glucose, without sodium pyruvate, with 10% FBS (Invitrogen 10091-148) and penicillin-streptomycin]. Cells were fixed in 10% formalin (Sigma-Aldrich HT501128-4L) for 10 min, permeabilized with 0.1% Triton X-100, and blocked in 5% FBS. Primary antibodies used were MafK (1:100, Abcam, ab50322) and Hsp60 (1:125, Santa Cruz, sc-1052). Secondary antibodies used were donkey anti-goat AF488 (Invitrogen A11055) and donkey anti-rabbit AF546 (Invitrogen A10040). Imaging on a Zeiss LSM 710 confocal microscope with PlanApochromat 63X/1.4 oil objective, and 0.7μm optical sections were acquired.

VI.   <u>Figure Legends</u>

Figure 4.1. Representative USCS Genome Browser snapshots of nuclear transcription factor ChIP-seq datasets exhibiting strong enrichment in the mitochondrial genome. (A) GM12878 GCN5 shows high signal intensity in the D-loop (the region between coordinates 16030 and 580, i.e. the non-coding regions on the left and right ends of the snapshot) representative of the D-loop enrichment observed for a large number of transcription factors (B) In contrast, a large MafK peak is observed in a coding region outside of the D-loop in HepG2 cells. Upper track (black) shows reads aligning to the forward strand, lower track (gray) shows read aligning to the reverse strand.

Figure 4.2. Unique mappability of the mitochondrial genome in ENCODE and modENCODE species. (A) human; (B) mouse; (C) *C. elegans*; (D) *D. melanogaster*. The 36 bp mappability track (see Methods for details) is shown. The annotated protein coding and rRNA and tRNA genes are shown in the inner circles as follows: forward-strand genes are shown as green lines, while reverse-strand genes are shown as red lines, with the exception of mouse and human rRNA and tRNAs (blue). The D-loop region in human is shown in black. Gene annotations were obtained from ENSEMBL (version 66).

Figure 4.3. Variation in mitochondrial DNA copy number in cell lines and tissues.
The fraction of reads mapping to the mitochondrial genome (chrM) is shown. (A,B) UW
human (A) and mouse (B) UW ENCODE digital genomic footprinting (DGF) data; (C)
UW human ChIP input datasets; (D) LICR mouse ChIP input datasets. "UW" and
"LICR" refers to the ENCODE production groups that generated the data. Inputs from the
UW and LICR groups were chosen because they are the largest ENCODE sets in terms of
number of cell lines/tissues assayed by the same production groups, thus avoiding
possible variation between different laboratories. A general positive correlation between
the expected metabolic demand of the tissue type and the relative amount of reads
mapping to chrM is observed.

**Figure 4.4. Signal distribution over the mitochondrial genome in human ChIP-seq
datasets.** The maximum z-score for each individual TF ChIP-seq replicate in each cell
line is shown on the left (factors are sorted by average z-score, with control datasets
always shown on the bottom in red, below the red horizontal line). The z-score profile
along the mitochondrial chromosome for the replicate with the highest z-score is shown
on the right. "SYDH" and "HA" refer to the ENCODE production groups which
generated the data. Z-scores ≥ 100 are shown as equal to 100. (A) GM12878 cells; (B)
K562 cells.

Figure 4.5. Signal distribution over the mitochondrial genome in human ChIP-seq datasets. The maximum z-score for each individual TF ChIP-seq replicate in each cell line is shown on the left (factors are sorted by average z-score, with control datasets always shown on the bottom in red, below the red horizontal line). The z-score profile along the mitochondrial chromosome for the replicate with the highest z-score is shown on the right. "SYDH" and "HA" refer to the ENCODE production groups which generated the data. Z-scores ≥100 are shown as equal to 100. (A) HepG2 cells; (B) HeLa cells; (C) A549 cells.

Figure 4.6. Signal distribution over the mitochondrial genome in human ChIP-seq datasets. The maximum z-score for each individual TF ChIP-seq replicate in each cell line is shown on the left (factors are sorted by average z-score, with control datasets always shown on the bottom in red, below the red horizontal line). The z-score profile along the mitochondrial chromosome for the replicate with the highest z-score is shown on the right. "SYDH" and "HA" refer to the ENCODE production groups which generated the data. Z-scores ≥100 are shown as equal to 100. (A) H1-hESC cells; (B) IMR90.

Figure 4.7. Signal distribution over the mitochondrial genome in human FAIRE-seq, DNAse-seq and MNAse-seq datasets. Shown is the maximum z-score for each

individual replicate for each cell line (left) and the z-score profile along the mitochondrial chromosome for the replicate with the highest z-score (right). (A) FAIRE data; (B) DNAse data; (C) MNAse data. "UNC", "UW" and "SYDH" refer to the ENCODE production groups which generated the data. Z-scores larger than 100 are shown as 100. No read enrichment over the D-loop is observed, suggesting that the D-loop signal found in TF ChIP-seq datasets is not due to sequencing biases but is a result of the immunoprecipitation process.

Figure 4.8. Combined signal distribution profile for the forward and reverse strand in the D-loop region. Shown is the average signal (in RPM) for each strand in human ChIP-seq datasets with z-scores ≥ 20 (A) and human IgG controls (B). Also shown for comparison is the plus and minus strand read distribution around nuclear CTCF binding sites in H1-hESC cells (C).

Figure 4.9. Human transcription factors with canonical ChIP-seq peaks (displaying the typical strand asymmetry in read distribution around the putative binding site) outside of the D-loop. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details). Protein-coding, rRNA and tRNA genes are shown as colored bars. The

innermost circle shows the motif occurrences in the mitochondrial genome for each factor as black vertical bars. (A) JunD (B) c-Jun; (C) Max; (D) CEBPβ. The reads per million (RPM) tracks are shown, scaled to the maximum signal level (for both strands) for each dataset.

Figure 4.10. Human transcription factors with canonical ChIP-seq peaks (displaying the typical strand asymmetry in read distribution around the putative binding site) outside of the D-loop. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details). Protein-coding, rRNA and tRNA genes are shown as colored bars. The innermost circle shows the motif occurrences in the mitochondrial genome for each factor as black vertical bars. (A) MafF; (B) MafK (note that MafK has been assayed using two different antibodies in HepG2, both of which are shown); (C) NFE2; (D) Rfx5. The reads per million (RPM) tracks are shown, scaled to the maximum signal level (for both strands) for each dataset.

Figure 4.11. Signal distribution over the mitochondrial genome in mouse ChIP-seq datasets. Shown is the maximum z-score for each individual replicate for each cell line (left) and the z-score profile along the mitochondrial chromosome for the replicate with

the highest z-score (right). Control datasets are shown in red on the bottom, below the red horizontal line. (A) CH12 cells; (B) MEL cells.

Figure 4.12. Mouse transcription factors with canonical ChIP-seq peaks (displaying the typical strand asymmetry in read distribution around the putative binding site) outside of the D-loop. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details). Protein-coding, rRNA and tRNA genes are shown as colored bars. The innermost circle shows the motif occurrences in the mitochondrial genome for each factor as black vertical bars. (A) MafK (note that the putative binding site is found in a region that is not completely mappable, thus the read profiles loses the canonical shape but the strand asymmetry is nevertheless apparent and a motif is present); (B) Max; (C) USF2. The reads per million (RPM) tracks are shown, scaled to the maximum signal level (for both strands) for each dataset.

Figure 4.13. Signal distribution over the mitochondrial genome in *C.elegans* ChIP-seq datasets. (A) Shown is the maximum z-score for each individual replicate for each cell line (left) and the z-score profile along the mitochondrial chromosome for the replicate with the highest z-score (right). Control datasets are shown in red on the bottom,

below the red horizontal line; (B) Forward and reverse strand read distribution over the *C.elegans* mitochondrial genome for W03F9.2 ("Young Adult" stage). Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details).

Figure 4.14. Mitochondrial ChIP-seq peaks are generally significantly stronger than nuclear peaks. Shown is the maximum signal (in RPM) for the top 10 nuclear peaks ("N", smaller black dots), and the maximum signal intensity (also in RPM) in the mitochondrial genome ("M", larger red dot) for representative ChIP-seq datasets for each factor. (A) Mouse datasets (B) Human datasets.

Figure 4.15. Localization of MafK to the mitochondria. (A) Immunocytochemistry showing MafK localization in HepG2 cells. Mitochondria were identified by HSP60 staining. Shown are two representative images of cells showing that MAFK localizes strongly to the nucleus and mitochondria, and exhibits diffuse staining in the cytoplasm. In 60% of cells (C), there is colocalization of HSP60 with MAFK staining at an intensity higher than that of the surrounding cytoplasm. (B) An example of a cell exhibiting only nuclear and cytoplasmic MAFK localization.

Figure 4.16. Distribution of reads over the human mitochondrial genome for factors previously reported to bind to mitochondria in ENCODE ChIP-seq data. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details). Protein-coding, rRNA and tRNA genes are shown as colored bars. The innermost circle shows the motif occurrences in the mitochondrial genome for each factor as black vertical bars. (A) CREB; (B) STAT3; (C) GR in A549 cells treated with different concentrations of dexamethasone (Dex) (Reddy, Pauli et al. 2009, Reddy, Gertz et al. 2012); (D) ERα in untreated (DMSO) ECC1 cells and ECC1 cells treated with bisphenol A (BPA), genistein (Gen) or 17β-estradiol (E2) (Gertz, Reddy et al. 2012); (E) IRF3; (F) NFκB in GM12878 cells treated with TNFα (Kasowski, Grubert et al. 2010). The reads per million (RPM) tracks are shown, scaled to the maximum signal level (for both strands) for each dataset.

Figure 4.17. Distribution of reads over the human and mouse mitochondrial genome for p53 in publicly available ChIP-seq datasets. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track. Protein-coding, rRNA and tRNA genes are shown as colored bars. The innermost circle shows the motif occurrences in the mitochondrial genome for each factor as black

vertical bars. (A) p53 in mouse embryionic fibroblasts (MEFs), data from (Kenzelmann Broz, Spano Mello et al. 2013), GSE46240. (B) p53 in mouse embryonic stem cells (mESC), data from (Li, He et al. 2012), GSE26361; (C) p53 in human IMR90 cells, data from (Aksoy, Chicas et al. 2012), GSE42728. The reads per million (RPM) tracks are shown, scaled to the maximum signal level (for both strands) for each dataset.

Figure 4.18. Distribution of reads over the mouse mitochondrial genome for MEF2D isoforms MEF2Da1 and MEF2Da2 in C2C12 myoblasts. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details). Protein-coding, rRNA and tRNA genes are shown as colored bars. The innermost circle shows the MEF2D motif occurrences in the mitochondrial genome as black vertical bars. Data was obtained from (Sebastian, Faralli et al. 2013), GSE43223.

Figure 4.19. Distribution of reads over the human mitochondrial genome for STAT1 and STAT5A in ENCODE ChIP-seq data. Reads mapping to the forward strand are represented in black, reads mapping to the reverse strand are represented in yellow. The unique mappability track for the mitochondrial genome is shown in red in the outside track (see Methods for details). Protein-coding rRNA and tRNA genes are shown as

colored bars. The innermost circle shows the motif occurrences in the mitochondrial genome for each factor as black vertical bars. (A) STAT1; (B) STAT5A; the RPM tracks are shown, scaled to the maximum signal level (for both strands) for each dataset.

Table 4.1. No mitochondrial targeting sites are predicted in positive mouse and human TFs. Analysis of all variants of the eight human transcription factors and three mouse transcription factors resulted in generally low predicted MTS values by TargetP and MitoProt. No predicted probability exceeded the threshold of 0.80. Criteria for TargetP: non-plant organism, no cutoffs.
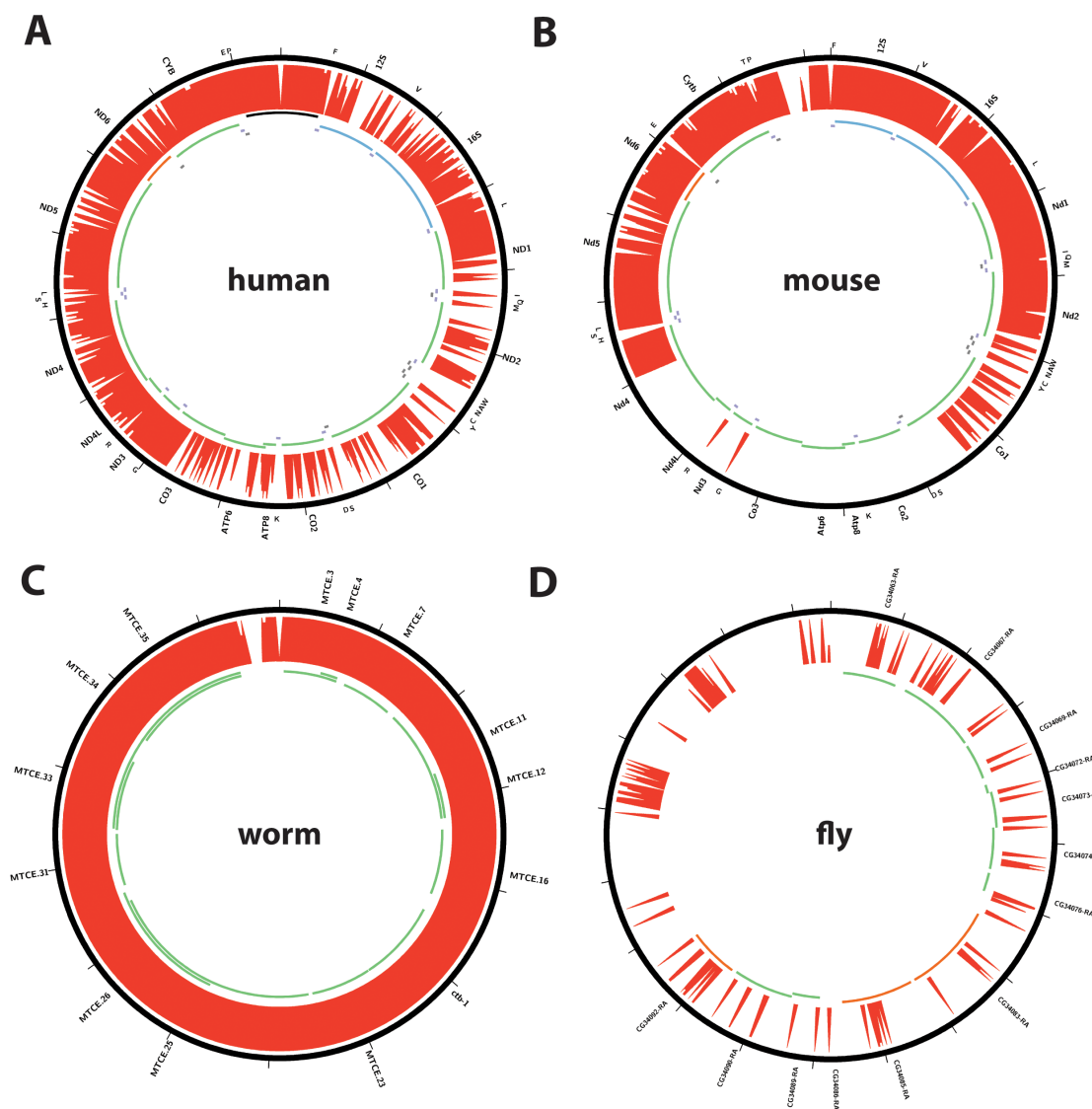
VII. <u>Figures</u>
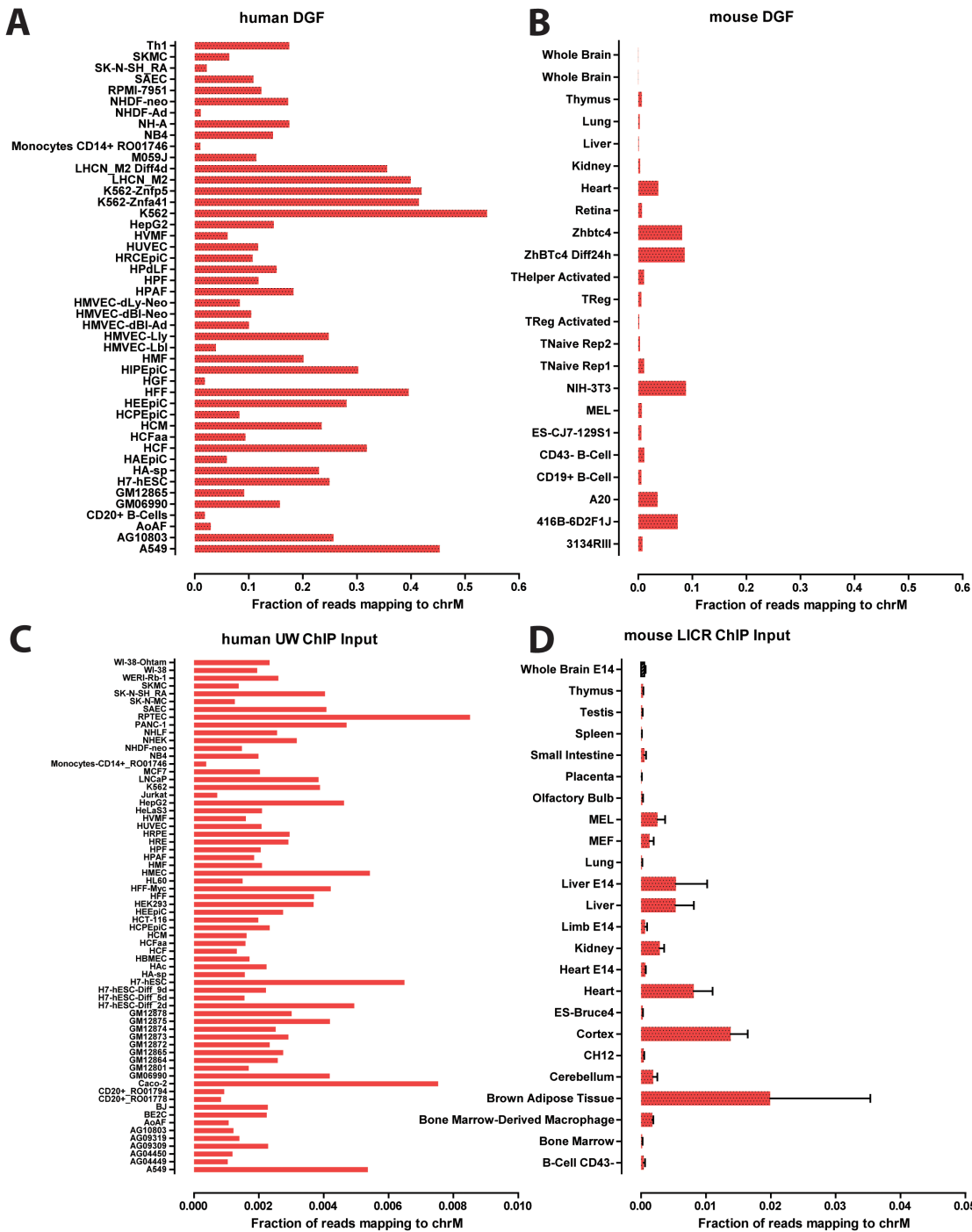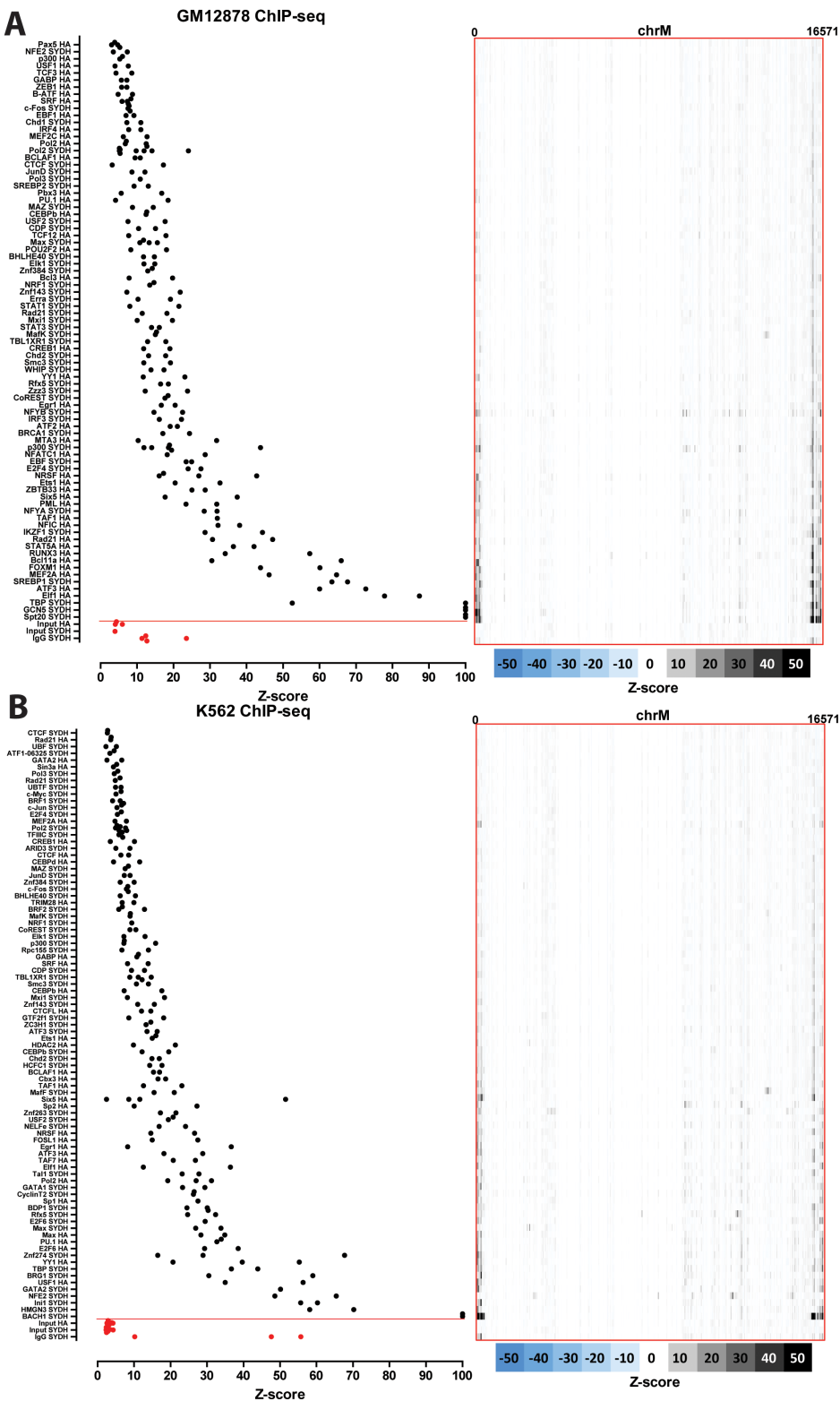
Figure 4.1.
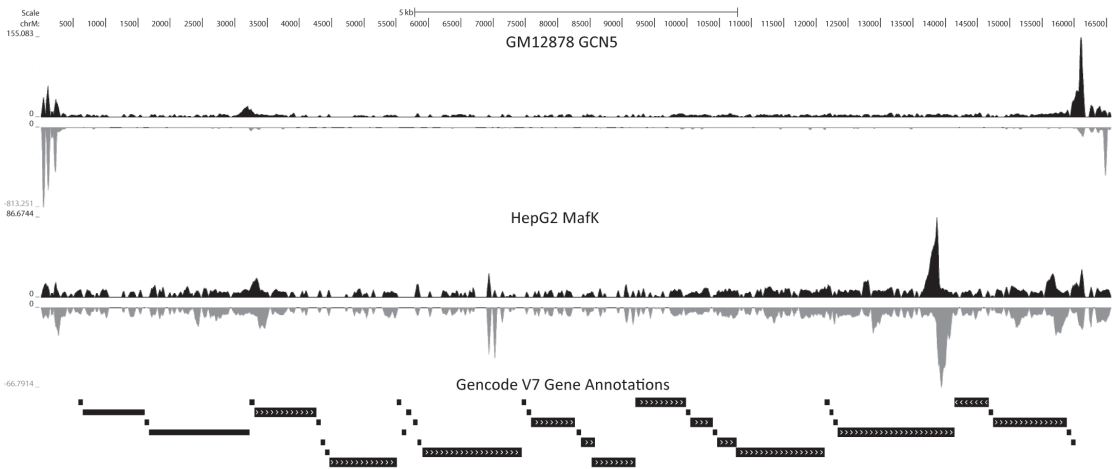
Figure 4.2.

