Role of EMG1 in Bowen-Conradi syndrome and in ribosome biogenesis

by

D. Joy Armistead

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the Degree of

Doctor of Philosophy

Department of Biochemistry and Medical Genetics

University of Manitoba

Winnipeg, Manitoba, Canada

Abstract

Bowen-Conradi syndrome is a lethal autosomal recessive disorder affecting Hutterite infants, with a phenotype including severe growth and psychomotor retardation, and leading to death at an average age of thirteen months. Characteristic physical features comprise microcephaly, micrognathia, prominent nose, limb contractures, and rocker-bottom feet. Although biochemical investigations did not provide any candidate genes, linkage analysis and sequencing identified an A>G mutation in *EMG1* as the probable cause of the disease. This gene is implicated in ribosome biogenesis, and the mutation results in an aspartate to glycine substitution causing an unstable EMG1 protein, leading to severely reduced levels in patient fibroblasts and lymphoblasts. Although the aspartate to glycine substitution does not affect EMG1 protein sub-cellular localization, it causes EMG1 to aggregate or become degraded and may alter its ability to associate with protein partners. The reduction in available EMG1 protein causes a transient delay in processing of the ribosomal small subunit 18S rRNA, leading to cell cycle delay at G2/M and a subsequent reduction in cell proliferation rates in Bowen-Conradi syndrome patient lymphoblasts. A mouse model of Bowen-Conradi syndrome also displayed severe developmental delay, with prominent effects in the cranial central nervous system. Embryos died prematurely during development, probably due to decreased proliferation rates accompanied by apoptosis. These results shed light on the etiology of Bowen-Conradi syndrome, and open the door for development of treatments.

Acknowledgements

I've learned that the best gift someone can give is their time, and I'd like to thank everyone who has so generously shared their time with me. First and foremost, my advisor Barb, who has a superhuman ability to retain a sense of humour in the most trying circumstances, and somehow makes everyone feel like she has all the time in the world for them. I'm so lucky and so proud to have been Barb's student, and I think the best way to thank her is to try and live up to her example. Thanks to my advisory committee, Hao Ding, Mary Lynn Duckworth, Cheryl Rockman-Greenberg, and particularly Steve Pind, who was always in my corner and ready to lend an ear.

My labmates Jit, Lara, Ramya, Naderah, and especially Rick, who has been there the whole way along, have been a great source of energy for me, with their infectious enthusiasm and excitement for science. I'll carry that with me wherever I go.

To my wonderful friends- Jen, Frieder, Selena, Sarah, Kert, Jill, Mike, Cheryl, Pamela, Desiree, Ludivine, Shannon, Christa- thank you for all your encouragement, and for sticking with me while I've been selfish with my own time. And because everyone knows that the way to my heart is through my stomach, thanks for feeding me and breaking all manner of bread with me- may Perogy Palooza live forever. Thanks for giving me a hand in so many ways over the years, and keeping me laughing!

Big thanks to my brother, who has always set the high bar for me to aim for. Finally, thanks to my parents, who are lovely, creative, funny, smart, and infinitely generous people. I look up to them more than they can imagine.

Table of Contents

List of abbreviations

Ψ: pseudouridine

5′ -TOP: 5′ terminal oligopyrimidine

Å: Angstrom

ABC: avidin-biotin complex

acp: amino-carboxy-propyl

BCA: bicinchoninic acid

BCS: Bowen-Conradi syndrome

BHK: baby hamster kidney

BLAST: basic local alignment search tool

BMP: bone morphogenic protein

Bop1: block of proliferation 1

bp: base pairs

BPTE: Bis-Tris PIPES EDTA buffer

BrdU: bromodeoxyuridine

BSA: bovine serum albumen

Bst: belly spot and tail

CAPS: *N*-cyclohexyl-3-aminopropanesulfonic acid

cDNA: complementary DNA

CNS: central nervous system

COFS: cerebro-oculofacioskeletal syndrome

CPM: counts per minute

Da: Dalton

DAB: 3, 3'- diaminobenzidine

DBA: Diamond-Blackfan anemia

DEPC: diethylpyrocarbonate

DFC: dense fibrillar component

DMEM: Dulbecco's modified Eagle medium

DMSO: dimethylsulfoxide

DNA: deoxyribonucleic acid

dNTP: deoxyribonucleotide triphosphate

Dsk: dark skin

DTA: diphtheria toxin A

E: embryonic day

EDTA: Ethylenediaminetetraacetic acid

EMG1: essential for mitotic growth 1

ENU: N-ethyl-N-nitrosourea

ES cells: embryonic stem cells

FBL: fibrillarin

FC: fibrillar centre

FLAG tag: DYKDDDDK

g: relative centrifugal force

GAPDH: glyceraldehyde 3-phosphate dehydrogenase

GC: granular component

GFP: green fluorescent protein

HA tag: human influenza hemagglutinin, YPYDVPDYA

H & E: hematoxylin and eosin

HEK: human embryonic kidney

HeLa: human cervical cancer cell line named for Henrietta Lacks

HPLC: high-performance liquid chromatography

HRP: horseradish peroxidase

IRES: internal ribosomal entry site

LB: lysogeny broth

MALDI MS: Matrix-assisted laser desorption/ionization mass spectrometry

MBP: maltose binding protein

MDM2: MDM2 oncogene, E3 ubiquitin protein ligase

MG132: Cbz-Leu-Leu-Leucinal

MOPS: 3-(N-morpholino)propanesulfonic acid

mRNA: messenger RNA

mTOR: mammalian target of rapamycin

MTT: (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

MYC: v-myc avian myelocytomatosis viral oncogene homolog

NOP14: nucleolar protein 14

NOR: nucleolar organiser region

ONPG: o-nitrophenyl-β-D-galactopyranoside

PBS: phosphate buffered saline

PCR: polymerase chain reaction

PI1: Proteasome inhibitor I, Z-Ile-Glu(OtBu)-Ala-Leucinal

POLR1: polymerase (RNA) I polypeptide

rDNA: ribosomal RNA genes

RIPA: radioimmunoprecipitation assay buffer

RNA: ribonucleic acid

RP-HPLC: reverse-phase high-performance liquid chromatography

RPL: large ribosomal subunit protein

RPMI: Roswell Park Memorial Institute medium

RPS: small ribosomal subunit protein

rRNA: ribosomal RNA

SA-IRES-βgeo: splice acceptor-internal ribosomal entry site-β-galactosidase-neomycin phosphotransferase