Figure 4.3.

Figure 4.4.

Figure 4.5.
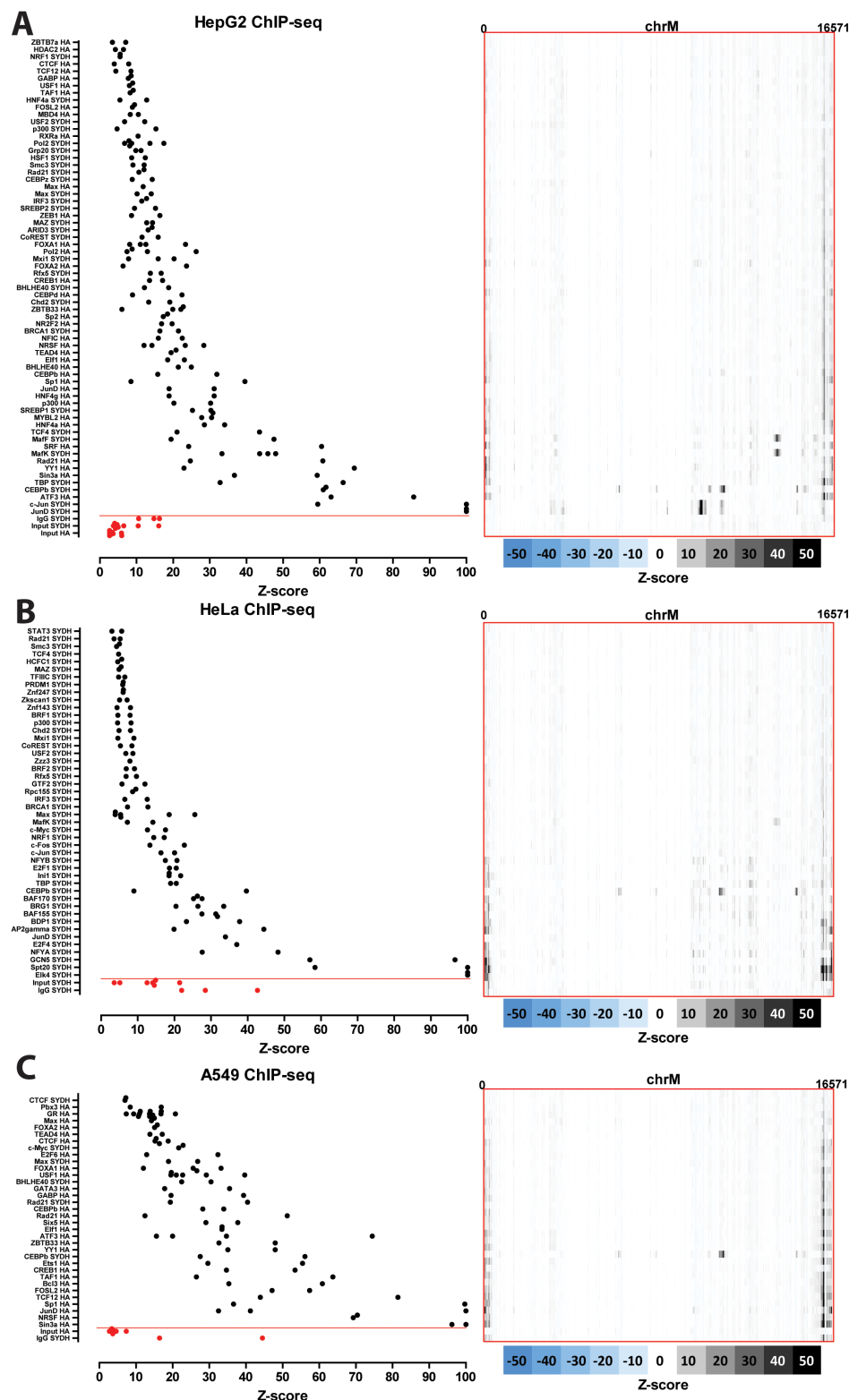
Figure 4.6.
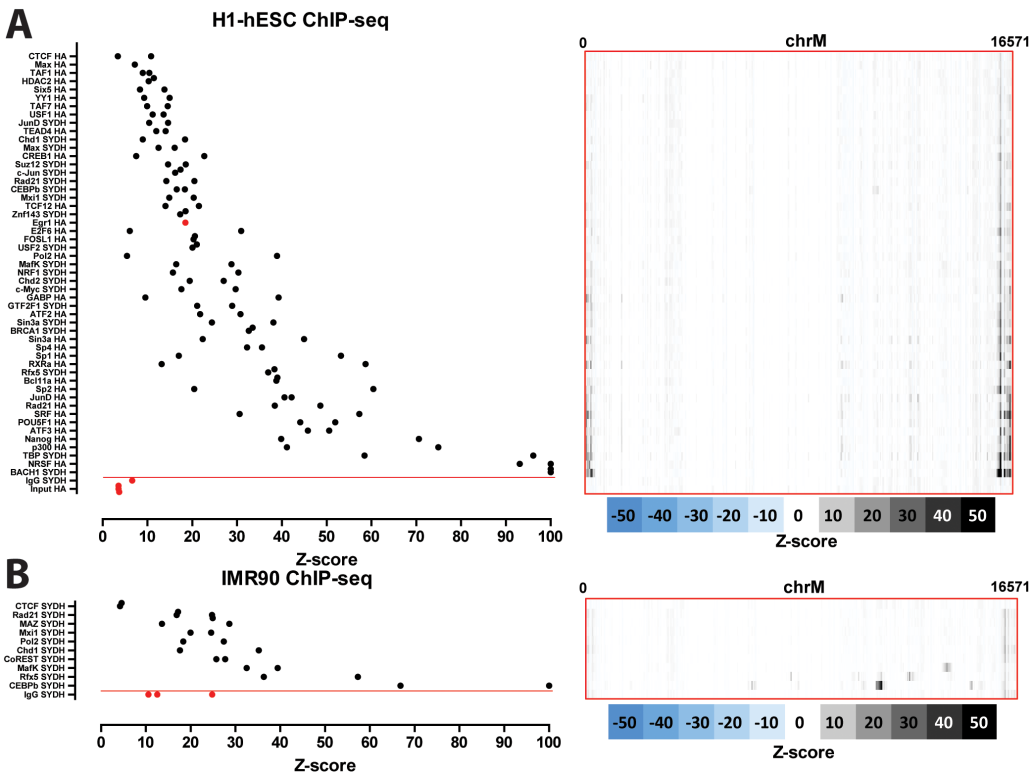
Figure 4.7.
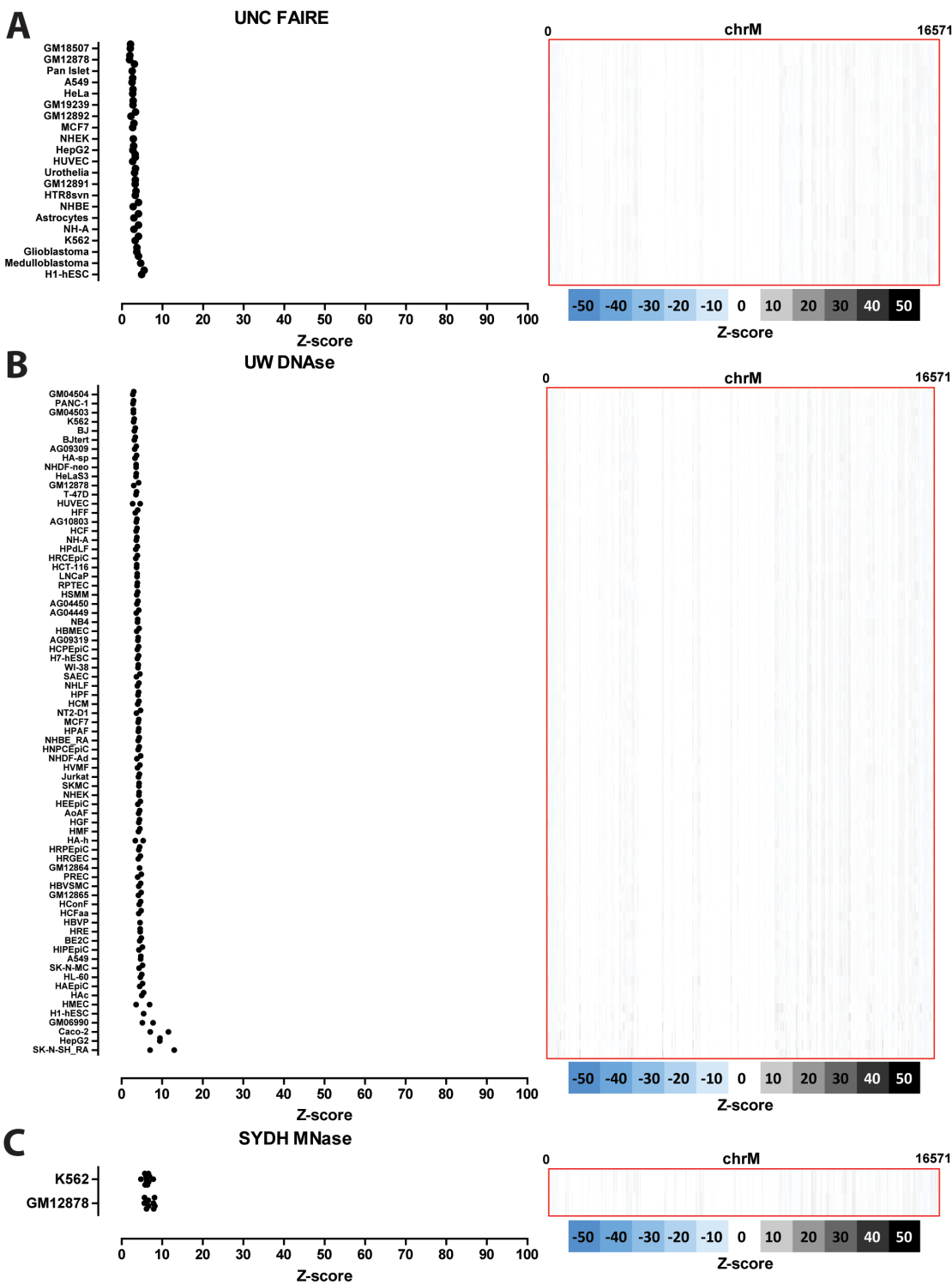
Figure 4.8.
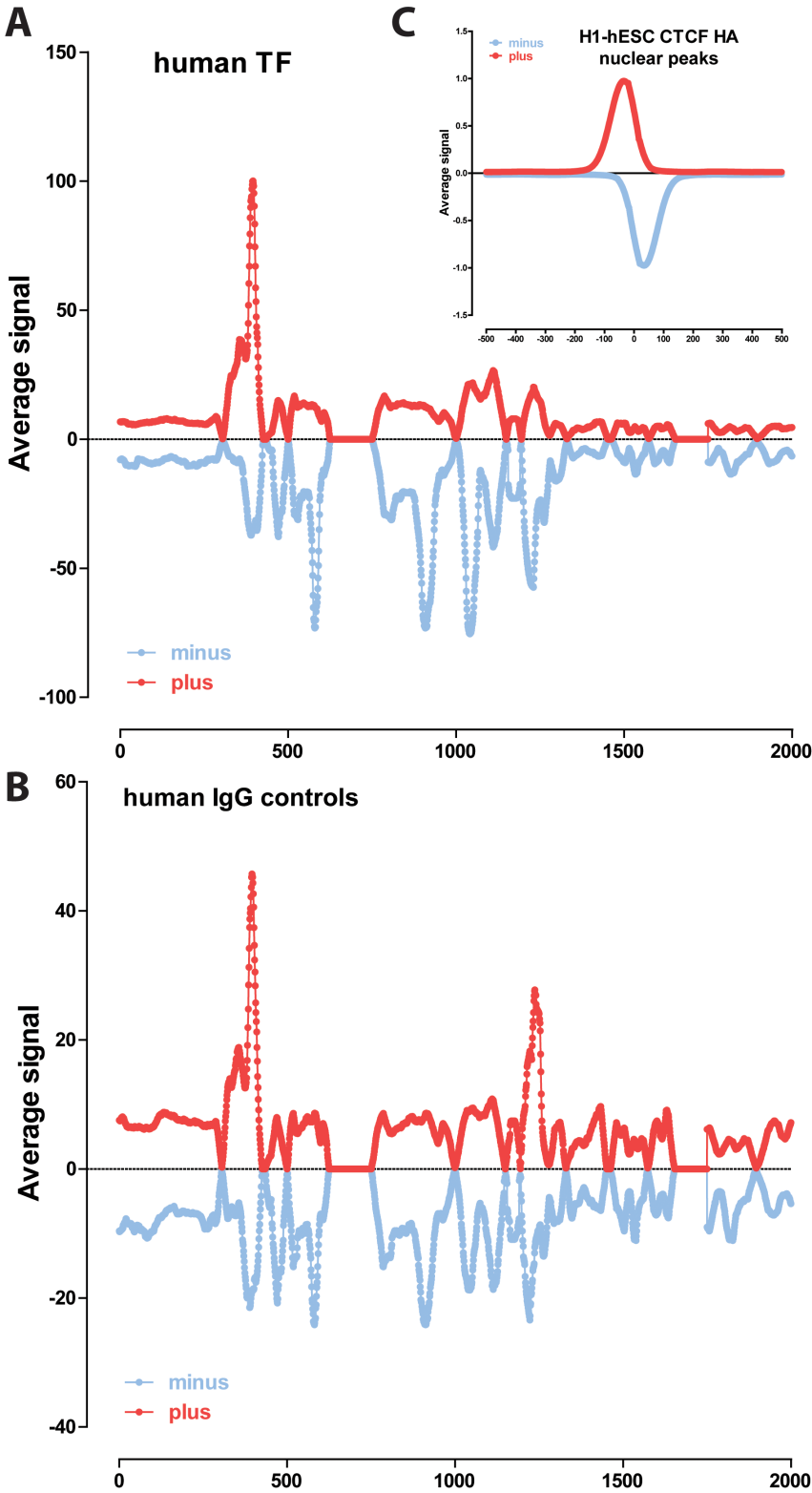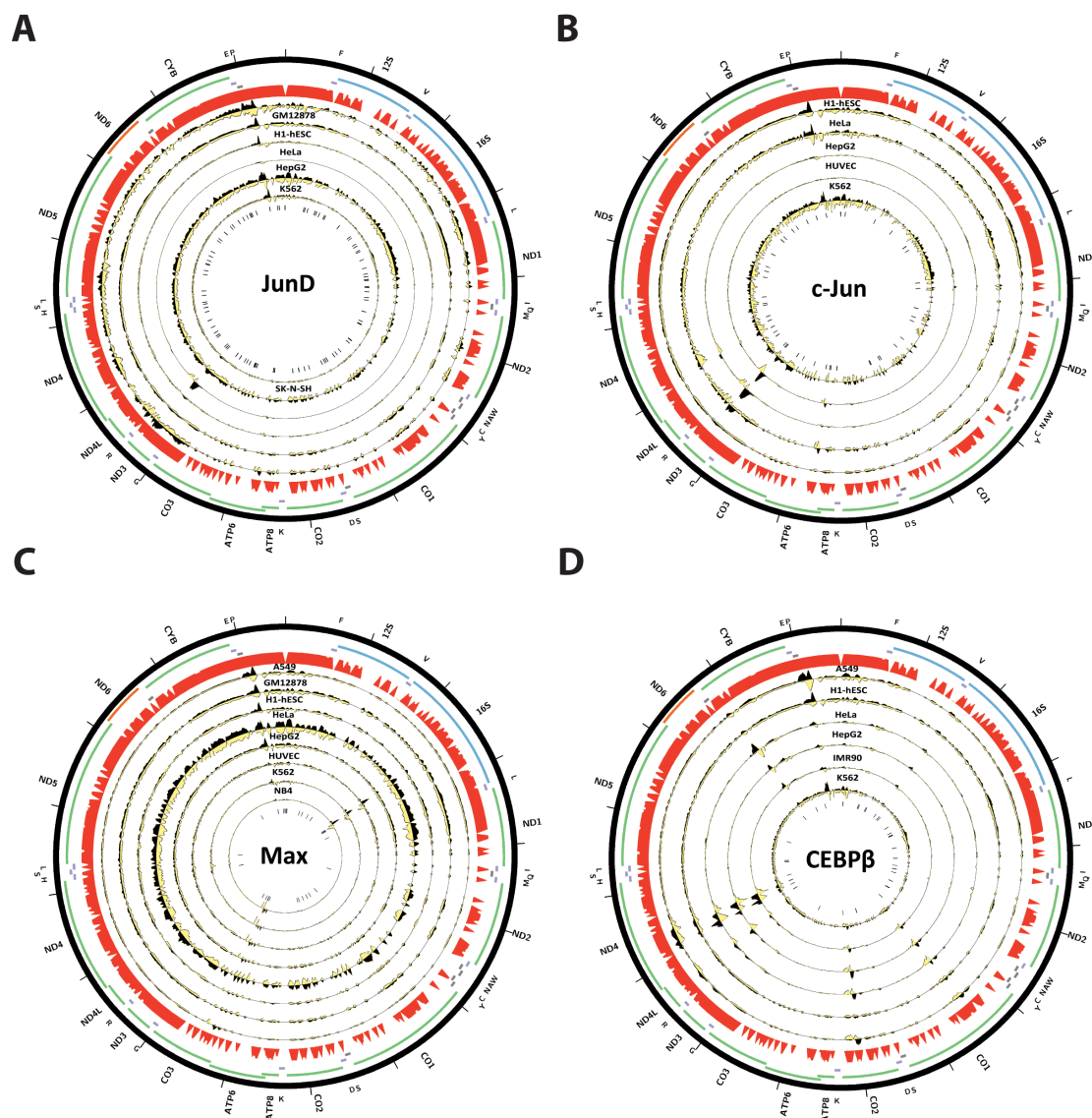
Figure 4.9.
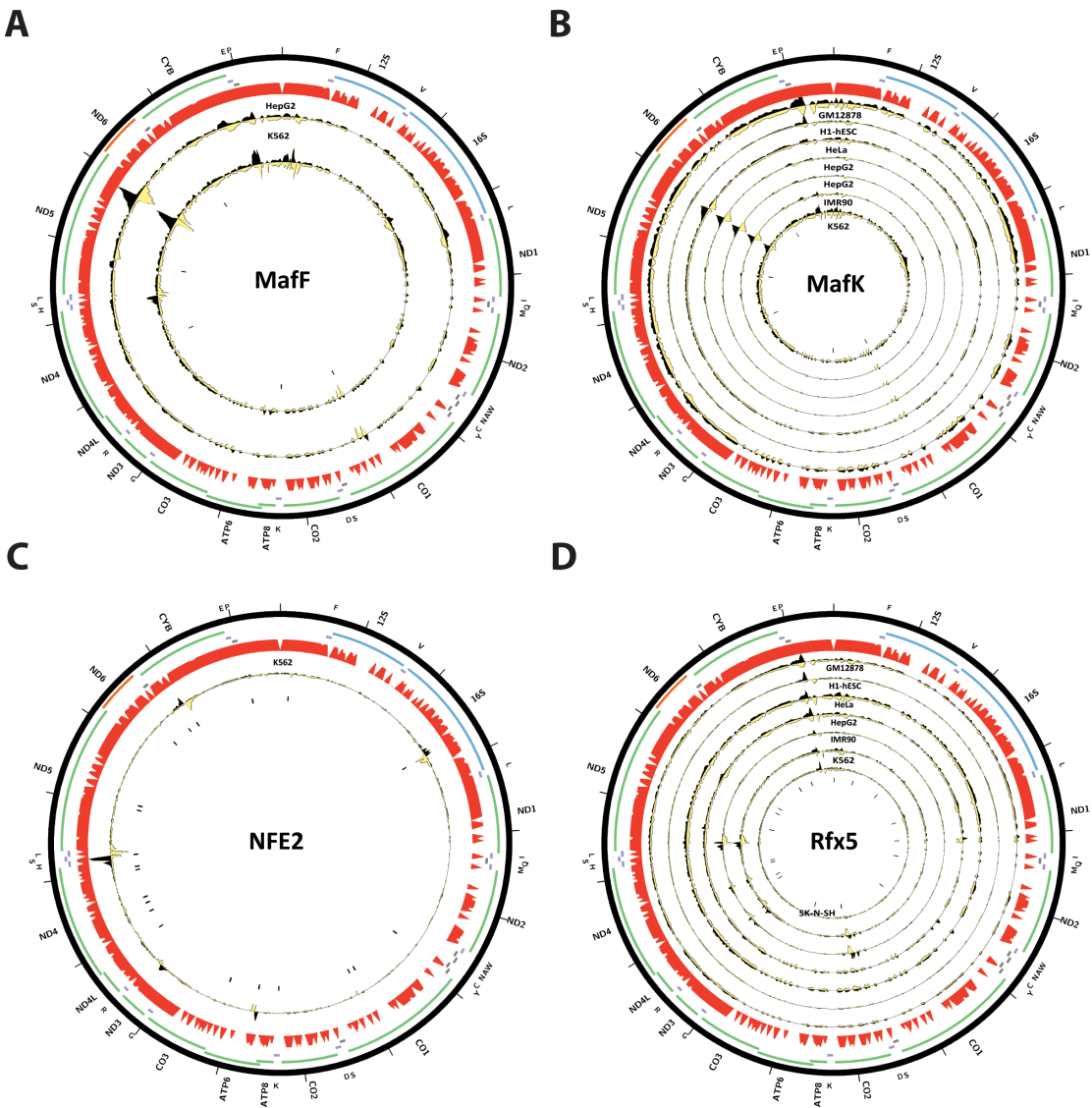
Figure 4.10.

Figure 4.11.
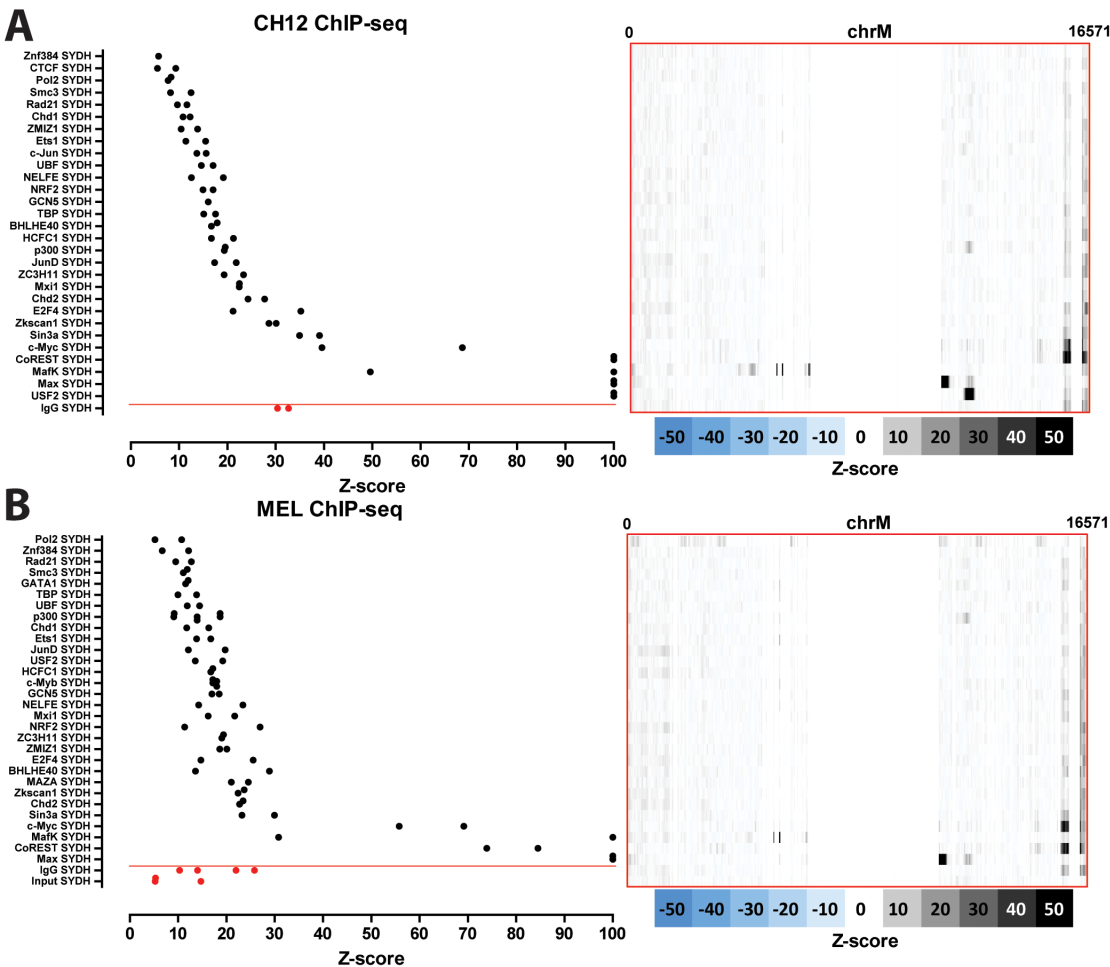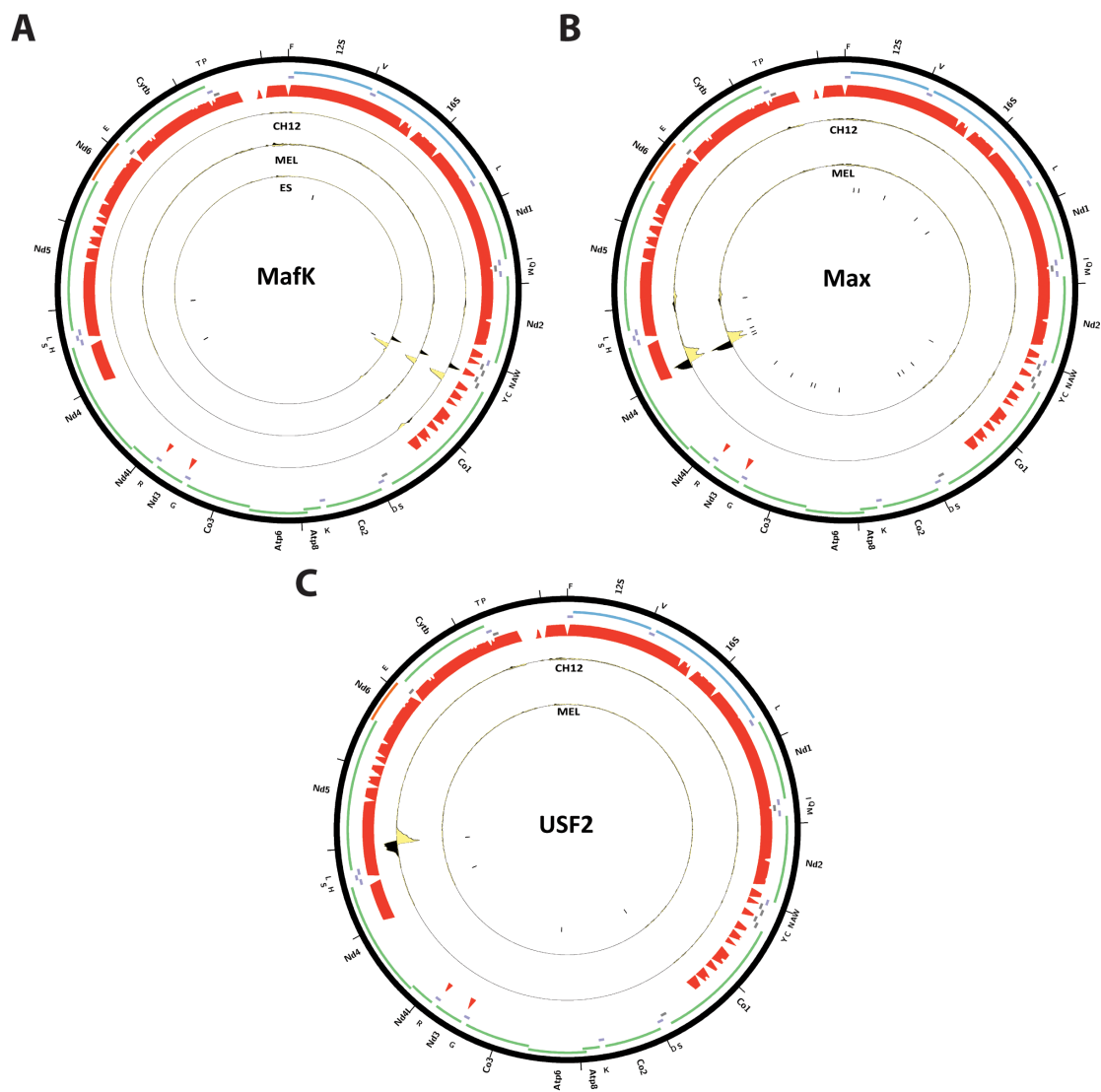
Figure 4.12.
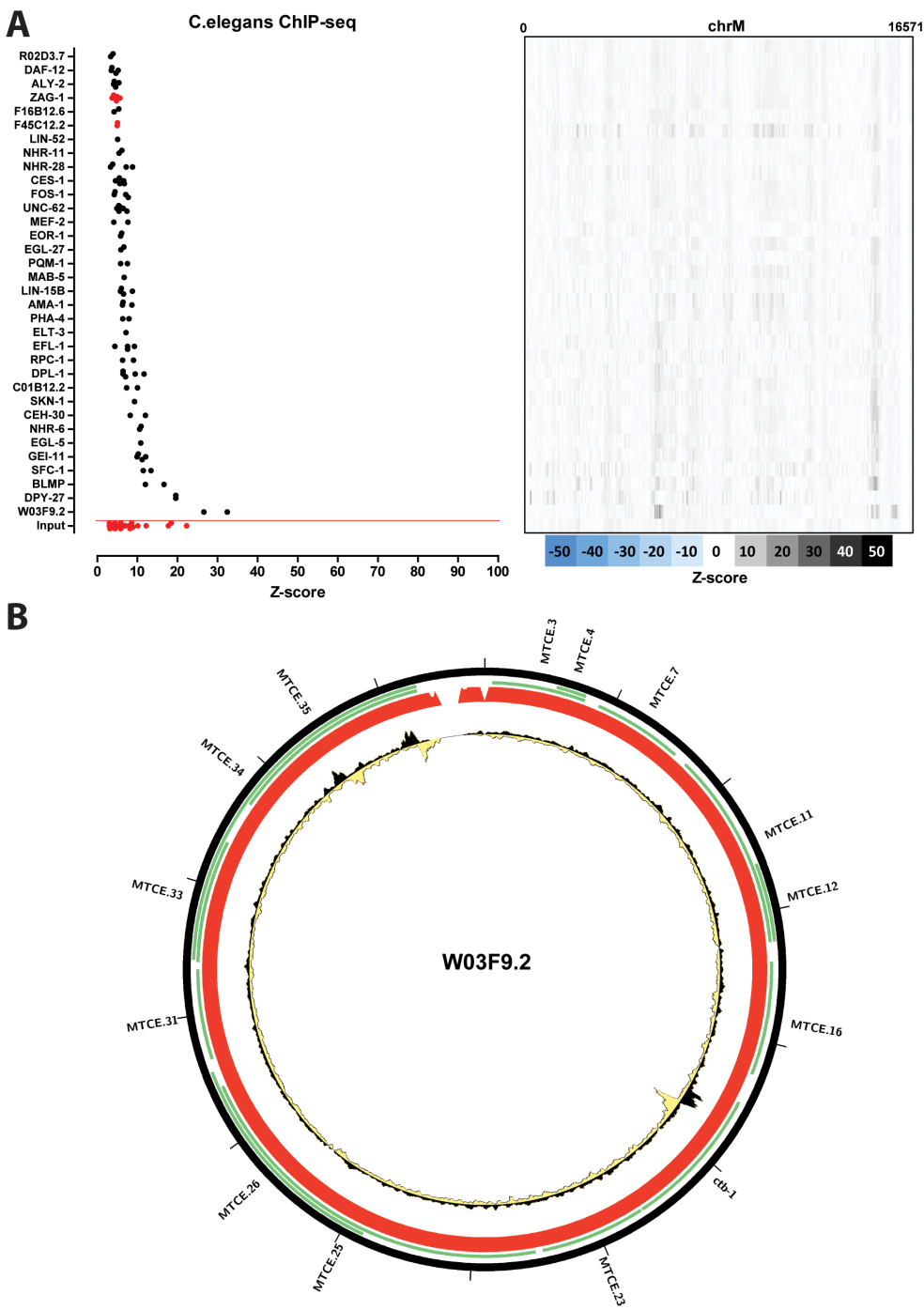
Figure 4.13.