SAM: S-adenosyl methionine

SBDS: Shwachman-Bodian-Diamond syndrome

SDS-PAGE: sodium dodecyl sufate polyacrylamide gel electrophoresis

Shh: sonic hedgehog

SMN: survival of motor neuron

snoRNA: small nucleolar RNA

snoRNP: small nucleolar ribonuclear particles

SNP: single nucleotide polymorphism

snRNP: small nuclear ribonuclear particles

SPOUT: SpoU and TrmD

SSC: saline-sodium citrate buffer

TAE: tris-acetate EDTA buffer

TBE: tris-borate EDTA buffer

TBST: Tris buffered saline with Tween

TCA: trichloroacetic acid

TCOF1: Treacher Collins-Franceschetti syndrome 1

TCS: Treacher Collins syndrome

TP53: tumor protein 53 (human P53 protein)

TRP53: transformation related protein 53 (mouse P53 protein)

tRNA: transfer RNA

TUNEL: terminal deoxynucleotidyl transferase dUTP nick end labelling

UBF: upstream binding factor

WT: wild type

X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

List of figures

List of tables

Chapter 1 : Introduction

1.1 Bowen-Conradi syndrome

1.1.1 Description and history of disease

Bowen-Conradi syndrome (BCS) was first described in two infant Hutterite brothers by Drs. Peter Bowen and Gerhard Conradi in 1976[1]. When they were born, the babies had difficulty feeding, were small for gestational age, and had similar physical characteristics including a prominent nose, microcephaly, micrognathia, rocker-bottom feet, flexion contractures of the fingers, and undescended testes. Karyotypes and blood tests, including amino acid, bilirubin, calcium, and glucose levels, were all normal. The first infant was unable to feed on his own and required gavage feeding. He died at 16 weeks of age, while the second infant died at age 6.5 months after failing to thrive and a bout of pneumonia. The parents of these two brothers were second cousins, and the authors noted six children within the parents' families who had died in infancy. Accordingly, they speculated that the anomalies in the two brothers could be the result of either an autosomal or an X-linked recessively inherited syndrome.

The syndrome was further described in 1979 and much later, in 2003[2,3] , in Hutterite families of Alberta, Saskatchewan, Manitoba, Montana, and North and South Dakota, confirming the major findings of Bowen and Conradi. Lowry *et al* estimated that the frequency of BCS in this population was 1 in 355 live births. Additional findings included contractures of the hips and knees, minor ear anomalies, hands clenched in a "trisomy 18" position, and dolicocephaly. Although most affected infants died within the first thirteen months of life, several children survived much longer, the oldest to nine

years of age, albeit with severe psychomotor retardation and growth impairment which was evident in all cases even with feeding via gastrostomy tube. Affected infants who survived beyond one year were generally non-responsive and unable to sit independently, although they seemed to recognize caregivers. The authors highlighted evidence that neurologic defects underlie the syndrome, including symptoms such as microcephaly, breech delivery, poor feeding and cry, and developmental delay[2]. The cause of death was described as simple asphyxia rather than complications from pneumonia.

Autopsies are rare, performed in only seven of the thirty-nine cases described by Lowry *et al*. Five of these patients showed a very small brain[3], while computerized axial tomograms in two infants revealed partial agenesis of the cerebellum consistent with a Dandy-Walker malformation in one, and hypoplasia of the structures in the posterior fossa in the other[2].

Although all of the affected infants had a normal karyotype, similarities exist between BCS and trisomy 18, which was often the original diagnosis. BCS has also been compared with the syndrome of camptodactyly, multiple ankyloses, facial abnormalities, and pulmonary hypoplasia, as well as cerebro-oculofacioskeletal (COFS) syndrome. All of these syndromes share growth retardation, microcephaly, micrognathia, limb contractures, and failure to thrive leading to early death[1,4-6] , making clinical diagnosis difficult. Several cases of BCS have been reported outside of the Hutterite population, in

Russian[7], German[8], Turkish[9], and Indian[10] infants. With the exception of the Indian infant which resembled BCS, these cases were refuted based on morphological differences[3].

Pedigrees of the affected Hutterite children showed that they were all related through four individuals born in the late 1700s, prior to the Hutterite emigration to North America[2]. It is therefore likely that the gene responsible for BCS is widely spread throughout the Hutterite population, and inherited in an autosomal recessive manner since both male and female infants are affected, while the parents do not show any symptoms.

1.1.2 Hutterite population and history

The overall history of the Hutterite population can today be considered a success story, despite several periods of persecution and population decline. The Hutterite faith was founded in Europe in 1528 during the Protestant Reformation on the belief in voluntary adult baptism, absolute pacifism, common ownership of goods and property, and communal, austere living separate from the world. Jacob Hutter, for whom the religious order is named, joined the sect five years after its foundation and became the leader, but was publicly executed for his beliefs in 1536 at Innsbruck. Because of their disbelief in militaristic or nationalistic causes, the Hutterites were forced to migrate multiple times, through Moravia, Slovakia, Hungary, and Romania, finally settling in what is now Ukraine in 1770 where they were granted immunity from military service by Catherine the Great. During this time, their population swelled to a maximum of approximately 20 000 people, but later dove to a low of only 19 people. The current population is likely

descended from a core group of 17 families numbering fewer than 100 people. In 1770, when the Hutterites migrated to Ukraine, the population numbered only 116. After a century in the same place, their immunity from military service was revoked in 1874, and approximately 800 Hutterites made the journey to North America, to what is now South Dakota. Roughly half of these became independent farmers, while 443 people established three colonies. Since then, groups formed from the descendants of the three colonies have come to be known as Dariusleut, Schmiedeleut, and Leherleut, all of which have maintained separate identities and slightly different customs since World War I. The Schmiedeleut, the group represented today in Manitoba, was founded by 215 settlers[11,12].