Figure 4.14.
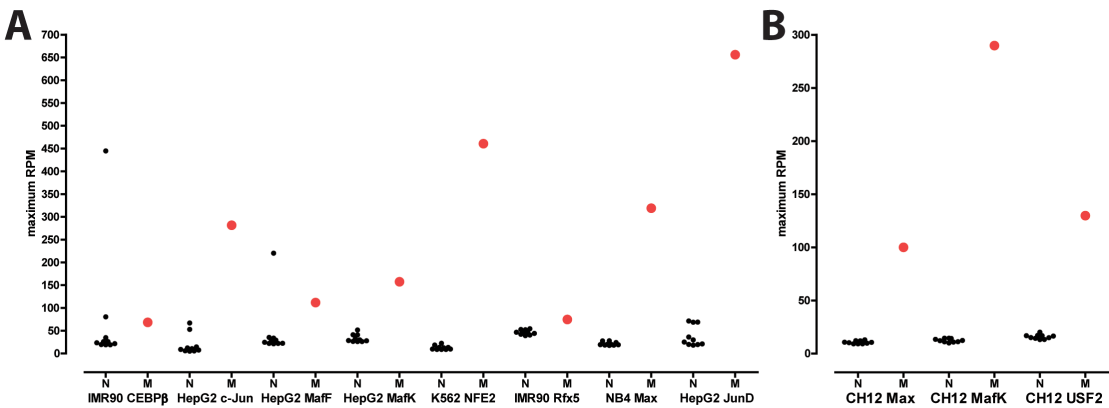
Figure 4.15.
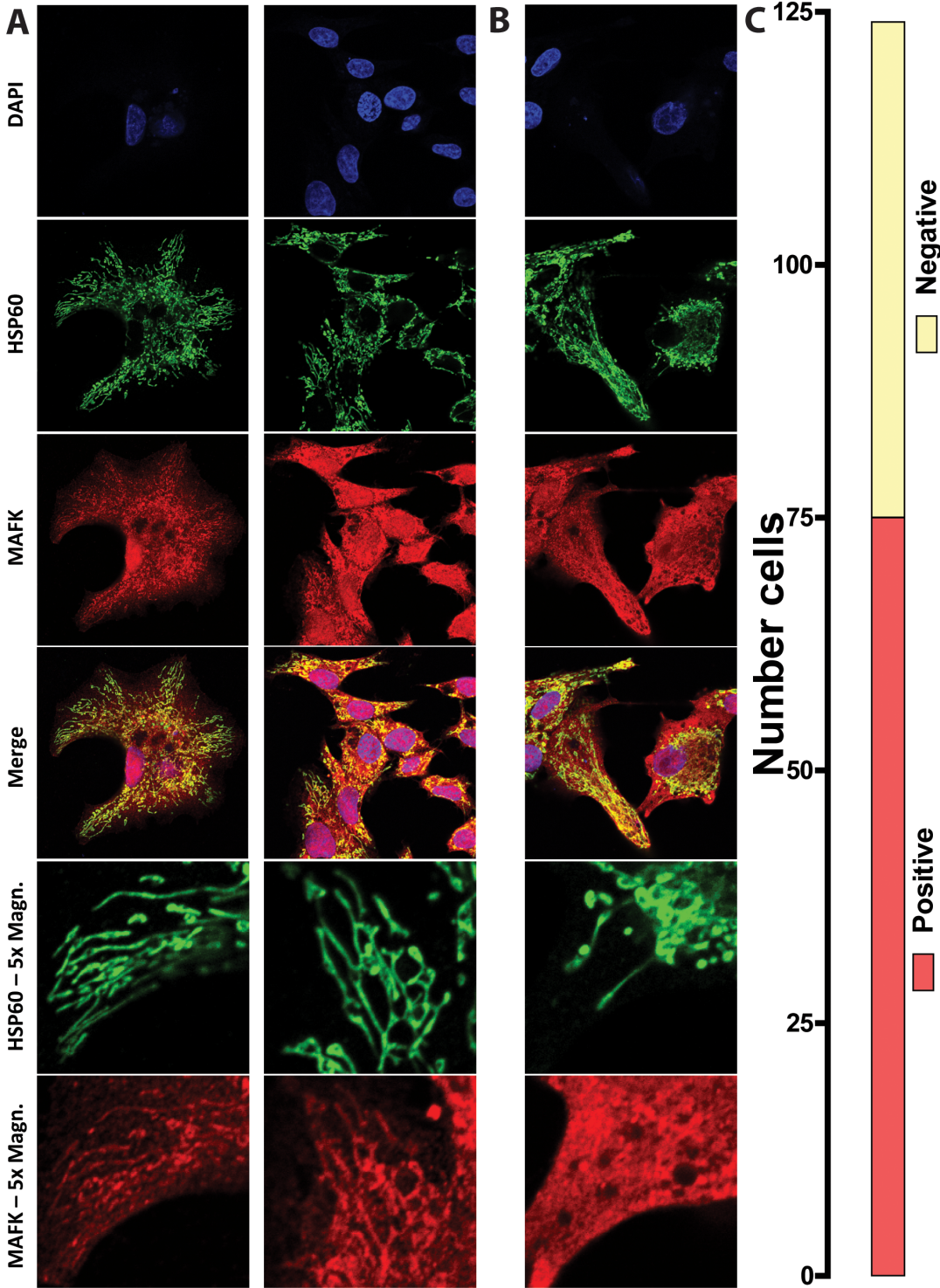
Figure 4.16.
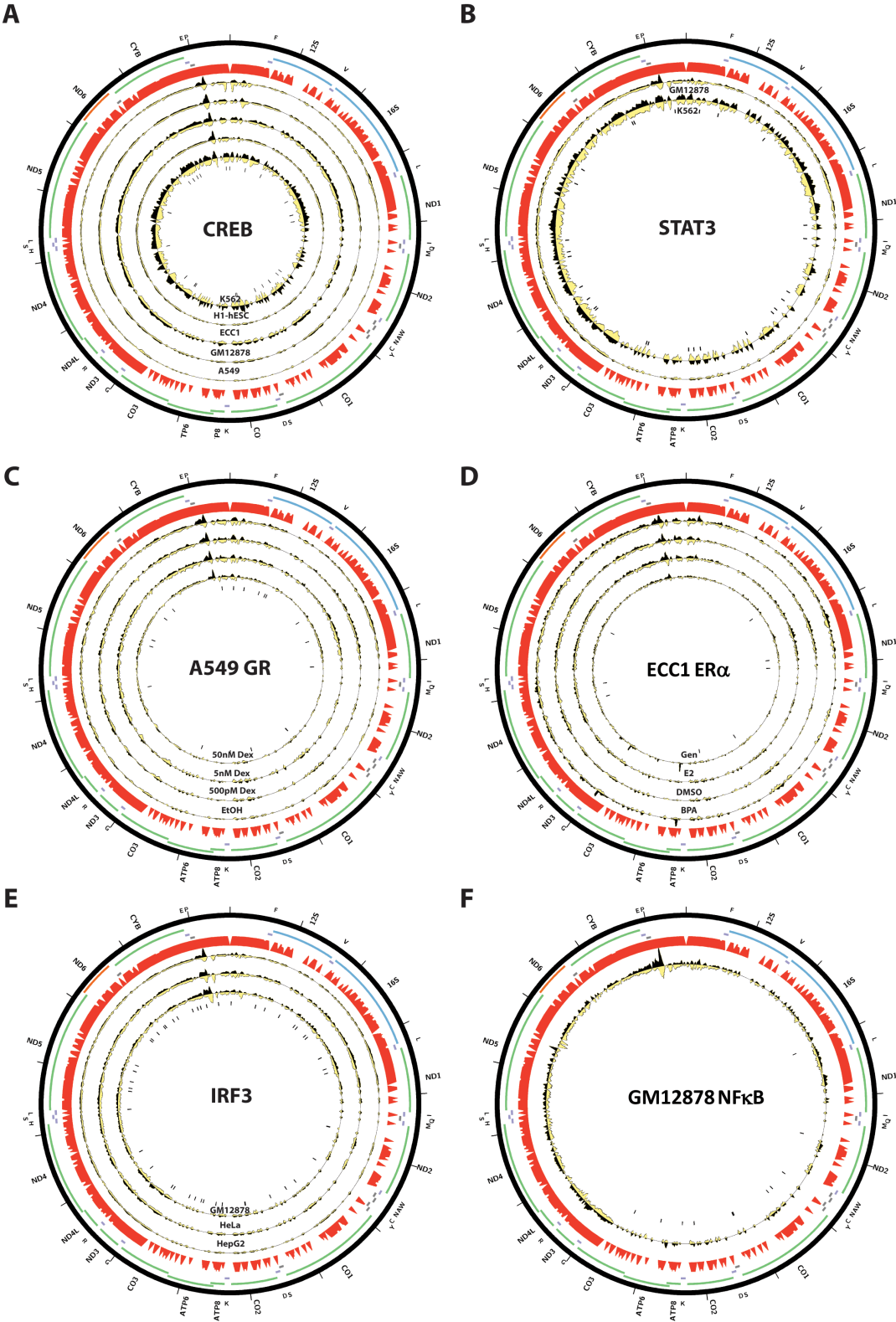
Figure 4.17.
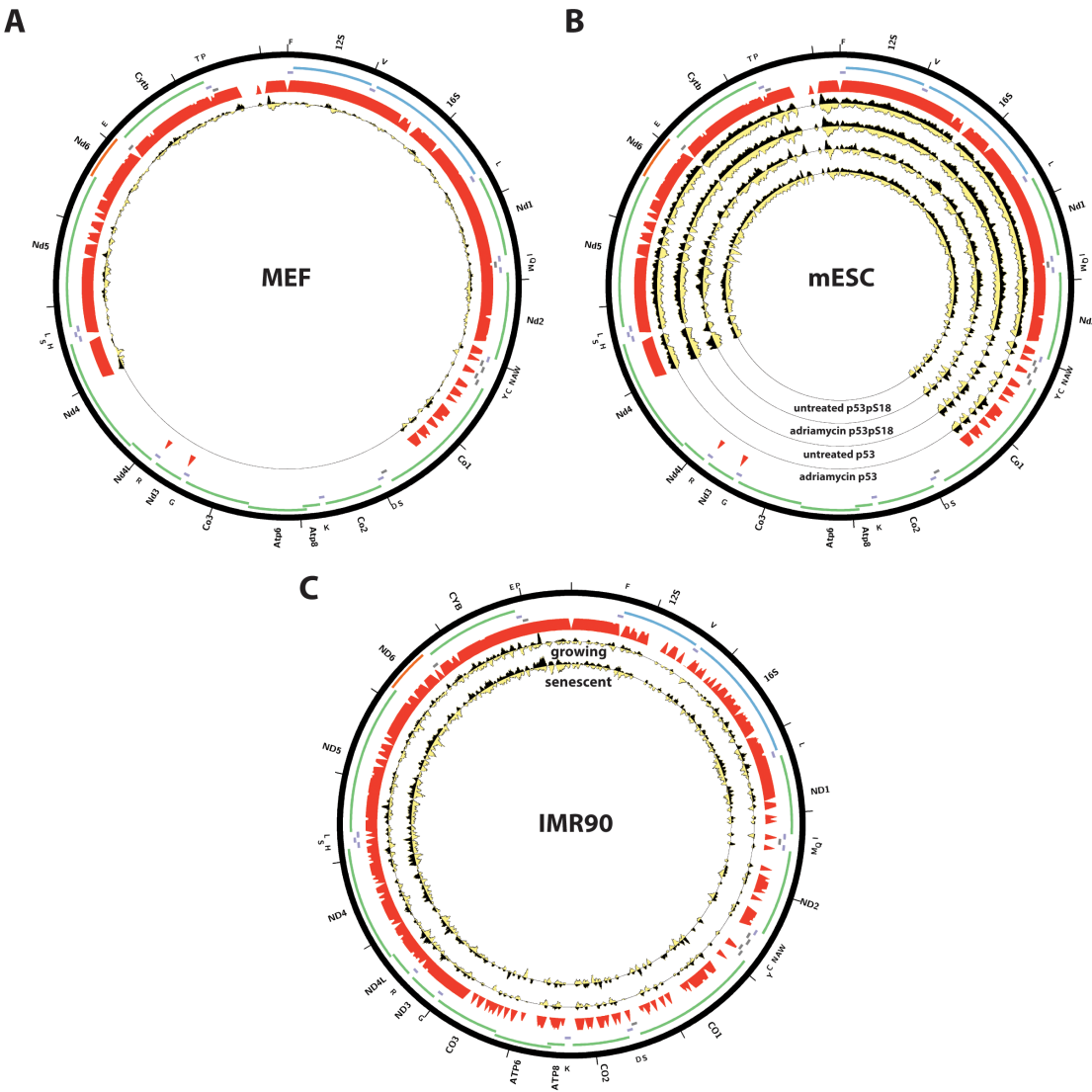
Figure 4.18.

Figure 4.19.
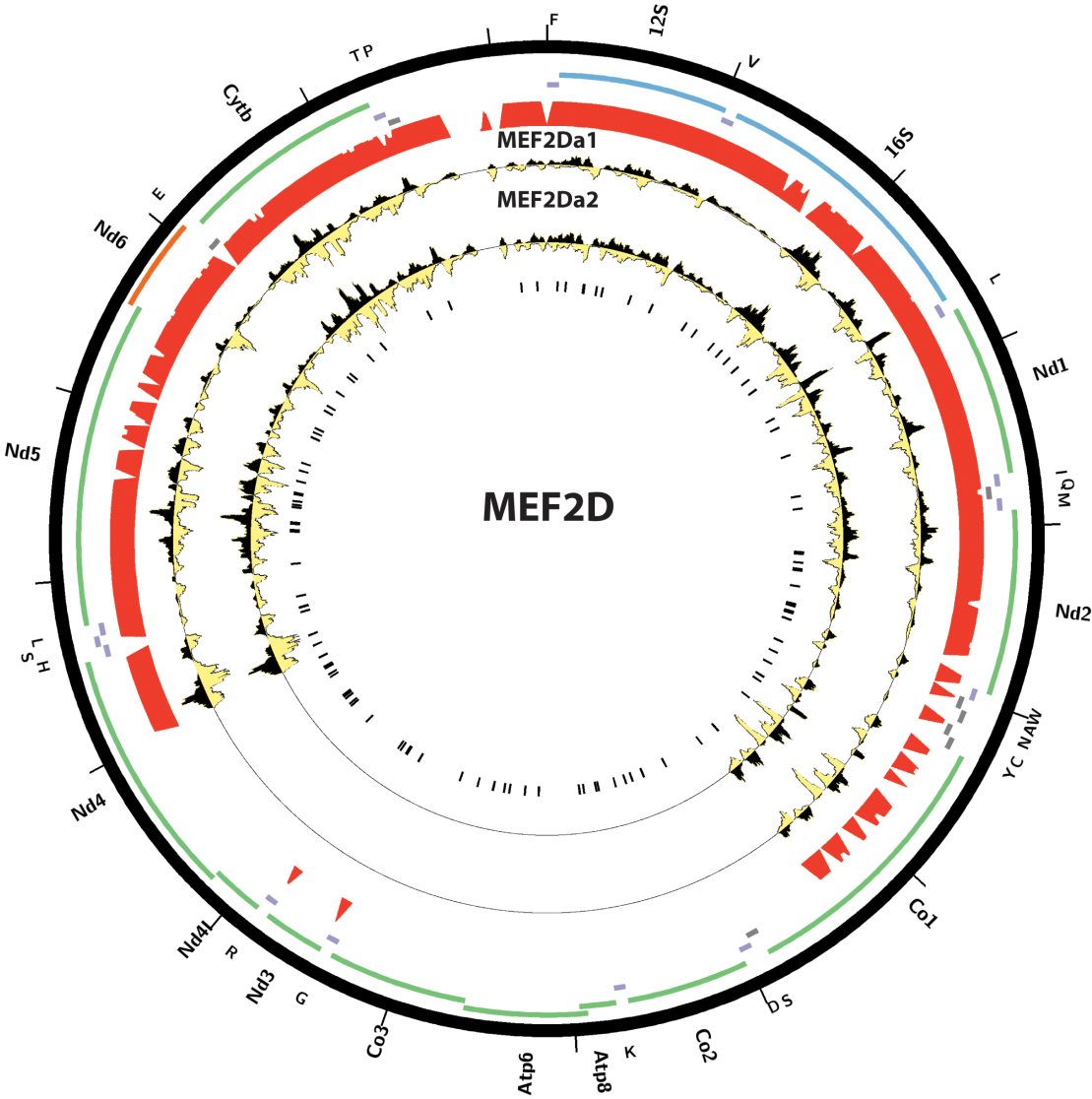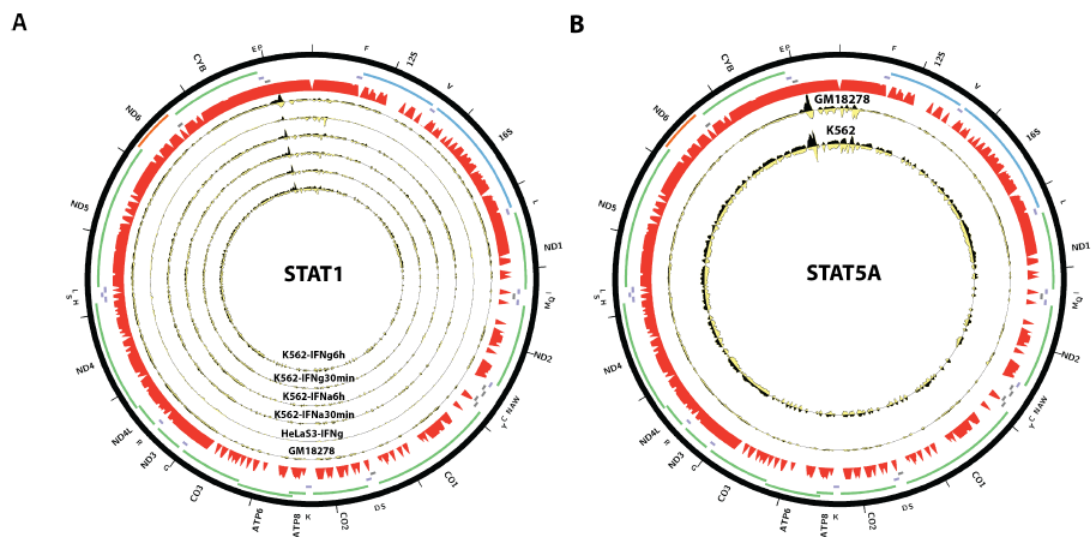
Table 4.1.