Although the move to North America afforded the Hutterites a degree of religious freedom and an abundance of arable land, during World War I their adherence to pacifism again necessitated a move. In 1918, 15 colonies moved to Alberta and Manitoba in Canada to avoid American military service. Since then, colonies have migrated back to the United States and further into Canada. The Hutterites still maintain their belief systems and communal, agrarian living practices. Marriages are monogamous, divorce is taboo, and a high cultural significance is placed on the family. The high fertility rate, the large family size[11], and the high standard of medical care in Hutterite colonies has led to an increase in population to over 40 000 today, residing in Alberta, Saskatchewan, Manitoba, Montana, North and South Dakota, and Minnesota. The average colony size is approximately 80 individuals, and when numbers reach roughly 120 people, more land is purchased and a new colony is formed.

Because the Hutterites derive from a small founding population and are genetically isolated, a number of autosomal recessive disorders are over-represented in the population, such as cystic fibrosis, or are unique, such as limb girdle muscular dystrophy type 2H[13,14]. At the same time, certain disorders occur at much lower rates or are unknown, such as multiple sclerosis and neural tube defects[14]. Extensive written genealogical records maintained by colony preachers have meant that the Hutterite population represents a unique opportunity for studying Mendelian disease. Indeed, this has enabled mapping of the genes responsible for several disorders including limb girdle muscular dystrophy type 2H[15].

1.1.3 Genome scan and linkage analysis

Increasingly detailed descriptions of the BCS phenotype[1-3] did not reveal any possible genetic or biochemical causes of the disease. In an effort to locate the gene responsible for BCS, Lamont *et al* [16] performed a genome-wide scan and linkage analysis in affected families. Due to the small number of founders and the closed, endogamous nature of the Hutterite population, it was hypothesized that the mutation responsible for BCS arose once in Hutterite history, and thus, that all BCS patients would be homozygous for the same mutation inherited from a common ancestor.

DNA was isolated from blood samples collected from BCS patients and their families in Manitoba, Alberta, and Saskatchewan, including affected children, their parents,

unaffected siblings, aunts, uncles, and grandparents. Control samples for estimating marker allele frequencies were obtained from Schmiedeleut Hutterites who were not a second-degree or closer relative of a BCS patient. A genome-wide scan was initially performed using 389 microsatellite markers, although additional microsatellite markers and informative single nucleotide polymorphisms were subsequently included in the analysis. After creating pedigrees for the most recent common ancestors, linkage analysis was performed and haplotypes were generated. An identically homozygous region on chromosome 12p13.3 was found in ten affected children, which spanned 1.9 megabases and included 59 known or predicted genes. This region did not include either *ERCC2* or *ERCC6,* genes associated with COFS syndrome[17-19], genetic evidence that COFS syndrome and BCS are distinct. As described in Chapter 3 of this thesis, this study provided the basis for the discovery of the gene responsible for BCS, which encoded a protein essential for ribosome biogenesis.

1.2 Ribosome biogenesis

1.2.1 Ribosome composition and function in eukaryotes

Ribosomes, the centres of protein synthesis in all living organisms, are huge, highly complex catalytic molecules requiring a functional relationship between nucleic acid and protein. A mature ribosome is composed of two subunits, the large or 60S, and the small or 40S subunit. The 60S subunit contains approximately 49 proteins and the 28S, 5.8S and 5S ribosomal RNAs, while the 40S subunit contains approximately 33 proteins and the 18S rRNA[20-22]. The proteins are generally found on the surface and are largely

structural while peptide bond formation is catalyzed by the rRNA[23-25], a remnant of the primitive "RNA world." The assembled mature or 80S ribosome, is a huge particle at approximately 4 MDa, visible by electron microscopy. Despite its complexity, an assembled ribosome must be very accurate when catalysing the formation of proteins; the error rate for incorporation of amino acids into a nascent protein is estimated to be one in 10 000[26-28].

Ribosome function has been studied most extensively in prokaryotes, where research of ribosome structure has been fuelled by the desire to design antibiotics specifically targeting the prokaryotic ribosome[29]. The much more complex eukaryotic ribosome is best known in yeast, where the ease of genetic manipulation and the development of cell-free systems have left the study of human ribosomes far behind. Indeed, the first example of a eukaryotic ribosomal X-ray crystal structure, published in 2010 at the relatively low resolution of 4.15Å, was from *Saccharomyces cerevisiae*[30], while a human ribosome crystal structure has yet to be elucidated. Though yeast and human ribosomes share many similarities, there may be differences in function and regulation of ribosomes in human cells. The recent realization that defects in ribosomal proteins or in ribosome assembly components can lead to human disease[31-38], coupled with the potential of the human ribosome as a target for chemotherapeutic drugs[39,40], have re-focussed interest in the human ribosome. Nevertheless, the understanding of the human ribosome currently relies heavily on interpolations from the more tractable yeast model, and it is generally assumed that the human ribosome functions in a similar manner.

1.2.2 Nucleolus formation and rDNA gene organization

The ribosomal RNA genes, which encode the 18S, 28S and 5.8S rRNA and are collectively known as the rDNA, are the centres of ribosome assembly. Reflecting their importance in cell growth and proliferation, rRNA genes are remarkably abundant, with an estimated 400 copies per genome in humans[41-44]. Copy number has been shown to vary between individuals, although the significance of this is not understood[43]. These copies are arranged as tandem repeats of varying number, each copy separated by an intergenic spacer. Repeats are found at five different loci in humans, at the secondary constrictions of acrocentric chromosomes[45,46]. Studies of the exact structure and composition of the rRNA gene loci have been hampered by their repetitive nature; rDNA was not included in the human genome project and is absent from reference sequences. It has only recently been discovered that the copies of rDNA are not identical, but may instead be composed of at least seven different variants[47]. Despite the large number of rDNA repeats, it has been shown that only a fraction undergoes transcription at a given time, even under conditions of rapid cell proliferation[48,49].