Gene	Variant	TargetP Probability	MitoProt Probability
Human			
CEBPb		0.091	0.0413
c-Jun		0.056	0.1973
JunD		0.147	0.0258
MafK		0.108	0.0167
MafF	1	0.112	0.0574
	2	0.112	0.0574
	3	0.112	0.0574
	4	0.112	0.0574
	5	0.635	0.1132
Max	1	0.066	0.0041
	2	0.064	0.0112
	3	0.064	0.0019
	4	0.055	0.0017
	5	0.068	0.0044
	6	0.057	0.0088
	7	0.071	0.0251
NFE2	1/2/3	0.282	0.1996
RFX5	1	0.092	0.0185
	2	0.092	0.0185
Mouse			
MafK		0.108	0.0124
MAX		0.064	0.0167
USF2	1	0.059	0.0484
	2	0.056	0.0563

BIBLIOGRAPHY

Achanta, G., R. Sasaki, L. Feng, J. S. Carew, W. Lu, H. Pelicano, M. J. Keating and P. Huang (2005). "Novel role of p53 in maintaining mitochondrial genetic stability through interaction with DNA Pol gamma." EMBO J **24**(19): 3482-3492.

Aksoy, O., A. Chicas, T. Zeng, Z. Zhao, M. McCurrach, X. Wang and S. W. Lowe (2012). "The atypical E2F family member E2F7 couples the p53 and RB pathways during cellular senescence." Genes Dev **26**(14): 1546-1557.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. Smith, R. Staden and I. G. Young (1981). "Sequence and organization of the human mitochondrial genome." Nature **290**(5806): 457-465.

Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li and W. S. Noble (2009). "MEME SUITE: tools for motif discovery and searching." Nucleic Acids Res **37**(Web Server issue): W202-208.

Bibb, M. J., R. A. Van Etten, C. T. Wright, M. W. Walberg and D. A. Clayton (1981). "Sequence and gene organization of mouse mitochondrial DNA." Cell **26**(2 Pt 2): 167-180.

Boengler, K., D. Hilfiker-Kleiner, G. Heusch and R. Schulz (2010). "Inhibition of permeability transition pore opening by mitochondrial STAT3 and its role in myocardial ischemia/reperfusion." Basic Res Cardiol **105**(6): 771-785.

Bogenhagen, D. and D. A. Clayton (1974). "The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid." J Biol Chem **249**(24): 7991-7995.

Cammarota, M., G. Paratcha, L. R. Bevilaqua, M. Levi de Stein, M. Lopez, A. Pellegrino de Iraldi, I. Izquierdo and J. H. Medina (1999). "Cyclic AMP-responsive element binding protein in brain mitochondria." J Neurochem **72**(6): 2272-2277.

Cantatore, P. and G. Attardi (1980). "Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA." Nucleic Acids Res **8**(12): 2605-2625.

Casas, F., L. Domenjoud, P. Rochard, R. Hatier, A. Rodier, L. Daury, A. Bianchi, P. Kremarik-Bouillaud, P. Becuwe, J. Keller, H. Schohn, C. Wrutniak-Cabello, G. Cabello and M. Dauca (2000). "A 45 kDa protein related to PPARgamma2, induced by peroxisome proliferators, is located in the mitochondrial matrix." FEBS Lett **478**(1-2): 4-8.

Casas, F., P. Rochard, A. Rodier, I. Cassar-Malek, S. Marchal-Victorion, R. J. Wiesner, G. Cabello and C. Wrutniak (1999). "A variant form of the nuclear triiodothyronine receptor c-ErbAalpha1 plays a direct role in regulation of mitochondrial RNA synthesis." Mol Cell Biol **19**(12): 7913-7924.

Chen, J. Q., M. Delannoy, C. Cooke and J. D. Yager (2004). "Mitochondrial localization of ERalpha and ERbeta in human MCF7 cells." Am J Physiol Endocrinol Metab **286**(6): E1011-1022.

Chueh, F. Y., K. F. Leong and C. L. Yu (2010). "Mitochondrial translocation of signal transducer and activator of transcription 5 (STAT5) in leukemic T cells and cytokine-stimulated cells." Biochem Biophys Res Commun **402**(4): 778-783.

Claros, M. G. and P. Vincens (1996). "Computational method to predict mitochondrially imported proteins and their targeting sequences." Eur J Biochem **241**(3): 779-786.

Cogswell, P. C., D. F. Kashatus, J. A. Keifer, D. C. Guttridge, J. Y. Reuther, C. Bristow, S. Roy, D. W. Nicholson and A. S. Baldwin, Jr. (2003). "NF-kappa B and I kappa B alpha are found in the mitochondria. Evidence for regulation of mitochondrial gene expression by NF-kappa B." J Biol Chem **278**(5): 2963-2968.

Consortium, E. P. (2011). "A user's guide to the encyclopedia of DNA elements (ENCODE)." PLoS Biol **9**(4): e1001046.

Consortium, E. P., B. E. Bernstein, E. Birney, I. Dunham, E. D. Green, C. Gunter and M. Snyder (2012). "An integrated encyclopedia of DNA elements in the human genome." Nature **489**(7414): 57-74.

De Rasmo, D., A. Signorile, E. Roca and S. Papa (2009). "cAMP response element-binding protein (CREB) is imported into mitochondria and promotes protein synthesis." FEBS J **276**(16): 4325-4333.

Demonacos, C., R. Djordjevic-Markovic, N. Tsawdaroglou and C. E. Sekeris (1995). "The mitochondrion as a primary site of action of glucocorticoids: the interaction of the glucocorticoid receptor with mitochondrial DNA sequences showing partial similarity to the nuclear glucocorticoid responsive elements." J Steroid Biochem Mol Biol **55**(1): 43-55.

Demonacos, C., N. C. Tsawdaroglou, R. Djordjevic-Markovic, M. Papalopoulou, V. Galanopoulos, S. Papadogeorgaki and C. E. Sekeris (1993). "Import of the glucocorticoid receptor into rat liver mitochondria in vivo and in vitro." J Steroid Biochem Mol Biol **46**(3): 401-413.

du Buy, H. G. and F. L. Riley (1967). "HYBRIDIZATION BETWEEN THE NUCLEAR AND KINETOPLAST DNA'S OF Leishmania enriettii AND BETWEEN NUCLEAR AND MITOCHONDRIAL DNA'S OF MOUSE LIVER." Proc Natl Acad Sci U S A **57**(3): 790-797.

Emanuelsson, O., S. Brunak, G. von Heijne and H. Nielsen (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." Nat Protoc **2**(4): 953-971.

Enriquez, J. A., P. Fernandez-Silva, N. Garrido-Perez, M. J. Lopez-Perez, A. Perez-Martos and J. Montoya (1999). "Direct regulation of mitochondrial RNA synthesis by thyroid hormone." Mol Cell Biol **19**(1): 657-670.

Enriquez, J. A., P. Fernandez-Silva and J. Montoya (1999). "Autonomous regulation in mammalian mitochondrial DNA transcription." Biol Chem **380**(7-8): 737-747.

Falkenberg, M., M. Gaspari, A. Rantanen, A. Trifunovic, N. G. Larsson and C. M. Gustafsson (2002). "Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA." Nat Genet **31**(3): 289-294.

Fisher, R. P. and D. A. Clayton (1985). "A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro." J Biol Chem **260**(20): 11330-11338.

Fisher, R. P. and D. A. Clayton (1988). "Purification and characterization of human mitochondrial transcription factor 1." Mol Cell Biol **8**(8): 3496-3509.

Fisher, R. P., T. Lisowsky, M. A. Parisi and D. A. Clayton (1992). "DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein." J Biol Chem **267**(5): 3358-3367.

Gaspari, M., N. G. Larsson and C. M. Gustafsson (2004). "The transcription machinery in mammalian mitochondria." Biochim Biophys Acta **1659**(2-3): 148-152.

Gerstein, M. B., Z. J. Lu, E. L. Van Nostrand, C. Cheng, B. I. Arshinoff, T. Liu, K. Y. Yip, R. Robilotto, A. Rechtsteiner, K. Ikegami, P. Alves, A. Chateigner, M. Perry, M. Morris, R. K. Auerbach, X. Feng, J. Leng, A. Vielle, W. Niu, K. Rhrissorrakrai, A. Agarwal, R. P. Alexander, G. Barber, C. M. Brdlik, J. Brennan, J. J. Brouillet, A. Carr, M. S. Cheung, H. Clawson, S. Contrino, L. O. Dannenberg, A. F. Dernburg, A. Desai, L. Dick, A. C. Dose, J. Du, T. Egelhofer, S. Ercan, G. Euskirchen, B. Ewing, E. A.

Feingold, R. Gassmann, P. J. Good, P. Green, F. Gullier, M. Gutwein, M. S. Guyer, L. Habegger, T. Han, J. G. Henikoff, S. R. Henz, A. Hinrichs, H. Holster, T. Hyman, A. L. Iniguez, J. Janette, M. Jensen, M. Kato, W. J. Kent, E. Kephart, V. Khivansara, E. Khurana, J. K. Kim, P. Kolasinska-Zwierz, E. C. Lai, I. Latorre, A. Leahey, S. Lewis, P. Lloyd, L. Lochovsky, R. F. Lowdon, Y. Lubling, R. Lyne, M. MacCoss, S. D. Mackowiak, M. Mangone, S. McKay, D. Mecenas, G. Merrihew, D. M. Miller, 3rd, A. Muroyama, J. I. Murray, S. L. Ooi, H. Pham, T. Phippen, E. A. Preston, N. Rajewsky, G. Ratsch, H. Rosenbaum, J. Rozowsky, K. Rutherford, P. Ruzanov, M. Sarov, R. Sasidharan, A. Sboner, P. Scheid, E. Segal, H. Shin, C. Shou, F. J. Slack, C. Slightam, R. Smith, W. C. Spencer, E. O. Stinson, S. Taing, T. Takasaki, D. Vafeados, K. Voronina, G. Wang, N. L. Washington, C. M. Whittle, B. Wu, K. K. Yan, G. Zeller, Z. Zha, M. Zhong, X. Zhou, E. C. mod, J. Ahringer, S. Strome, K. C. Gunsalus, G. Micklem, X. S. Liu, V. Reinke, S. K. Kim, L. W. Hillier, S. Henikoff, F. Piano, M. Snyder, L. Stein, J. D. Lieb and R. H. Waterston (2010). "Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project." Science **330**(6012): 1775-1787.

Gertz, J., T. E. Reddy, K. E. Varley, M. J. Garabedian and R. M. Myers (2012). "Genistein and bisphenol A exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner." Genome Res **22**(11): 2153-2162.

Gough, D. J., A. Corlett, K. Schlessinger, J. Wegrzyn, A. C. Larner and D. E. Levy (2009). "Mitochondrial STAT3 supports Ras-dependent oncogenic transformation." Science **324**(5935): 1713-1716.

Hazkani-Covo, E., R. M. Zeller and W. Martin (2010). "Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes." PLoS Genet **6**(2): e1000834.

Heyne, K., S. Mannebach, E. Wuertz, K. X. Knaup, M. Mahyar-Roemer and K. Roemer (2004). "Identification of a putative p53 binding sequence within the human mitochondrial genome." FEBS Lett **578**(1-2): 198-202.

Johnson, D. S., A. Mortazavi, R. M. Myers and B. Wold (2007). "Genome-wide mapping of in vivo protein-DNA interactions." Science **316**(5830): 1497-1502.

Johnson, R. F., Witzel, II and N. D. Perkins (2011). "p53-dependent regulation of mitochondrial energy production by the RelA subunit of NF-kappaB." Cancer Res **71**(16): 5588-5597.

Kasowski, M., F. Grubert, C. Heffelfinger, M. Hariharan, A. Asabere, S. M. Waszak, L. Habegger, J. Rozowsky, M. Shi, A. E. Urban, M. Y. Hong, K. J. Karczewski, W. Huber, S. M. Weissman, M. B. Gerstein, J. O. Korbel and M. Snyder (2010). "Variation in transcription factor binding among humans." Science **328**(5975): 232-235.

Kenzelmann Broz, D., S. Spano Mello, K. T. Bieging, D. Jiang, R. L. Dusek, C. A. Brady, A. Sidow and L. D. Attardi (2013). "Global genomic profiling reveals an extensive p53-regulated autophagy program contributing to key p53 responses." Genes Dev **27**(9): 1016-1031.

Koufali, M. M., P. Moutsatsou, C. E. Sekeris and K. C. Breen (2003). "The dynamic localization of the glucocorticoid receptor in rat C6 glioma cell mitochondria." Mol Cell Endocrinol **209**(1-2): 51-60.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome Biol **10**(3): R25.

Lee, J., C. H. Kim, D. K. Simon, L. R. Aminova, A. Y. Andreyev, Y. E. Kushnareva, A. N. Murphy, B. E. Lonze, K. S. Kim, D. D. Ginty, R. J. Ferrante, H. Ryu and R. R. Ratan (2005). "Mitochondrial cyclic AMP response element-binding protein (CREB) mediates mitochondrial gene expression and neuronal survival." J Biol Chem **280**(49): 40398-40401.

Leigh-Brown, S., J. A. Enriquez and D. T. Odom (2010). "Nuclear transcription factors in mammalian mitochondria." Genome Biol **11**(7): 215.

Li, M., Y. He, W. Dubois, X. Wu, J. Shi and J. Huang (2012). "Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells." Mol Cell **46**(1): 30-42.

Marchenko, N. D., A. Zaika and U. M. Moll (2000). "Death signal-induced localization of p53 protein to mitochondria. A potential role in apoptotic signaling." J Biol Chem **275**(21): 16202-16212.

Marinov, G. K., Y. E. Wang, D. Chan and B. J. Wold (2014). "Evidence for site-specific occupancy of the mitochondrial genome by nuclear transcription factors." PLoS One **9**(1): e84713.

Metodiev, M. D., N. Lesko, C. B. Park, Y. Camara, Y. Shi, R. Wibom, K. Hultenby, C. M. Gustafsson and N. G. Larsson (2009). "Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome." Cell Metab **9**(4): 386-397.

mod, E. C., S. Roy, J. Ernst, P. V. Kharchenko, P. Kheradpour, N. Negre, M. L. Eaton, J. M. Landolin, C. A. Bristow, L. Ma, M. F. Lin, S. Washietl, B. I. Arshinoff, F. Ay, P. E. Meyer, N. Robine, N. L. Washington, L. Di Stefano, E. Berezikov, C. D. Brown, R. Candeias, J. W. Carlson, A. Carr, I. Jungreis, D. Marbach, R. Sealfon, M. Y. Tolstorukov, S. Will, A. A. Alekseyenko, C. Artieri, B. W. Booth, A. N. Brooks, Q. Dai, C. A. Davis, M. O. Duff, X. Feng, A. A. Gorchakov, T. Gu, J. G. Henikoff, P. Kapranov, R. Li, H. K. MacAlpine, J. Malone, A. Minoda, J. Nordman, K. Okamura, M. Perry, S. K. Powell, N. C. Riddle, A. Sakai, A. Samsonova, J. E. Sandler, Y. B. Schwartz, N. Sher, R. Spokony, D. Sturgill, M. van Baren, K. H. Wan, L. Yang, C. Yu, E. Feingold, P. Good, M. Guyer, R. Lowdon, K. Ahmad, J. Andrews, B. Berger, S. E. Brenner, M. R. Brent, L. Cherbas, S. C. Elgin, T. R. Gingeras, R. Grossman, R. A. Hoskins, T. C. Kaufman, W. Kent, M. I. Kuroda, T. Orr-Weaver, N. Perrimon, V. Pirrotta, J. W. Posakony, B. Ren, S. Russell, P. Cherbas, B. R. Graveley, S. Lewis, G. Micklem, B. Oliver, P. J. Park, S. E. Celniker, S. Henikoff, G. H. Karpen, E. C. Lai, D. M. MacAlpine, L. D. Stein, K. P.

White and M. Kellis (2010). "Identification of functional elements and regulatory circuits by Drosophila modENCODE." Science **330**(6012): 1787-1797.

Monje, P. and R. Boland (2001). "Subcellular distribution of native estrogen receptor alpha and beta isoforms in rabbit uterus and ovary." J Cell Biochem **82**(3): 467-479.

Montoya, J., T. Christianson, D. Levens, M. Rabinowitz and G. Attardi (1982). "Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA." Proc Natl Acad Sci U S A **79**(23): 7195-7199.

Mortazavi, A., E. C. Leeper Thompson, S. T. Garcia, R. M. Myers and B. Wold (2006). "Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire." Genome Res **16**(10): 1208-1221.