Each of the rDNA loci in the human genome has the potential to form a nucleolus during transcription, and these sites were therefore originally termed nucleolar organiser regions (NORs)[50]. The nucleolus forms in the nucleus during rDNA transcription and constitutes a membraneless, highly dense structure in which smaller compartments can be observed by electron microscopy. At the centre of the nucleolus is a compartment called the

fibrillar centre (FC), which is surrounded by the dense fibrillar component (DFC). These are both surrounded by the granular component (GC) (Figure 1.1). The transcription of rRNA genes is an ordered process, thought to take place at the interface between the FC and the DFC, with the nascent transcript moving through the DFC and the GC as it matures[51-53]. rDNA transcription can be visualised in Miller chromatin spreads, which reveals a remarkable "Christmas tree" structure, where the rDNA is the trunk and nascent rRNAs of increasing length form the branches, joined to the trunk by RNA polymerase I. Miller spreads can include a tree at each rDNA repeat, like beads on a necklace. How this ordered structure compacts into the FC, DFC, and GC remains unclear, although complex models exist[54].

1.2.3 Ribosome synthesis in yeast and human cells

Ribosome biosynthesis seems to be similar in yeast and human cells, yet the timing and even the order of some of the steps either differ or are simply unknown in human cells. The general process in mammalian cells is the same: the precursor 45S rRNA is transcribed by RNA polymerase I as a long polycistronic transcript which is then extensively processed through cleavage and modification events to yield the mature 18S, 28S and 5.8S rRNAs forms, ensuring equimolar amounts of these rRNA species. The order and timing of nucleolytic cleavages is well-defined in yeast and can follow several alternative pathways[22,55]. The process is less understood in human cells, and likely follows one of two pathways depending on the timing of the first two cleavage events[22,56]. A simplified diagram of 45S rRNA processing is found in Figure 1.2. The

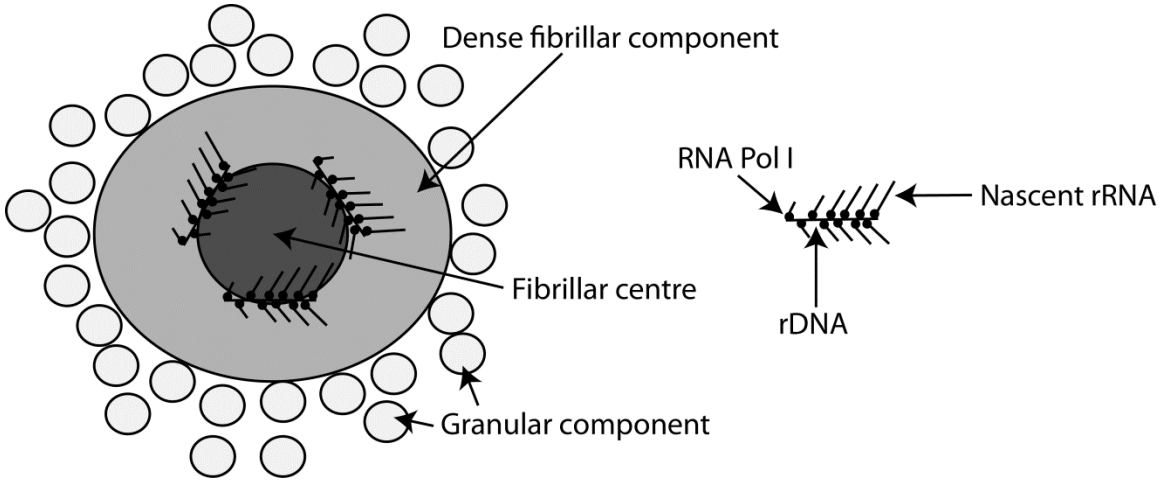


Figure 1.1 Structure of the nucleolus and rDNA transcription.

The nucleolus is centred on the fibrillar centre, the site of chromatin loops containing the rDNA repeats. The FC is surrounded by the dense fibrillar component, where the nascent rRNA is modified and processed. This is surrounded by the granular component, where the rRNA is assembled with ribosomal proteins and other components. The nascent rRNA at transcription sites forms “Christmas tree” structures, with the rDNA comprising the trunk of the tree, and rRNA transcripts connected to the trunk by RNA Pol I forming the branches.

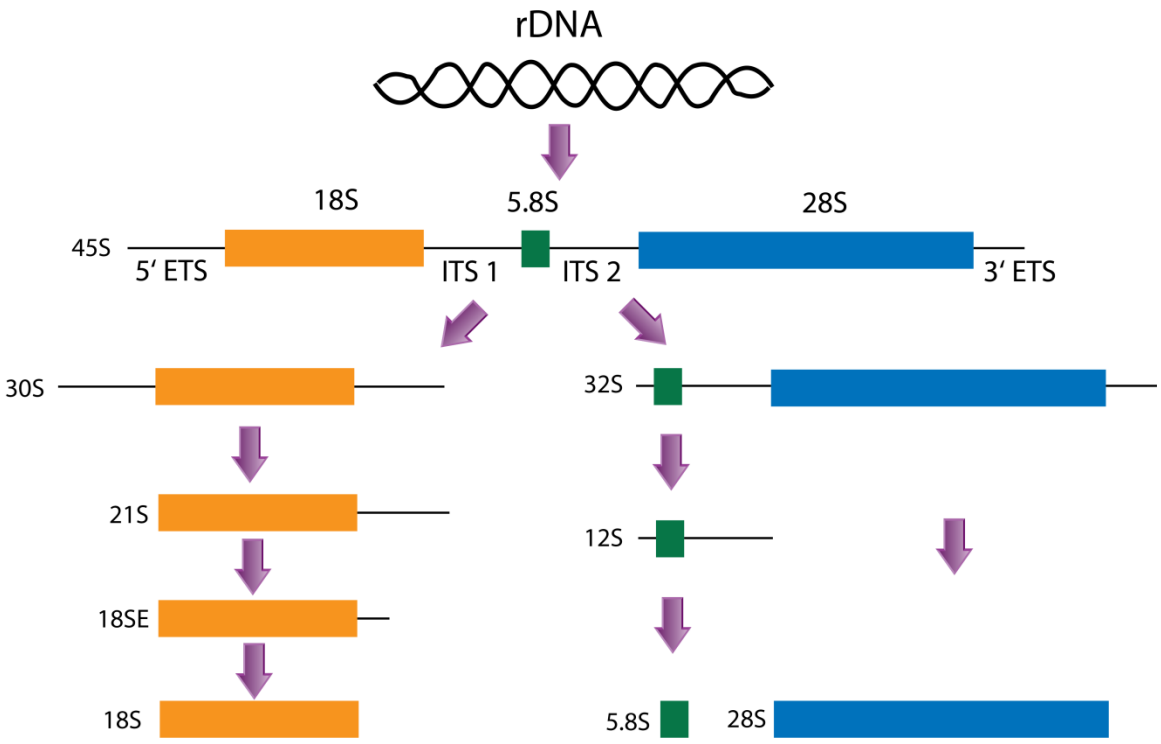


Figure 1.2 Ribosomal RNA processing.