Mouse, E. C., J. A. Stamatoyannopoulos, M. Snyder, R. Hardison, B. Ren, T. Gingeras, D. M. Gilbert, M. Groudine, M. Bender, R. Kaul, T. Canfield, E. Giste, A. Johnson, M. Zhang, G. Balasundaram, R. Byron, V. Roach, P. J. Sabo, R. Sandstrom, A. S. Stehling, R. E. Thurman, S. M. Weissman, P. Cayting, M. Hariharan, J. Lian, Y. Cheng, S. G. Landt, Z. Ma, B. J. Wold, J. Dekker, G. E. Crawford, C. A. Keller, W. Wu, C. Morrissey, S. A. Kumar, T. Mishra, D. Jain, M. Byrska-Bishop, D. Blankenberg, B. R. Lajoie, G. Jain, A. Sanyal, K. B. Chen, O. Denas, J. Taylor, G. A. Blobel, M. J. Weiss, M. Pimkin, W. Deng, G. K. Marinov, B. A. Williams, K. I. Fisher-Aylor, G. Desalvo, A. Kiralusha, D. Trout, H. Amrhein, A. Mortazavi, L. Edsall, D. McCleary, S. Kuan, Y. Shen, F. Yue, Z. Ye, C. A. Davis, C. Zaleski, S. Jha, C. Xue, A. Dobin, W. Lin, M. Fastuca, H. Wang, R. Guigo, S. Djebali, J. Lagarde, T. Ryba, T. Sasaki, V. S. Malladi, M. S. Cline, V. M.

Kirkup, K. Learned, K. R. Rosenbloom, W. J. Kent, E. A. Feingold, P. J. Good, M. Pazin, R. F. Lowdon and L. B. Adams (2012). "An encyclopedia of mouse DNA elements (Mouse ENCODE)." Genome Biol **13**(8): 418.

Nass, S., M. M. Nass and U. Hennix (1965). "Deoxyribonucleic Acid in Isolated Rat-Liver Mitochondria." Biochim Biophys Acta **95**: 426-435.

Neph, S., J. Vierstra, A. B. Stergachis, A. P. Reynolds, E. Haugen, B. Vernot, R. E. Thurman, S. John, R. Sandstrom, A. K. Johnson, M. T. Maurano, R. Humbert, E. Rynes, H. Wang, S. Vong, K. Lee, D. Bates, M. Diegel, V. Roach, D. Dunn, J. Neri, A. Schafer, R. S. Hansen, T. Kutyavin, E. Giste, M. Weaver, T. Canfield, P. Sabo, M. Zhang, G. Balasundaram, R. Byron, M. J. MacCoss, J. M. Akey, M. A. Bender, M. Groudine, R. Kaul and J. A. Stamatoyannopoulos (2012). "An expansive human regulatory lexicon encoded in transcription factor footprints." Nature **489**(7414): 83-90.

Ogita, K., Y. Fujinami, M. Kitano and Y. Yoneda (2003). "Transcription factor activator protein-1 expressed by kainate treatment can bind to the non-coding region of mitochondrial genome in murine hippocampus." J Neurosci Res **73**(6): 794-802.

Ogita, K., H. Okuda, M. Kitano, Y. Fujinami, K. Ozaki and Y. Yoneda (2002). "Localization of activator protein-1 complex with DNA binding activity in mitochondria of murine brain after in vivo treatment with kainate." J Neurosci **22**(7): 2561-2570.

Psarra, A. M., S. Solakidi and C. E. Sekeris (2006). "The mitochondrion as a primary site of action of steroid and thyroid hormones: presence and action of steroid and thyroid

hormone receptors in mitochondria of animal cells." Mol Cell Endocrinol **246**(1-2): 21-33.

Reddy, T. E., J. Gertz, G. E. Crawford, M. J. Garabedian and R. M. Myers (2012). "The hypersensitive glucocorticoid response specifically regulates period 1 and expression of circadian genes." Mol Cell Biol **32**(18): 3756-3767.

Reddy, T. E., F. Pauli, R. O. Sprouse, N. F. Neff, K. M. Newberry, M. J. Garabedian and R. M. Myers (2009). "Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation." Genome Res **19**(12): 2163-2171.

Ryu, H., J. Lee, S. Impey, R. R. Ratan and R. J. Ferrante (2005). "Antioxidants modulate mitochondrial PKA and increase CREB binding to D-loop DNA of the mitochondrial genome in neurons." Proc Natl Acad Sci U S A **102**(39): 13915-13920.

Sagan, L. (1967). "On the origin of mitosing cells." J Theor Biol **14**(3): 255-274.

Satoh, M. and T. Kuroiwa (1991). "Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell." Exp Cell Res **196**(1): 137-140.

Sebastian, S., H. Faralli, Z. Yao, P. Rakopoulos, C. Palii, Y. Cao, K. Singh, Q. C. Liu, A. Chu, A. Aziz, M. Brand, S. J. Tapscott and F. J. Dilworth (2013). "Tissue-specific splicing of a ubiquitously expressed transcription factor is essential for muscle differentiation." Genes Dev **27**(11): 1247-1259.

Shadel, G. S. and D. A. Clayton (1997). "Mitochondrial DNA maintenance in vertebrates." <u>Annu Rev Biochem</u> **66**: 409-435.

She, H., Q. Yang, K. Shepherd, Y. Smith, G. Miller, C. Testa and Z. Mao (2011). "Direct regulation of complex I by mitochondrial MEF2D is disrupted in a mouse model of Parkinson disease and in human patients." <u>J Clin Invest</u> **121**(3): 930-940.

Shulga, N. and J. G. Pastorino (2012). "GRIM-19-mediated translocation of STAT3 to mitochondria is necessary for TNF-induced necroptosis." <u>J Cell Sci</u> **125**(Pt 12): 2995-3003.

Shutt, T. E., M. Bestwick and G. S. Shadel (2011). "The core human mitochondrial transcription initiation complex: It only takes two to tango." <u>Transcription</u> **2**(2): 55-59.

Song, L., Z. Zhang, L. L. Grasfeder, A. P. Boyle, P. G. Giresi, B. K. Lee, N. C. Sheffield, S. Graf, M. Huss, D. Keefe, Z. Liu, D. London, R. M. McDaniell, Y. Shibata, K. A. Showers, J. M. Simon, T. Vales, T. Wang, D. Winter, Z. Zhang, N. D. Clarke, E. Birney, V. R. Iyer, G. E. Crawford, J. D. Lieb and T. S. Furey (2011). "Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity." <u>Genome Res</u> **21**(10): 1757-1767.

Szczepanek, K., Q. Chen, M. Derecka, F. N. Salloum, Q. Zhang, M. Szelag, J. Cichy, R. C. Kukreja, J. Dulak, E. J. Lesnefsky and A. C. Larner (2011). "Mitochondrial-targeted Signal transducer and activator of transcription 3 (STAT3) protects against ischemia-

induced changes in the electron transport chain and the generation of reactive oxygen species." J Biol Chem **286**(34): 29610-29620.

Szczepanek, K., Q. Chen, A. C. Larner and E. J. Lesnefsky (2012). "Cytoprotection by the modulation of mitochondrial electron transport chain: the emerging role of mitochondrial STAT3." Mitochondrion **12**(2): 180-189.

Szklarczyk, R. and M. A. Huynen (2010). "Mosaic origin of the mitochondrial proteome." Proteomics **10**(22): 4012-4024.

Thurman, R. E., E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, N. C. Sheffield, A. B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T. K. Canfield, M. Diegel, D. Dunn, A. K. Ebersol, T. Frum, E. Giste, A. K. Johnson, E. M. Johnson, T. Kutyavin, B. Lajoie, B. K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E. D. Nguyen, H. Qu, A. P. Reynolds, V. Roach, A. Safi, M. E. Sanchez, A. Sanyal, A. Shafer, J. M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, Z. Zhang, B. Lenhard, M. Tewari, M. O. Dorschner, R. S. Hansen, P. A. Navas, G. Stamatoyannopoulos, V. R. Iyer, J. D. Lieb, S. R. Sunyaev, J. M. Akey, P. J. Sabo, R. Kaul, T. S. Furey, J. Dekker, G. E. Crawford and J. A. Stamatoyannopoulos (2012). "The accessible chromatin landscape of the human genome." Nature **489**(7414): 75-82.

Wang, J., J. Zhuang, S. Iyer, X. Lin, T. W. Whitfield, M. C. Greven, B. G. Pierce, X. Dong, A. Kundaje, Y. Cheng, O. J. Rando, E. Birney, R. M. Myers, W. S. Noble, M. Snyder and Z. Weng (2012). "Sequence features and chromatin structure around the

genomic regions bound by 119 human transcription factors." <u>Genome Res</u> **22**(9): 1798-1812.

Wang, Y. E., G. K. Marinov, B. J. Wold and D. C. Chan (2013). "Genome-wide analysis reveals coating of the mitochondrial genome by TFAM." <u>PLoS One</u> **8**(8): e74513.

Wegrzyn, J., R. Potla, Y. J. Chwae, N. B. Sepuri, Q. Zhang, T. Koeck, M. Derecka, K. Szczepanek, M. Szelag, A. Gornicka, A. Moh, S. Moghaddas, Q. Chen, S. Bobbili, J. Cichy, J. Dulak, D. P. Baker, A. Wolfman, D. Stuehr, M. O. Hassan, X. Y. Fu, N. Avadhani, J. I. Drake, P. Fawcett, E. J. Lesnefsky and A. C. Larner (2009). "Function of mitochondrial Stat3 in cellular respiration." <u>Science</u> **323**(5915): 793-797.

Williams, R. S. (1986). "Mitochondrial gene expression in mammalian striated muscle. Evidence that variation in gene dosage is the major regulatory event." <u>J Biol Chem</u> **261**(26): 12390-12394.

Wrutniak, C., I. Cassar-Malek, S. Marchal, A. Rascle, S. Heusser, J. M. Keller, J. Flechon, M. Dauca, J. Samarut, J. Ghysdael and et al. (1995). "A 43-kDa protein related to c-Erb A alpha 1 is located in the mitochondrial matrix of rat liver." <u>J Biol Chem</u> **270**(27): 16347-16354.

Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen and C. R. Woese (1985). "Mitochondrial origins." <u>Proc Natl Acad Sci U S A</u> **82**(13): 4443-4447.

Yoshida, Y., H. Izumi, T. Torigoe, H. Ishiguchi, H. Itoh, D. Kang and K. Kohno (2003). "P53 physically interacts with mitochondrial transcription factor A and differentially regulates binding to damaged DNA." Cancer Res **63**(13): 3729-3734.

Chapter 5

FUTURE DIRECTIONS

Maintenance of mtDNA stability is required for cellular and organismal viability; mutations in the genome lead to cellular respiratory defects, resulting in neuromuscular impairment, movement disorders, myopathy, and eye disorders (Zeviani and Di Donato 2004). Furthermore, loss of mtDNA is known to cause Alpers' Disease, ataxia neuropathy spectrum, progressive external ophthalmoplegia, and mitochondrial depletion syndrome (Suomalainen and Isohanni 2010, Stumpf, Saneto et al. 2013). mtDNA instability has also been implicated in cardiovascular disease (Sugiyama, Hattori et al. 1991, Ide, Tsutsui et al. 2001, Karamanlidis, Nascimben et al. 2010, Karamanlidis, Bautista-Hernandez et al. 2011), diabetes (Maassen, LM et al. 2004, Simmons, Suponitsky-Kroyter et al. 2005, Gauthier, Wiederkehr et al. 2009), neurodegenerative disorders such as Alzheimer's (Coskun, Beal et al. 2004, Coskun, Wyrembak et al. 2012, Rice, Keeney et al. 2014), and aging (Corral-Debrinski, Shoffner et al. 1992, Trifunovic and Larsson 2008). Therefore, it is essential to understand the mechanisms of mtDNA maintenance.

Mitochondrial Dynamics and mtDNA Copy Number

Work in our lab (Chen, Vermulst et al. 2010) has demonstrated that loss of mitochondrial fusion leads to not only loss of mtDNA copy number, but also of its integrity, with mutations and deletions accumulating in affected skeletal muscle. A recent study has tied the mitochondrial fission protein Drp1 to nucleoid morphology, where knockdown causes aggregation and misdistribution of the nucleoids (Ban-Ishihara, Ishihara et al. 2013). Although mtDNA copy number does not change, it is unclear if the incomplete loss of fission seen here is the reason for this. However, it is evident that mitochondrial fission and fusion proteins play an important role in the regulation of the genome. Further understanding of the pathways initiated when fusion and fission are abrogated will provide necessary insight into understanding how control of mtDNA maintenance is tied into the proteins regulating fusion and fission. One amenable approach is through proteomics analysis of cells with deleted fusion proteins via SILAC analysis, and comparison of the labeled mitochondrial proteome to that found in wildtype counterparts. The technique has previously been utilized effectively by our own lab (Chan, Salazar et al. 2011) to characterize changes in mitochondrial protein content upon CCCP treatment. If significant differential import or export of protein levels exists, SILAC will be able to detect this. An alternative, and perhaps more directed, method of identifying changes in the proteome specifically at the nucleoid, possibly with greater sensitivity given the lower number of targets, would be an approach analogous to that described by Bogenhagen et al., who have used both immunoaffinity purification using

anti-TFAM and anti-mtSSB antibodies, as well as purification of formaldehyde-crosslinked nucleoids exposed to harsh ionic detergents to decrease nonspecific binding in order to selectively isolate mtDNA-associated proteins for identification (Wang and Bogenhagen 2006, Bogenhagen, Rousseau et al. 2008). Proteins within the nucleoids isolated by either means can then be sequenced for identification. Although this method is more specific to the proteins directly involved with the nucleoid, it is certainly less quantitative and does not allow for accurate elucidation of changes in protein levels within the nucleoid. Furthermore, it is more susceptible to experimental variation; because of the need to compare two independent samples, this method requires additional attention to consistency of procedure. A comparison between wild-type and knockout cells is not ideal due to potential proteasomal differences between cell lines, and an RNAi – based assay would be more preferable in order to abrogate inter-cell line variability, and also allow for a time-course study to follow changes in the mitochondrial proteome as fusion proteins are depleted. Unfortunately, because levels of the fusion proteins must be extremely low before defects are seen, a knockdown-based model could be difficult to attain.

SILAC analysis of fusion protein-deficient cells, especially if performed as a time course after successful RNAi knockdown, could prove extremely informative in itself to understand how loss of fusion shifts protein levels. However, another potential result of these studies is the identification of novel proteins associated with the nucleoid, imported into the mitochondria, or absent from each upon fusion protein depletion that could then

be further characterized. Immunocytochemical analysis could confirm such proteomic results via localization studies, and subcellular fractionation followed by western blotting could confirm differential levels of each target within the mitochondria. Furthermore, because it is thought that some level of defect in fusion-deficient cells could be due to loss of distribution of mitochondrial contents and thus inter-mitochondrial protein heterogeneity (Chen, Vermulst et al. 2010), immunocytochemical analysis should be able to determine the level of this heterogeneity for each target. Finally, over-expression and knockdown studies of the target followed by rt-qPCR, qPCR, or western blotting could then elucidate the effect of the target on transcription, replication, or translation of mtDNA products.

Dynamics of the Mitochondrial Nucleoid

The mitochondrial nucleoid is normally composed of many proteins (Wang and Bogenhagen 2006, Bogenhagen, Rousseau et al. 2008). Of those, the transcriptional and replication machineries are perhaps best characterized (Falkenberg, Gaspari et al. 2002, Korhonen, Pham et al. 2004, Falkenberg, Larsson et al. 2007). We focused our analysis on TFAM, which is essential for mtDNA copy number maintenance. Our ChIP-seq data reveals the nonspecific binding nature of TFAM across the genome under resting conditions in HeLa cells, as well as a specific binding site upstream of the origin of light strand replication, previously unidentified in human mtDNA. Strikingly, we did not

observe increased read intensity in the known binding sites in LSP and HSP. These unexpected observations call for a revision of the known model of TFAM; under these cellular conditions, TFAM does appear not bind to its sites of transcriptional activation, or if it does, that it is a transient or low-probability event. Furthermore, a novel site with higher TFAM occupancy has been identified, one whose function is thus far unknown. Currently, TFAM is tied to replication of the genome through the requirement for an RNA primer for initiation (Lee and Clayton 1998). Due to its proximity to the origin of light strand replication, it can be postulated that this specific binding locus may play a direct role in replication. Further analysis of this site will be required to understand its significance. Specifically, biochemical assays may be performed to understand the function of TFAM binding to this site. FRET and SPR studies could be performed to understand the binding affinity and level of DNA bending by TFAM at this locus as compared to that for nonspecific sequences. Replication run-off assays may also be utilized to understand the role of TFAM in replication at this locus, although interpretation of results from such an experiment could be difficult due to the role of TFAM in transcription in facilitating the RNA primers required for replication, and the dependence of initiation of light strand replication on heavy strand synthesis (Holt and Reyes 2012). There are widely used cell-free replication assays utilizing cell lysate that appear to work for the mitochondrial system (Jui and Wong 1991, Rizwani and Chellappan 2009, Svitin and Chesnokov 2010), and which could be leveraged to study the effect of TFAM in the context of a relatively complete cellular system. Because TFAM knockout cells are not viable, wild-type lysate cannot be compared to that of a

knockout for replication activity. Instead, TFAM from wild-type mitochondrial lysate could be removed via immunoaffinity pull-down. Because it has been shown that POLRMT is sufficient for RNA primer formation at O_L (Wanrooij, Fuste et al. 2008), POLRMT would also need to be removed from the lysate and the lagging strand primer added to drive continued replication. Levels of replication, as measured by qPCR, in the absence of POLMRT, presence of the RNA primer, and presence or absence of TFAM should inform whether TFAM has a direct effect on replication from O_L.

Mitochondrial *in vitro* replication assays have also previously been utilized to characterize the minimum replisome (Wong and Clayton 1985, Korhonen, Pham et al. 2004), and could also be used to understand the role of TFAM at the O_L. Testing the role of TFAM in transcription at the O_L locus could be performed using well-known *in vitro* transcription assays. Decoupling TFAM's potential role in direct modulation of replication from its role in transcription of the primer may then be performed in an experiment where the RNA primer is provided along with the minimum replisome and the presence or absence of TFAM, similarly to the experiments using cell lysate. While the true *in vitro* replication system may yield "cleaner" experimental data, lack of other components of the mitochondrial proteome could preclude potential protein-protein interactions necessary for direct modulation of replication by TFAM. Thus, lack of effect in the *in vitro* system will not rule out the possibility that TFAM modulates replication through protein adapters.

Although we have identified a cell state where TFAM binds across the entire mitochondrial genome (Wang, Marinov et al. 2013), it is possible that under different conditions, the nucleoid could undergo remodeling, with differential TFAM binding patterns. Indeed, such remodeling is observed with the yeast homolog, Abf2p (Kucej, Kucejova et al. 2008), under different metabolic conditions. Given the homologous systems, it is reasonable to postulate that similar remodeling would occur if human cells were placed under different metabolic conditions. This could be tested via growth of cells in galactose-containing media as opposed to glucose-containing media, which drives cells towards reliance on oxidative phosphorylation because no net ATP is created when pyruvate is produced via glycolytic metabolism of galactose (Reitzer, Wice et al. 1979, Robinson, Petrova-Benedict et al. 1992, Marroquin, Hynes et al. 2007). Another state in which mtDNA-TFAM interactions could be altered is when active transcription and replication are occurring. While these activities generally occur continuously throughout the cell cycle (Bogenhagen and Clayton 1977, Magnusson, Orth et al. 2003) with some variation from nucleoid to nucleoid (Wai, Teoli et al. 2008), general upregulation of replication may be achieved by first depleting mtDNA and then allowing for recovery, presumably driving the mtDNA population towards a more uniform, actively replicative state. Such experiments were attempted in our lab using EtBr, which is well-established to cause mtDNA depletion (King and Attardi 1996), followed by recovery during which TFAM levels recover more slowly than mtDNA levels (Seidel-Rogol and Shadel 2002), implying differential binding characteristics. Treatment of HeLa cells resulted in rapid mtDNA and TFAM depletion and striking changes in nucleoid morphology, with

nucleoids consolidating into extremely large super-nucleoids, suggesting differential nucleoid structure. However, evidence that TFAM-DNA interactions were modified during the recovery phase were inconsistent, with some experiments showed a strong, highly specific binding of TFAM at distinct loci, and others demonstrated nonspecific binding across the entire genome as seen under normal conditions. It is possible that the state of TFAM-DNA interactions is highly sensitive to cellular growth conditions and is highly transient, thus rendering capture of any differential state more difficult. Further ChIP-seq studies will need to be performed to fully elucidate the dynamics of binding under highly replicative conditions.

It has recently been shown that TFAM may bind differentially in aged rat frontal cortex, soleus muscle, and liver as compared to young and calorie-restricted aged rats (Picca, Fracasso et al. 2013, Picca, Pesce et al. 2013, Picca, Pesce et al. 2014) via analysis of 5 specific regions of the genome. However, the far more powerful tool is our ChIP-seq analysis, which allows for much greater resolution across the entire genome, and such an experiment could readily be performed using rat tissue samples. Interestingly, Picca et al. assayed an area proximal to the O_L and found decreased occupancy in cortex but increased occupancy in the soleus and liver in aged rats. Relative amounts of TFAM were increased, in contrast to the decreased TFAM amounts in soleus and liver. Thus, it appears that there is some dynamics of binding at O_L as well as elsewhere in the genome that could be further characterized with ChIP-seq analysis. In addition, similar ChIP-seq experiments may be performed to understand differential TFAM binding in other

pathologies where mtDNA instability has been implicated previously, such as mammalian models of cardiovascular disease, diabetes, and Alzheimer's disease. Although TFAM has not been directly implicated in causing any of these pathologies, that TFAM is capable of rescuing phenotypic defects upon overexpression suggests that mis-regulation of TFAM could be a phenotype of the diseases.

Beyond the current TFAM ChIP-seq analysis, it is evident that such experiments will be an important tool for understanding how other proteins interact with the mitochondrial genome. ChIP-seq analysis such as that described in Chapter 2 could be utilized for many of the core components of the nucleoid in order to better understand binding patterns during many cell states. Key targets of interest include: Opa1, for which there is some evidence that it acts as an anchor for the mitochondrial nucleoid to the inner mitochondrial membrane (Elachouri, Vidoni et al. 2011); ATAD3, which has been implicated in regulation of nucleoid morphology and may bind to the D-loop (He, Mao et al. 2007); the prohibitins, which have also been implicated in nucleoid morphology and genome stabilization (Merkwirth, Martinelli et al. 2012); the Lon protease, which is known to bind to the D-loop, hypothetically to remain close to TFAM to effect rapid degradation as needed (Liu, Lu et al. 2004, Lu, Yadav et al. 2007); mTERF1, to extend our current understanding of transcriptional termination beyond its already-known site of high occupancy (Kruse, Narasimhan et al. 1989) toward additional sites (Hyvarinen, Pohjoismaki et al. 2007); mTERF2 and mTERF3, which have been shown to bind to the D-loop and appear to modulate transcription (Wenz, Luca et al. 2009); TFB1M, TFB2M,

RNA polymerase, Twinkle, and DNA polymerase, in order to understand potentially novel start sites as well as pause sites across the genome to better understand the dynamics of transcription and replication. While there are other core nucleoid proteins that may also be of interest (Bogenhagen, Rousseau et al. 2008), ChIP-seq data for binding by these targets is likely to lead to better understanding of the regulation of the genome, supplementing existing biochemical characterization studies.

Finally, there is some evidence for localization of classically nuclear transcription factors to mitochondria and associating with mtDNA (Leigh-Brown, Enriquez et al. 2010). We identified both human and mouse transcription factors exhibiting canonical ChIP-seq peaks. None of these peaks had previously been identified as areas of protein localization by the transcriptional machinery; however, further ChIP-seq analysis of the transcriptional machinery, as mentioned above, could aid in determining the underlying rationale for the noncanonical localization of these TFs outside of the NCR. The main mitochondrial transcripts were first identified using guanylyltransferase capping followed by S1 nuclease protection experiments, identifying the 3 main transcripts from LSP, HSP1, and HSP2. However, it is possible that lower-abundance transcripts exist, not captured by these methods. Due to the polycistronic nature of the mitochondrial transcripts, identification of true transcription initiation from sequencing methods can be difficult (Ojala, Montoya et al. 1981, Temperley, Wydro et al. 2010). Previous experiments have relied on pulse-chase labeling of nascent RNA transcripts (Cantatore and Attardi 1980), and could potentially be utilized for identification of the potential

lower-occurrence sites of transcriptional initiation. Such identification of novel origins of transcription by nascent transcript characterization, in conjunction with analysis of existing knowledge of the transcriptome and novel pause sites for POLRMT through ChIP-seq, could aid in identifying novel transcriptional start sites or stop sites coinciding with the non-D-loop binding sites of the identified nuclear transcription factors, forwarding understanding of their function.

Interestingly, several of the nuclear transcription factors we identified as binding to mtDNA have been previously identified as affecting mitochondrial function. It is well-known that MafF and MafK, part of the Maf family of proteins, effect function as a heterodimer with NRF-2 (Dhakshinamoorthy and Jaiswal 2000), a transcription factor which is a key activator of the oxidative stress response and modulates expression of cytochrome oxidase (Igarashi, Kataoka et al. 1994, Itoh, Chiba et al. 1997). Indeed, an NRF-2 consensus sequence has been identified in the promoters of TFAM, TBFIM, and TBFIIM (Rantanen, Jansson et al. 2001, Falkenberg, Gaspari et al. 2002, McCulloch, Seidel-Rogol et al. 2002). The proteins are somewhat redundant, leading to only a minor defect when a single gene is knocked out, with a moderate neurological defect in MafG knockout mice, perhaps due to its more ubiquitous expression pattern. Strikingly, a MafK/MafG double knockout displays severe motor ataxia and spastic hind legs (Kotkow and Orkin 1996, Shavit, Motohashi et al. 1998, Onodera, Shavit et al. 1999, Onodera, Shavit et al. 2000), while the triple knockout is embryonic lethal (Yamazaki, Katsuoka et al. 2012). Although the neurological defects have been currently attributed to reduction in

glycine receptor A1 expression, it is possible that the role of Maf proteins modulating mitochondrial function may also play a role, whether it be nuclear- or mitochondrial-genome driven. Interestingly, NFE2, also identified in our analysis, is a CNC family protein as well, forming a dimer with MafF to effect transcriptional regulation in erythrocytes (Blank, Kim et al. 1997). However, loss of NFE2 leads only to a mild erythropoietic effect, unsurprising given that its expression is limited to erythrocytes, and no mitochondrial-related phenotype is known (Shivdasani and Orkin 1995).

Two other nuclear transcription factors identified, c-Jun and JunD, are members of the JUN family of proteins and are components of the AP1 transcription factor (Mechta-Grigoriou, Gerald et al. 2001). Both of these transcription factors have been shown to dimerize with NRF-1 and NRF-2, driving induction of response to antioxidants and xenobiotics (Novotny, Prieschl et al. 1998, Venugopal and Jaiswal 1998). NRF-1 has been linked to expression of genes in all five subunits of the respiratory complexes (Virbasius, Virbasius et al. 1993, Tiranti, Rossi et al. 1995, Scarpulla 2002, Scarpulla 2002, Kelly and Scarpulla 2004), and also drives expression of TFAM, TFBIM, and TFB2M (Virbasius and Scarpulla 1994, Gleyzer, Vercauteren et al. 2005). Furthermore, AP-1, a Jun/Fos protein family complex, has been shown to translocate to mitochondria and bind to mtDNA (Ogita, Okuda et al. 2002, Ogita, Fujinami et al. 2003). c-Jun-lacking mice are embryonic lethal due to hematopoietic and hepatic apoptosis that does not appear to be related to hepatopathies seen in some patients with mitochondrial disease (Eferl, Sibilia et al. 1999). Interestingly, a JunD-/- mouse exhibits accelerated aging-

induced endothelial dysfunction, increased mitochondrial superoxide formation, increased mtDNA fragmentation, and swelling of the mitochondria (Paneni, Osto et al. 2013). Although these phenotypic effects could be due to nuclear effect alone, it is clear that c-Jun and JunD function in mitochondrial maintenance.

CEBPβ, a bZIP transcription factor, has been implicated in increasing mitochondrial biogenesis and expression of ETC components through elevated expression of the G protein α subunit in white adipose tissue of mice where CEBPα has been replaced by CEBPβ (Chiu, Lin et al. 2004). Interestingly, expression of Complex I, II, COX3, and NDUFS3 proteins was elevated in a CEBPβ knockout mouse (Rahman, Choudhury et al. 2013). While expression levels of the individual components of Complexes I and II, which contain mitochondrially encoded proteins, was not probed, COX3 expression levels were elevated in the knockout mouse. Because COX3 is mitochondrially encoded, this implies that CEBPβ knockout has an effect on transcription from the mitochondrial genome.

There is limited evidence for regulation of mitochondrial function by the other identified nuclear transcription factors. Max is required for almost all functions of Myc family proteins and is a member of the E2F6 complex, which is involved in gene repression via histone modification. A Max-/- knockout mouse exhibits embryonic lethality due to general developmental arrest (Shen-Li, O'Hagan et al. 2000). USF2, also a Myc family protein, has been linked to glucose-mediated induction of hepatic gene expression. Although USF2 has been implicated in transcription of some mitochondrial

genes (Breen and Jordan 1999), no mitochondrial defect has been identified in USF2 knockout mice, where the main defects are in altered kinetics of glucose response upon administration of a high carbohydrate diet and iron accumulation due to hepcidin deficiency (Vallet, Henrion et al. 1997, Nicolas, Bennoun et al. 2001). Finally, RFX5, a subunit of the RFX complex, has largely been implicated in regulation of the MHC class II molecules (Clausen, Waldburger et al. 1998, Gobin, Peijnenburg et al. 1998). Of all the nuclear transcription factors identified to have mitochondrial binding sites, none have been implicated in human diseases with phenotypes similar to that of mitochondrial diseases.

Several of the identified nuclear transcription factors have been identified as modulating mitochondrial gene expression from at least the nuclear genome. Interestingly, CEBPβ has been shown to regulate levels of COX3 transcripts, which originate from the mitochondrial genome. However, it is not clear whether these transcription factors act directly on mtDNA, and can only be decoupled from the nuclear effects through further biochemical characterization Although we were able to demonstrate mitochondrial localization of MafK to mitochondria using immunofluorescence, the other transcription factors we also identified in our analysis also require evidence of mitochondrial localization. Of note, we were able to localize c-Jun to mitochondria, but significant background fluorescence of the antibody used prevented inclusion of the results in our publication. Other methods of demonstrating localization to mitochondria have also previously included subcellular fractionation followed by

Western blot (Cammarota, Paratcha et al. 1999, Milanesi, Vasconsuelo et al. 2009), immunoelectron microscopy (Cammarota, Paratcha et al. 1999), and evidence for import of radiolabeled protein into the mitochondria (Casas, Rochard et al. 1999). And, of course – the mitochondrial effect of the transcription factors must be separated from the nuclear. This has previously been shown by observation of isolated mitochondria when exposed to the ligand which activates the TF in question (Enriquez, Fernandez-Silva et al. 1999), or by overexpression of mitochondrially-targeted protein (Lee, Kim et al. 2005, Casas, Pessemesse et al. 2008) in conjunction with monitoring of mtDNA and nuclear transcript levels. The addition of these two pieces of evidence – localization to the mitochondria and mtDNA-specific regulation agnostic of the nuclear genome – will confirm the identified transcription factors as acting on the mitochondrial genome directly.

The first utilization of ChIP-seq in the mitochondrial setting has clearly demonstrated that it is a powerful tool for visualizing binding of mitochondrial proteins, with great potential for use in understanding regulation within the mitochondrial nucleoid, especially given the unique size of the mitochondrial genome, which allows for coverage many fold above that of similar experiments in the nuclear genome. While the experiments detailed here demonstrate the power of the tool, its uses extend beyond simply understanding where transcription factors bind and into understanding the dynamics of binding. Future studies will surely further our understanding of not only how the many components of the nucleoid interact with the genome, but also how the dynamics of nucleoid as driven changes in cell state, whether they be metabolic or

pathological. Given the importance of mitochondrial function in cellular and organismal viability, understanding and optimizing these dynamics of the nucleoid could certainly have future therapeutic implications that remain to be explored.

BIBLIOGRAPHY

Ban-Ishihara, R., T. Ishihara, N. Sasaki, K. Mihara and N. Ishihara (2013). "Dynamics of nucleoid structure regulated by mitochondrial fission contributes to cristae reformation and release of cytochrome c." Proc Natl Acad Sci U S A **110**(29): 11863-11868.

Blank, V., M. J. Kim and N. C. Andrews (1997). "Human MafG is a functional partner for p45 NF-E2 in activating globin gene expression." Blood **89**(11): 3925-3935.

Bogenhagen, D. and D. A. Clayton (1977). "Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle." Cell **11**(4): 719-727.

Bogenhagen, D. F., D. Rousseau and S. Burke (2008). "The layered structure of human mitochondrial DNA nucleoids." J Biol Chem **283**(6): 3665-3675.

Breen, G. A. and E. M. Jordan (1999). "Transcriptional activation of the F(1)F(0) ATP synthase alpha-subunit initiator element by USF2 is mediated by p300." Biochim Biophys Acta **1428**(2-3): 169-176.

Cammarota, M., G. Paratcha, L. R. Bevilaqua, M. Levi de Stein, M. Lopez, A. Pellegrino de Iraldi, I. Izquierdo and J. H. Medina (1999). "Cyclic AMP-responsive element binding protein in brain mitochondria." J Neurochem **72**(6): 2272-2277.

Cantatore, P. and G. Attardi (1980). "Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA." Nucleic Acids Res **8**(12): 2605-2625.

Casas, F., L. Pessemesse, S. Grandemange, P. Seyer, N. Gueguen, O. Baris, L. Lepourry, G. Cabello and C. Wrutniak-Cabello (2008). "Overexpression of the mitochondrial T3 receptor p43 induces a shift in skeletal muscle fiber types." PLoS One **3**(6): e2501.

Casas, F., P. Rochard, A. Rodier, I. Cassar-Malek, S. Marchal-Victorion, R. J. Wiesner, G. Cabello and C. Wrutniak (1999). "A variant form of the nuclear triiodothyronine receptor c-ErbAalpha1 plays a direct role in regulation of mitochondrial RNA synthesis." Mol Cell Biol **19**(12): 7913-7924.

Chan, N. C., A. M. Salazar, A. H. Pham, M. J. Sweredoski, N. J. Kolawa, R. L. Graham, S. Hess and D. C. Chan (2011). "Broad activation of the ubiquitin-proteasome system by Parkin is critical for mitophagy." Hum Mol Genet **20**(9): 1726-1737.

Chen, H., M. Vermulst, Y. E. Wang, A. Chomyn, T. A. Prolla, J. M. McCaffery and D. C. Chan (2010). "Mitochondrial fusion is required for mtDNA stability in skeletal muscle and tolerance of mtDNA mutations." Cell **141**(2): 280-289.

Chiu, C. H., W. D. Lin, S. Y. Huang and Y. H. Lee (2004). "Effect of a C/EBP gene replacement on mitochondrial biogenesis in fat cells." Genes Dev **18**(16): 1970-1975.

Clausen, B. E., J. M. Waldburger, F. Schwenk, E. Barras, B. Mach, K. Rajewsky, I. Forster and W. Reith (1998). "Residual MHC class II expression on mature dendritic cells and activated B cells in RFX5-deficient mice." Immunity **8**(2): 143-155.

Corral-Debrinski, M., J. M. Shoffner, M. T. Lott and D. C. Wallace (1992). "Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease." Mutat Res **275**(3-6): 169-180.

Coskun, P., J. Wyrembak, S. E. Schriner, H. W. Chen, C. Marciniack, F. Laferla and D. C. Wallace (2012). "A mitochondrial etiology of Alzheimer and Parkinson disease." Biochim Biophys Acta **1820**(5): 553-564.

Coskun, P. E., M. F. Beal and D. C. Wallace (2004). "Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication." Proc Natl Acad Sci U S A **101**(29): 10726-10731.

Dhakshinamoorthy, S. and A. K. Jaiswal (2000). "Small maf (MafG and MafK) proteins negatively regulate antioxidant response element-mediated expression and antioxidant induction of the NAD(P)H:Quinone oxidoreductase1 gene." J Biol Chem **275**(51): 40134-40141.

Eferl, R., M. Sibilia, F. Hilberg, A. Fuchsbichler, I. Kufferath, B. Guertl, R. Zenz, E. F. Wagner and K. Zatloukal (1999). "Functions of c-Jun in liver and heart development." J Cell Biol **145**(5): 1049-1061.

Elachouri, G., S. Vidoni, C. Zanna, A. Pattyn, H. Boukhaddaoui, K. Gaget, P. Yu-Wai-Man, G. Gasparre, E. Sarzi, C. Delettre, A. Olichon, D. Loiseau, P. Reynier, P. F. Chinnery, A. Rotig, V. Carelli, C. P. Hamel, M. Rugolo and G. Lenaers (2011). "OPA1

links human mitochondrial genome maintenance to mtDNA replication and distribution." Genome Res **21**(1): 12-20.

Enriquez, J. A., P. Fernandez-Silva, N. Garrido-Perez, M. J. Lopez-Perez, A. Perez-Martos and J. Montoya (1999). "Direct regulation of mitochondrial RNA synthesis by thyroid hormone." Mol Cell Biol **19**(1): 657-670.

Falkenberg, M., M. Gaspari, A. Rantanen, A. Trifunovic, N. G. Larsson and C. M. Gustafsson (2002). "Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA." Nat Genet **31**(3): 289-294.