The ribosomal RNA is transcribed as a long precursor 45S rRNA at sites of ribosomal RNA gene repeats, collectively known as the rDNA. The 45S rRNA contains the 18S, 5.8S and 28S rRNAs, flanked by two external transcribed spacers (ETS) and separated by two internal transcribed spacers (ITS). There are numerous pathways of maturation; this diagram depicts a simplified version. Cleavage in ITS 1 separates the 30S and 32S precursors, and the two pathways mature largely independently of one another thereafter. Cleavages by endo- and exo-nucleases result in the mature 18S, 5.8S, and 28S rRNA species.

initial endonucleolytic cleavages take place in the transcribed spacer regions, and the transcripts are then processed further in both directions by exonucleases.

The 5S rRNA is transcribed independently by RNA polymerase III in the nucleoplasm. It undergoes its own maturation pathway and is re-imported into the nucleolus as a complex with large ribosomal protein RPL5, to be incorporated into the assembling ribosome[57]. The ribosomal protein genes are transcribed by RNA polymerase II, and must be transcribed in the nucleus, their mRNA exported to the cytoplasm where they are translated, and the proteins re-imported into the nucleus to be assembled with the nascent rRNA. How the cell coordinates ribosomal protein transcription and translation with rRNA synthesis is not well understood.

All of these elements come together in the nucleolus to form the pre-90S ribosome, which is further processed to yield the pre-60S and pre-40S ribosomes. These intermediates are exported into the cytoplasm via nuclear pore complexes, where they undergo final maturation or "unpacking" steps to become the mature 60S and 40S subunits, which join to form the 80S ribosome (Figure 1.3).

1.2.4 rRNA modification

Ribosomal RNA maturation is a complex yet well-described process which requires the participation of hundreds of non-ribosomal factors. These factors, which include proteins and small nucleolar RNAs (snoRNAs), cleave, unwind, fold, and covalently modify the

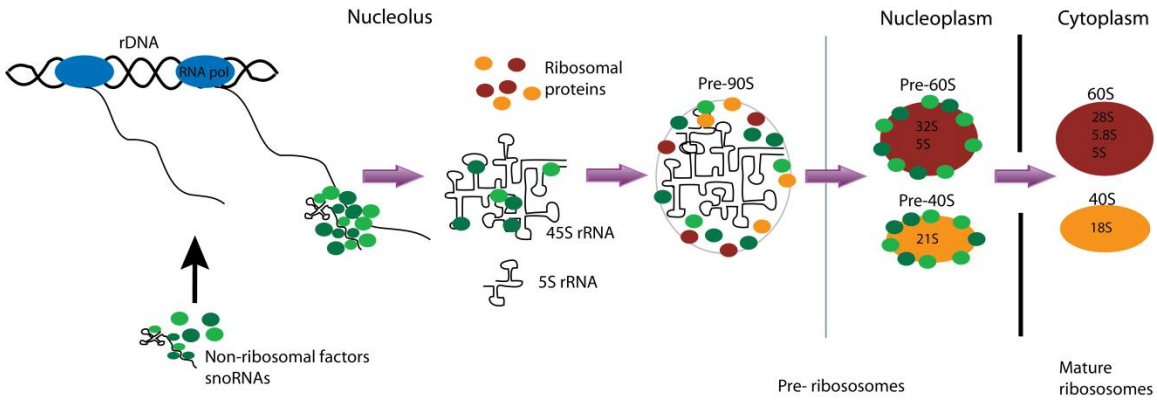


Figure 1.3 Ribosome biosynthesis.

Ribosome biosynthesis takes place in the nucleolus, starting with the transcription of RNA genes by RNA Polymerase I. The nascent transcript is joined by non-ribosomal factors including proteins and snoRNAs which modify the rRNA co-transcriptionally. The 45S rRNA precursor is joined by ribosomal proteins translated in the cytoplasm and re-imported into the nucleus, as well as by the 5S rRNA which is transcribed in the nucleoplasm. The pre-90S ribosome undergoes many modification and folding events, and is cleaved into the pre-60S ribosome containing the 32S and 5S rRNAs, and the pre-40S ribosome containing the 21S rRNA. These are exported into the cytoplasm before undergoing final modifications and cleavages which result in the mature 60S and 40S ribosome subunits, no longer associated with non-ribosomal factors.

nascent rRNA co-transcriptionally[58-60], and may also aid in its assembly with ribosomal proteins. Modifications to the rRNA are akin to post-translational modifications in proteins and include pseudouridylation, 2′-OH ribose methylation, base alterations, or combinations of these termed hypermodifications[61]. The most common sequence-specific RNA modifications are isomerisation of uridine to pseudouridine (at approximately 95 sites) and 2′-OH ribose methylation (at approximately 115 sites)[61-63]. Pseudoruridylation is catalysed by box H/ACA small nucleolar ribonuclear particles (snoRNPs), while 2′-OH ribose methylation is catalysed by box C/D snoRNPs. These snoRNPs gain their sequence specificity from a guide snoRNA which is complementary to the rRNA sequence to be modified, and is complexed with a catalytic protein. By employing a universal enzyme but specific guide snoRNAs, the cell avoids the necessity of having hundreds of different enzymes for site-specific pseudouridylation and methylation. On the other hand, there are a minority of rRNA modifications that require unique site-specific enzymes, including base modifications. The exact function of the rRNA modifications is somewhat unclear, but it has been shown that, in the context of the assembled ribosome tertiary structure, they cluster at important sites such as the A and P sites and the polypeptide exit tunnel[61,64]. It is therefore thought that rRNA modifications aid in its proper folding and function, by altering and optimizing interactions with neighbouring rRNA and ribosomal proteins, or even with tRNA and mRNA[61,65,66].

The hypermodified nucleotide N1-methyl-N3-(3-amino-3-carboxypropyl) pseudouridine (m1acp3ψ) is unique in the mature rRNA. While loss of modification at most single sites in the rRNA does not appear to have a significant impact on ribosome function, the

hypermodified m1acp3ψ is an exception[64,67,68]. Impairment of the pseudouridylation reaction results in a delay of 18S rRNA processing and a decrease in ribosome fidelity[64]. The formation of this base is initiated in the nucleolus by isomerization of uridine to pseudouridine by an H/ACA-class snoRNP containing the snoRNA snR35. A second reaction methylates N1 of the pseudouridine, and the modification is completed in the cytoplasm by the addition of the amino-carboxypropyl group (Figure 1.4)[69-71]. In the absence of the pseudouridylation reaction, the amino-carboxypropyl modification can still take place, but methylation is abrogated[64,72].

1.2.5 Regulation and importance in cell proliferation and survival

The centrality of ribosomes to cell proliferation and survival is underscored by the fact that the majority of ribosome biogenesis genes are essential and non-redundant in yeast[59,60,73]. It should come as no surprise that ribosome biogenesis is highly regulated at several different levels. Initiation of rDNA transcription, elongation rate, and rate of 45S rRNA transcript processing must be coordinated with ribosomal protein transcription, translation, post-translational modification, and transport back into the nucleus. Synthesis of the approximately 80 ribosomal proteins somehow occurs at equimolar quantities, and in fact, a reduction in the levels of a single ribosomal protein leads to the concomitant reduction of all the other proteins in the same subunit[74]. These processes must also be modulated according to environmental stimuli.

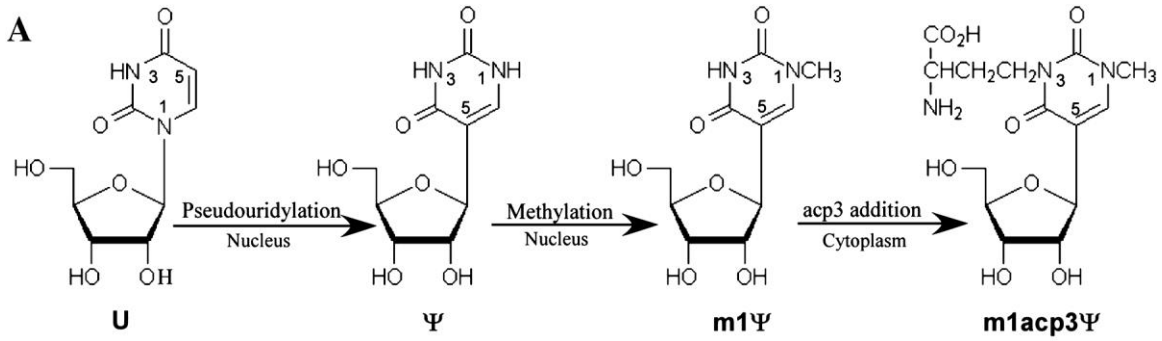


Figure 1.4 Modification of uridine.

The uridine at position 1248 in human 18S rRNA is extensively modified. First, the base undergoes isomerisation by pseudouridine synthase. The N1 atom is then available for the addition of a methyl group by a methyltransferase (EMG1). Finally, after the maturing small ribosomal subunit is transported into the cytoplasm, an amino-carboxypropyl (acp) moiety is added to the N3 atom. The final hypermodified pseudouridine, known as m1acp3Ψ, may play an important role in the P site of the mature ribosome. Reprinted with permission from Liang X, Liu Q, and Fournier MJ. Loss of rRNA modification in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. RNA (2009), 15: 1716-1728.

In a proliferating cell, an increased requirement for protein synthesis means an increased requirement for ribosomes. Ribosome biogenesis and cellular environment are linked by the mammalian target of rapamycin (mTOR) pathway, which senses nutrient availability, hormones, energy levels and cellular stress. In situations where increased ribosome biogenesis is required, mTOR complex 1 promotes rDNA transcription by positively regulating RNA polymerase I. In addition, it phosphorylates ribosomal protein S6 kinase 1 (S6K1), which in turn phosphorylates small ribosomal subunit protein RPS6. The phosphorylated RPS6 appears to alter the specificity of the ribosome, conferring it with a preference for translating 5 ′ terminal oligopyrimidine (5′-TOP) mRNAs, a group of mRNAs which includes ribosomal protein genes[75,76].

Another master coordinator of cell growth and division, the transcription factor MYC, has been shown to directly bind and stimulate the transcription of the rRNA genes[77]. MYC recruits transcription factor SL1 and influences transcription by RNA polymerase I. It appears that ribosome biogenesis and MYC are interdependent, as cells haploinsufficient for ribosomal protein genes are able to suppress the oncogenic activity of MYC[78].

The tumor protein 53 (TP53) tumor suppressor pathway stands as a strong checkpoint to verify that ribosome biogenesis is intact before allowing cell proliferation. In a healthy cell, the E3 ubiquitin ligase MDM2 binds and ubiquitinates TP53, thereby targeting it for degradation by the proteasome and maintaining it at low levels. In a cell where ribosome

biogenesis has been perturbed, the balance between rRNA and ribosomal protein synthesis becomes uncoupled, leading to free ribosomal proteins that cannot be incorporated into a functional ribosome. Certain free ribosomal proteins such as RPL5, RPL11, or RPL23 bind and segregate MDM2, which leads to the stabilization of TP53. Cell cycle arrest and/or apoptosis ensue, ensuring that cells with deficiencies in ribosome biogenesis do not survive[79-83].