Falkenberg, M., N. G. Larsson and C. M. Gustafsson (2007). "DNA replication and transcription in mammalian mitochondria." Annu Rev Biochem **76**: 679-699.

Gauthier, B. R., A. Wiederkehr, M. Baquie, C. Dai, A. C. Powers, J. Kerr-Conte, F. Pattou, R. J. MacDonald, J. Ferrer and C. B. Wollheim (2009). "PDX1 deficiency causes mitochondrial dysfunction and defective insulin secretion through TFAM suppression." Cell Metab **10**(2): 110-118.

Gleyzer, N., K. Vercauteren and R. C. Scarpulla (2005). "Control of mitochondrial transcription specificity factors (TFB1M and TFB2M) by nuclear respiratory factors (NRF-1 and NRF-2) and PGC-1 family coactivators." Mol Cell Biol **25**(4): 1354-1366.

Gobin, S. J., A. Peijnenburg, M. van Eggermond, M. van Zutphen, R. van den Berg and P. J. van den Elsen (1998). "The RFX complex is crucial for the constitutive and CIITA-

mediated transactivation of MHC class I and beta2-microglobulin genes." <u>Immunity</u> **9**(4): 531-541.

He, J., C. C. Mao, A. Reyes, H. Sembongi, M. Di Re, C. Granycome, A. B. Clippingdale, I. M. Fearnley, M. Harbour, A. J. Robinson, S. Reichelt, J. N. Spelbrink, J. E. Walker and I. J. Holt (2007). "The AAA+ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization." <u>J Cell Biol</u> **176**(2): 141-146.

Holt, I. J. and A. Reyes (2012). "Human mitochondrial DNA replication." <u>Cold Spring Harb Perspect Biol</u> **4**(12).

Hyvarinen, A. K., J. L. Pohjoismaki, A. Reyes, S. Wanrooij, T. Yasukawa, P. J. Karhunen, J. N. Spelbrink, I. J. Holt and H. T. Jacobs (2007). "The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA." <u>Nucleic Acids Res</u> **35**(19): 6458-6474.

Ide, T., H. Tsutsui, S. Hayashidani, D. Kang, N. Suematsu, K. Nakamura, H. Utsumi, N. Hamasaki and A. Takeshita (2001). "Mitochondrial DNA damage and dysfunction associated with oxidative stress in failing hearts after myocardial infarction." <u>Circ Res</u> **88**(5): 529-535.

Igarashi, K., K. Kataoka, K. Itoh, N. Hayashi, M. Nishizawa and M. Yamamoto (1994). "Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins." <u>Nature</u> **367**(6463): 568-572.

Itoh, K., T. Chiba, S. Takahashi, T. Ishii, K. Igarashi, Y. Katoh, T. Oyake, N. Hayashi, K. Satoh, I. Hatayama, M. Yamamoto and Y. Nabeshima (1997). "An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements." Biochem Biophys Res Commun **236**(2): 313-322.

Jui, H. Y. and T. W. Wong (1991). "In vitro replication of heavy strand DNA in permeabilized human mitochondria." Nucleic Acids Res **19**(4): 905-911.

Karamanlidis, G., V. Bautista-Hernandez, F. Fynn-Thompson, P. Del Nido and R. Tian (2011). "Impaired mitochondrial biogenesis precedes heart failure in right ventricular hypertrophy in congenital heart disease." Circ Heart Fail **4**(6): 707-713.

Karamanlidis, G., L. Nascimben, G. S. Couper, P. S. Shekar, F. del Monte and R. Tian (2010). "Defective DNA replication impairs mitochondrial biogenesis in human failing hearts." Circ Res **106**(9): 1541-1548.

Kelly, D. P. and R. C. Scarpulla (2004). "Transcriptional regulatory circuits controlling mitochondrial biogenesis and function." Genes Dev **18**(4): 357-368.

King, M. P. and G. Attardi (1996). "Isolation of human cell lines lacking mitochondrial DNA." Methods Enzymol **264**: 304-313.

Korhonen, J. A., X. H. Pham, M. Pellegrini and M. Falkenberg (2004). "Reconstitution of a minimal mtDNA replisome in vitro." EMBO J **23**(12): 2423-2429.

Kotkow, K. J. and S. H. Orkin (1996). "Complexity of the erythroid transcription factor NF-E2 as revealed by gene targeting of the mouse p18 NF-E2 locus." Proc Natl Acad Sci U S A **93**(8): 3514-3518.

Kruse, B., N. Narasimhan and G. Attardi (1989). "Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination." Cell **58**(2): 391-397.

Kucej, M., B. Kucejova, R. Subramanian, X. J. Chen and R. A. Butow (2008). "Mitochondrial nucleoids undergo remodeling in response to metabolic cues." J Cell Sci **121**(Pt 11): 1861-1868.

Lee, D. Y. and D. A. Clayton (1998). "Initiation of mitochondrial DNA replication by transcription and R-loop processing." J Biol Chem **273**(46): 30614-30621.

Lee, J., C. H. Kim, D. K. Simon, L. R. Aminova, A. Y. Andreyev, Y. E. Kushnareva, A. N. Murphy, B. E. Lonze, K. S. Kim, D. D. Ginty, R. J. Ferrante, H. Ryu and R. R. Ratan (2005). "Mitochondrial cyclic AMP response element-binding protein (CREB) mediates mitochondrial gene expression and neuronal survival." J Biol Chem **280**(49): 40398-40401.

Leigh-Brown, S., J. A. Enriquez and D. T. Odom (2010). "Nuclear transcription factors in mammalian mitochondria." Genome Biol **11**(7): 215.

Liu, T., B. Lu, I. Lee, G. Ondrovicova, E. Kutejova and C. K. Suzuki (2004). "DNA and RNA binding by the mitochondrial lon protease is regulated by nucleotide and protein substrate." J Biol Chem **279**(14): 13902-13910.

Lu, B., S. Yadav, P. G. Shah, T. Liu, B. Tian, S. Pukszta, N. Villaluna, E. Kutejova, C. S. Newlon, J. H. Santos and C. K. Suzuki (2007). "Roles for the human ATP-dependent Lon protease in mitochondrial DNA maintenance." J Biol Chem **282**(24): 17363-17374.

Maassen, J. A., T. H. LM, E. Van Essen, R. J. Heine, G. Nijpels, R. S. Jahangir Tafrechi, A. K. Raap, G. M. Janssen and H. H. Lemkes (2004). "Mitochondrial diabetes: molecular mechanisms and clinical presentation." Diabetes **53 Suppl 1**: S103-109.

Magnusson, J., M. Orth, P. Lestienne and J. W. Taanman (2003). "Replication of mitochondrial DNA occurs throughout the mitochondria of cultured human cells." Exp Cell Res **289**(1): 133-142.

Marroquin, L. D., J. Hynes, J. A. Dykens, J. D. Jamieson and Y. Will (2007). "Circumventing the Crabtree effect: replacing media glucose with galactose increases susceptibility of HepG2 cells to mitochondrial toxicants." Toxicol Sci **97**(2): 539-547.

McCulloch, V., B. L. Seidel-Rogol and G. S. Shadel (2002). "A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine." Mol Cell Biol **22**(4): 1116-1125.

Mechta-Grigoriou, F., D. Gerald and M. Yaniv (2001). "The mammalian Jun proteins: redundancy and specificity." Oncogene **20**(19): 2378-2389.

Merkwirth, C., P. Martinelli, A. Korwitz, M. Morbin, H. S. Bronneke, S. D. Jordan, E. I. Rugarli and T. Langer (2012). "Loss of prohibitin membrane scaffolds impairs mitochondrial architecture and leads to tau hyperphosphorylation and neurodegeneration." PLoS Genet **8**(11): e1003021.

Milanesi, L., A. Vasconsuelo, A. R. de Boland and R. Boland (2009). "Expression and subcellular distribution of native estrogen receptor beta in murine C2C12 cells and skeletal muscle tissue." Steroids **74**(6): 489-497.

Nicolas, G., M. Bennoun, I. Devaux, C. Beaumont, B. Grandchamp, A. Kahn and S. Vaulont (2001). "Lack of hepcidin gene expression and severe tissue iron overload in upstream stimulatory factor 2 (USF2) knockout mice." Proc Natl Acad Sci U S A **98**(15): 8780-8785.

Novotny, V., E. E. Prieschl, R. Csonga, G. Fabjani and T. Baumruker (1998). "Nrf1 in a complex with fosB, c-jun, junD and ATF2 forms the AP1 component at the TNF alpha promoter in stimulated mast cells." Nucleic Acids Res **26**(23): 5480-5485.

Ogita, K., Y. Fujinami, M. Kitano and Y. Yoneda (2003). "Transcription factor activator protein-1 expressed by kainate treatment can bind to the non-coding region of mitochondrial genome in murine hippocampus." J Neurosci Res **73**(6): 794-802.

Ogita, K., H. Okuda, M. Kitano, Y. Fujinami, K. Ozaki and Y. Yoneda (2002). "Localization of activator protein-1 complex with DNA binding activity in mitochondria of murine brain after in vivo treatment with kainate." J Neurosci **22**(7): 2561-2570.

Ojala, D., J. Montoya and G. Attardi (1981). "tRNA punctuation model of RNA processing in human mitochondria." Nature **290**(5806): 470-474.

Onodera, K., J. A. Shavit, H. Motohashi, F. Katsuoka, J. E. Akasaka, J. D. Engel and M. Yamamoto (1999). "Characterization of the murine mafF gene." J Biol Chem **274**(30): 21162-21169.

Onodera, K., J. A. Shavit, H. Motohashi, M. Yamamoto and J. D. Engel (2000). "Perinatal synthetic lethality and hematopoietic defects in compound mafG::mafK mutant mice." EMBO J **19**(6): 1335-1345.

Paneni, F., E. Osto, S. Costantino, B. Mateescu, S. Briand, G. Coppolino, E. Perna, P. Mocharla, A. Akhmedov, R. Kubant, L. Rohrer, T. Malinski, G. G. Camici, C. M. Matter, F. Mechta-Grigoriou, M. Volpe, T. F. Luscher and F. Cosentino (2013). "Deletion of the activated protein-1 transcription factor JunD induces oxidative stress and accelerates age-related endothelial dysfunction." Circulation **127**(11): 1229-1240, e1221-1221.

Picca, A., F. Fracasso, V. Pesce, P. Cantatore, A. M. Joseph, C. Leeuwenburgh, M. N. Gadaleta and A. M. Lezza (2013). "Age- and calorie restriction-related changes in rat brain mitochondrial DNA and TFAM binding." Age (Dordr) **35**(5): 1607-1620.

Picca, A., V. Pesce, F. Fracasso, A. M. Joseph, C. Leeuwenburgh and A. M. Lezza (2013). "Aging and calorie restriction oppositely affect mitochondrial biogenesis through TFAM binding at both origins of mitochondrial DNA replication in rat liver." PLoS One **8**(9): e74644.

Picca, A., V. Pesce, F. Fracasso, A. M. Joseph, C. Leeuwenburgh and A. M. Lezza (2014). "A comparison among the tissue-specific effects of aging and calorie restriction on TFAM amount and TFAM-binding activity to mtDNA in rat." Biochim Biophys Acta.

Rahman, S. M., M. Choudhury, R. C. Janssen, K. C. Baquero, M. Miyazaki and J. E. Friedman (2013). "CCAAT/enhancer binding protein beta deletion increases mitochondrial function and protects mice from LXR-induced hepatic steatosis." Biochem Biophys Res Commun **430**(1): 336-339.

Rantanen, A., M. Jansson, A. Oldfors and N. G. Larsson (2001). "Downregulation of Tfam and mtDNA copy number during mammalian spermatogenesis." Mamm Genome **12**(10): 787-792.

Reitzer, L. J., B. M. Wice and D. Kennell (1979). "Evidence that glutamine, not sugar, is the major energy source for cultured HeLa cells." J Biol Chem **254**(8): 2669-2676.

Rice, A. C., P. M. Keeney, N. K. Algarzae, A. C. Ladd, R. R. Thomas and J. P. Bennett Jr (2014). "Mitochondrial DNA Copy Numbers in Pyramidal Neurons are Decreased and Mitochondrial Biogenesis Transcriptome Signaling is Disrupted in Alzheimer's Disease Hippocampi." J Alzheimers Dis.

Rizwani, W. and S. P. Chellappan (2009). "In vitro replication assay with mammalian cell extracts." Methods Mol Biol **523**: 203-216.

Robinson, B. H., R. Petrova-Benedict, J. R. Buncic and D. C. Wallace (1992). "Nonviability of cells with oxidative defects in galactose medium: a screening test for affected patient fibroblasts." Biochem Med Metab Biol **48**(2): 122-126.

Scarpulla, R. C. (2002). "Nuclear activators and coactivators in mammalian mitochondrial biogenesis." Biochim Biophys Acta **1576**(1-2): 1-14.

Scarpulla, R. C. (2002). "Transcriptional activators and coactivators in the nuclear control of mitochondrial function in mammalian cells." Gene **286**(1): 81-89.

Seidel-Rogol, B. L. and G. S. Shadel (2002). "Modulation of mitochondrial transcription in response to mtDNA depletion and repletion in HeLa cells." Nucleic Acids Res **30**(9): 1929-1934.

Shavit, J. A., H. Motohashi, K. Onodera, J. Akasaka, M. Yamamoto and J. D. Engel (1998). "Impaired megakaryopoiesis and behavioral defects in mafG-null mutant mice." Genes Dev **12**(14): 2164-2174.

Shen-Li, H., R. C. O'Hagan, H. Hou, Jr., J. W. Horner, 2nd, H. W. Lee and R. A. DePinho (2000). "Essential role for Max in early embryonic growth and development." Genes Dev **14**(1): 17-22.

Shivdasani, R. A. and S. H. Orkin (1995). "Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2." Proc Natl Acad Sci U S A **92**(19): 8690-8694.

Simmons, R. A., I. Suponitsky-Kroyter and M. A. Selak (2005). "Progressive accumulation of mitochondrial DNA mutations and decline in mitochondrial function lead to beta-cell failure." J Biol Chem **280**(31): 28785-28791.

Stumpf, J. D., R. P. Saneto and W. C. Copeland (2013). "Clinical and molecular features of POLG-related mitochondrial disease." Cold Spring Harb Perspect Biol **5**(4): a011395.

Sugiyama, S., K. Hattori, M. Hayakawa and T. Ozawa (1991). "Quantitative analysis of age-associated accumulation of mitochondrial DNA with deletion in human hearts." Biochem Biophys Res Commun **180**(2): 894-899.

Suomalainen, A. and P. Isohanni (2010). "Mitochondrial DNA depletion syndromes--many genes, common mechanisms." Neuromuscul Disord **20**(7): 429-437.

Svitin, A. and I. Chesnokov (2010). "Study of DNA replication in Drosophila using cell free in vitro system." Cell Cycle **9**(4): 815-819.

Temperley, R. J., M. Wydro, R. N. Lightowlers and Z. M. Chrzanowska-Lightowlers (2010). "Human mitochondrial mRNAs--like members of all families, similar but different." Biochim Biophys Acta **1797**(6-7): 1081-1085.

Tiranti, V., E. Rossi, M. Rocchi, S. DiDonato, O. Zuffardi and M. Zeviani (1995). "The gene (NFE2L1) for human NRF-1, an activator involved in nuclear-mitochondrial interactions, maps to 7q32." Genomics **27**(3): 555-557.

Trifunovic, A. and N. G. Larsson (2008). "Mitochondrial dysfunction as a cause of ageing." J Intern Med **263**(2): 167-178.

Vallet, V. S., A. A. Henrion, D. Bucchini, M. Casado, M. Raymondjean, A. Kahn and S. Vaulont (1997). "Glucose-dependent liver gene expression in upstream stimulatory factor 2 -/- mice." J Biol Chem **272**(35): 21944-21949.

Venugopal, R. and A. K. Jaiswal (1998). "Nrf2 and Nrf1 in association with Jun proteins regulate antioxidant response element-mediated expression and coordinated induction of genes encoding detoxifying enzymes." Oncogene **17**(24): 3145-3156.

Virbasius, C. A., J. V. Virbasius and R. C. Scarpulla (1993). "NRF-1, an activator involved in nuclear-mitochondrial interactions, utilizes a new DNA-binding domain conserved in a family of developmental regulators." Genes Dev **7**(12A): 2431-2445.

Virbasius, J. V. and R. C. Scarpulla (1994). "Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis." Proc Natl Acad Sci U S A **91**(4): 1309-1313.

Wai, T., D. Teoli and E. A. Shoubridge (2008). "The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes." Nat Genet **40**(12): 1484-1488.

Wang, Y. and D. F. Bogenhagen (2006). "Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane." J Biol Chem **281**(35): 25791-25802.

Wang, Y. E., G. K. Marinov, B. J. Wold and D. C. Chan (2013). "Genome-wide analysis reveals coating of the mitochondrial genome by TFAM." PLoS One **8**(8): e74513.

Wanrooij, S., J. M. Fuste, G. Farge, Y. Shi, C. M. Gustafsson and M. Falkenberg (2008). "Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro." Proc Natl Acad Sci U S A **105**(32): 11122-11127.

Wenz, T., C. Luca, A. Torraco and C. T. Moraes (2009). "mTERF2 regulates oxidative phosphorylation by modulating mtDNA transcription." Cell Metab **9**(6): 499-511.

Wong, T. W. and D. A. Clayton (1985). "In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis." Cell **42**(3): 951-958.

Yamazaki, H., F. Katsuoka, H. Motohashi, J. D. Engel and M. Yamamoto (2012). "Embryonic lethality and fetal liver apoptosis in mice lacking all three small Maf proteins." Mol Cell Biol **32**(4): 808-816.

Zeviani, M. and S. Di Donato (2004). "Mitochondrial disorders." Brain **127**(Pt 10): 2153-2172.

APPENDIX 1 – LIST OF ACCEPTED PUBLICATIONS

Chen, H., M. Vermulst, Y. E. Wang, A. Chomyn, T. A. Prolla, J. M. McCaffery and D. C. Chan (2010). "Mitochondrial fusion is required for mtDNA stability in skeletal muscle and tolerance of mtDNA mutations." <u>Cell</u> **141[2]**: 280-289.

Marinov, G. K., Y. E. Wang, D. Chan and B. J. Wold (2014). "Evidence for site-specific occupancy of the mitochondrial genome by nuclear transcription factors." <u>PLoS One</u> **9**(1): e84713.

Wang, Y. E., G. K. Marinov, B. J. Wold and D. C. Chan (2013). "Genome-wide analysis reveals coating of the mitochondrial genome by TFAM." <u>PLoS One</u> **8**(8): e74513.