Cell cycle and ribosome biogenesis are directly linked, as it has been shown in yeast not only that decreases in ribosome levels delay progression through the Start checkpoint[84], but also that the small subunit processome proteins are necessary for exit from G1 and entry into G2[59]. In mammalian cells, deletion of *Bop1*, a gene involved in large ribosomal subunit biosynthesis, results in cell cycle arrest, also at the G1 checkpoint[79]. Significantly, perturbation of the TP53 pathway reverses this effect, allowing progression through the cell cycle.

1.2.6 Ribosome biogenesis in brain development

Bowen-Conradi syndrome is predominantly a neurological disorder, as affected infants are severely neurologically impaired, even lacking the ability to suck. The defect causing BCS clearly has a profound deleterious effect on brain development. Given the essential role of ribosome biogenesis in cell proliferation, it seems logical that a defect in a ribosome assembly protein could alter the process of rapid cell growth and division crucial for the early stages of central nervous system development during embryogenesis.

Examples from the literature are, however, rare. In zebrafish, mutations in ribosomal biogenesis proteins *bap28* and *pescadillo* result in p53-mediated apoptosis in the central nervous system[85], and a disruption of oligodendrocyte formation due to cell cycle delay, respectively[86]. The *bap28* mutation is accompanied by a delay in rRNA synthesis, however this was not examined in the case of the *pescadillo* mutation. In a mouse model of Diamond-Blackfan anemia, mutations in ribosomal protein gene *Rps7* lead to enlarged ventricles in the brain and cortical thinning caused by increased apoptosis[87]. Modifications to the rRNA can also affect brain development, as removal of sequence-specific snoRNAs associated with rRNA modification can lead to brain malformation in zebrafish[88].

1.3 Ribosomopathies

1.3.1 Description of diseases

As the molecular causes of previously described diseases become known, diseases of ribosome biogenesis have recently begun to be grouped. Diseases arising from defects in ribosomal protein genes or in genes associated with ribosomal biogenesis are called "ribosomopathies". These diseases are highly heterogeneous in both their physical manifestations and modes of inheritance, yet share several characteristics, including microcephaly, hematological defects, and predisposition to cancer, most commonly leukemia[31,89,90]. It has also been proposed that immune defects are a hallmark of ribosomopathies[91]. Table 1.1 summarizes the features of the currently known ribosomopathies. Research into the etiology of ribosome biogenesis disorders has aided

the examination of the causes of BCS, as these diseases tend to share certain defects at a molecular level. Processing delays or defects in rRNA maturation are often found, resulting in an imbalance of mature ribosome subunit levels, which frequently causes a reduction in rates of protein synthesis and cell proliferation.

Disease (OMIM)	**Inheritance**	**Clinical manifestations**	**Gene**	**Function in ribosome biogenesis**
Diamond Blackfan anemia (#105650)	autosomal dominant, sporadic	anemia, bone marrow failure, craniofacial abnormalities, cardiac defects, cancer predisposition	*RPS7,RPS10, RPS17, RPS19, RPS24,RPS26 RPL5, RPL11, RPL35A,*	Ribosomal protein
Shwachman-Diamond syndrome (#260400)	autosomal recessive	growth retardation, exocrine pancreas insufficiency, hematologic defects, skeletal abnormalities, cancer predisposition	*SBDS*	Binding of 40S and 60S subunits
Dyskeratosis congenita (X-linked: #305000)	X-linked recessive, autosomal dominant, autosomal recessive	mucocutaneous abnormalities, bone marrow failure, cancer predisposition, growth retardation, immunodeficiency	*DKC1, TERT, TERC, NOP10*	Pseudouridine synthase
Treacher Collins syndrome (AD: #154500)	autosomal dominant, autosomal recessive	craniofacial abnormalities	*TCOF1, POLR1D, POLR1C*	Transcription of rRNA genes
Cartilage hair hypoplasia	autosomal recessive	short stature, hypoplastic hair, anemia,	*RMRP*	Cleavage of rRNA

(#250250)		cancer predisposition		
Anauxetic dysplasia (#607095)	autosomal recessive	extreme short stature	*RMRP*	Cleavage of rRNA
5q- syndrome (#153550)	sporadic somatic deletion	severe macrocytic anemia, cancer predisposition	*RPS14*	Ribosomal protein
Alopecia, neurological defects, and endocrinopathy syndrome (#612079)	autosomal recessive	hair loss, short stature, microcephaly, mental retardation, progressive motor retardation, delayed puberty, central adrenal insufficiency	*RBM28*	Nucleolar component of snRNP complex
North American Indian childhood cirrhosis (#604901)	autosomal recessive	Transient neonatal jaundice progressing to cirrhosis	*CIRH1A*	Pre-rRNA processing
Isolated congenital asplenia (#271400)	autosomal dominant	Asplenia leading to severe bacterial infections	*RPSA*	Ribosomal protein

Table 1.1 The ribosomopathies

One of the most intriguing, yet confounding, discoveries of ribosomopathy research has been that they display a degree of tissue specificity. Ribosome biogenesis is a universal process (with the rRNA genes commonly referred to as "housekeeping" genes), yet ribosome biogenesis defects often manifest themselves only in particular tissues. For example, Treacher Collins syndrome is caused by mutations in *TCOF1*, a gene associated

with rDNA transcription, or in *POLR1C* and *POLR1D*, subunits of RNA polymerase I, which transcribes the rDNA[92,93]. Facial defects seen in Treacher Collins syndrome are a result of a reduced proliferating neural crest cell population- other cells are unaffected during development[94]. Similarly, Diamond Blackfan anemia is caused by mutations in ribosomal protein genes, yet the disease is most commonly characterized by hematological defects. Even more confusingly, mutations in mouse ribosomal protein genes *Rps19* and *Rps20* cause a reduction in melanocyte numbers during development, but results in their accumulation in the adult[95]. Bowen Conradi syndrome seems to strongly affect the central nervous system during development, as evidenced by the severe developmental delay seen in affected children, while avoiding clear hematological defects often seen in other ribosomopathies. There are obvious spatial and temporal aspects to ribosome biogenesis disorders, areas of avid research interest at this time.

Recently, it has become increasingly clear that the ribosome itself plays an active regulatory role, rather than being a simple protein synthesis machine. Several studies have shown that under certain conditions the ribosome displays a preference for translating specific types of mRNAs. While this concept has been around for at least a decade as the "ribosome filter hypothesis"[96], it has only recently begun to gain compelling evidence. One example previously mentioned is that 5′-TOP mRNAs are preferentially translated via mTOR complex 1-mediated signalling when nutrient levels are high, increasing ribosome protein production[75,76]. Unexpectedly, a study by Kondrashov *et al* in 2011[97] found by screening patterning defects in mice that a mutation in the large subunit protein RPL38 leads to a reduction in the translation of a subset of

homeobox mRNAs, although global protein synthesis rates were unchanged. RPL38 expression varies from tissue to tissue, and the authors speculated that alterations of ribosomal protein expression might result in specialised ribosomes, which could in turn exert transcript-specific translational control. In the same vein, mutations causing X-linked dyskeratosis congenita result in a reduction of internal ribosome entry site (IRES)-mediated mRNA translation[98]. In other ribosomopathies where no evidence of altered global protein synthesis rates has been found, it remains to be seen whether the specificity of the ribosome is affected, resulting in preferential translation of a subset of mRNA transcripts.

1.3.2 Animal models

The first indications that defects in ribosomes could lead to morphological abnormalities in a multicellular organism were found in *Drosophila.* Mutations at 50-60 loci known as the *Minute* genes caused a similar phenotype, consisting of short bristles, slow development, reduced viability, rough eyes, small body, and etched tergites. The common characteristics for all of the mutations suggested that a single pathway was involved. In 1985, Kongsuwan *et al*[99] showed for the first time that one of the *Minute* loci encoded a ribosomal protein, and speculated that the rate of ribosome assembly was lowered, affecting cell proliferation rates. Subsequently, most of the approximately 50 *Minute* genes were identified as ribosomal protein genes[100].

In the mouse, the majority of research into ribosome defects has come from recent knockout and targeted knockin models designed specifically to mimic human ribosomopathies. Knockouts of the genes responsible for Diamond-Blackfan anemia[101], Shwachman-Diamond syndrome[102], dyskeratosis congenita[103], and Treacher Collins syndrome[104] all result in embryonic lethality. Mouse models of ribosomopathies have nonetheless uncovered essential information about the importance of ribosome biogenesis in development. For example, mouse models of the autosomal dominant form of Treacher Collins syndrome have shown that a defect in the *Tcof1* gene, encoding the rDNA transcription-associated protein Treacle, causes decreased production of ribosomes in neural crest cells. Subsequent reductions in neural crest cell number and proliferation rate lead to hypoplasia of the cranial and facial bones[94]. Significantly, loss of one allele of the transformation related protein 53 gene (*Trp53*) in these mice completely rescues the phenotype by decreasing apoptosis in neuroepithelial and neural crest cells[105].

Several more serendipitous discoveries have also contributed to the field. A homozygous lethal mutation causing pigmentation and retinal defects known as the Belly spot and tail mouse (*BstI)*, was termed a mouse *Minute* when it was determined to be caused by a mutation in a gene encoding a large ribosomal subunit protein[106]. More recently, an N-ethyl-N nitrosurea mutagenesis screen for pigmentation defects led to the discovery that two different mutants known as Dark skin 3 and 4 (*Dsk3* and *Dsk4*) were caused by mutations in ribosomal protein genes *Rps19* and *Rps20*[95]. Ablation of the TRP53 checkpoint rescued the pigmentation defect, as well as hematopoietic and growth

abnormalities. Since *RPS19* is mutated in approximately 25% of cases of Diamond-Blackfan anemia[107], these results could influence treatment of human disease.

Zebrafish models of ribosome biogenesis disorders, usually employing morpholino knockdown of genes of interest, have proven extremely useful in untangling the developmental basis of tissue-specific phenotypes of ribosomopathies. In a Shwachman-Diamond syndrome study it was shown that *sbds* expression in the developing zebrafish pancreas persists over time, while it diminishes in other tissues including the adjacent gut and liver[108]. Knockdown of *sbds* resulted in reduced proliferation specifically in exocrine pancreatic cells[108,109], which could explain the exocrine pancreatic deficiency in Shwachman-Diamond syndrome patients. Similarly, a zebrafish study of erythropoietic defects in Diamond-Blackfan anemia found that ablation of TP53 rescued morphological abnormalities, but did not improve erythropoiesis[110,111] suggesting tissue-dependent roles for TP53.

1.3.3 Comparison with BCS

Of the characteristic defects seen in multiple ribosomopathies, microcephaly is the only one that has been observed in BCS patients. Neither hematopoietic defects nor cancer predisposition have been reported; however this may be a reflection of the early lethality of the disease. It is intriguing that immune system deficiencies have been seen in a number of ribosomopathies[91], as BCS patients in the literature are often reported to have reoccurring cases of pneumonia[1-3].

1.4 EMG1

1.4.1 EMG1 function in yeast

EMG1 was first described in 1996 by Hakuno *et al*[112], who isolated *Mra1*, the *Schizosaccharomyces pombe* orthologue of *EMG1*, as a novel high-copy-number suppressor of a *Ras1* deficiency. This paper showed for the first time that *Mra1*was essential for cell growth, as *S. pombe* spores lacking *Mra1* can germinate and divide 1-3 cell cycles, but subsequently fail to grow. The function of *Mra1* was however not elucidated.

In 2001, Liu *et al*[113] isolated *EMG1* in *Saccharomyces cerevisiae* in a screen for novel genes repressed during heat shock, naming it Essential for Mitotic Growth 1. An effort to identify suppressors of the lethal Δ*emg1* defect only retrieved *EMG1*, suggesting that its function is both essential and non-redundant. The gene is highly conserved, and mouse, human and *Candida albicans EMG1* are able to compensate for the deletion of the *S. cerevisiae* gene[113,114]. An examination of Emg1 sub-cellular localization by immunoblot showed that it localizes both to the nuclear and the cytoplasmic fractions[113], while GFP-tagged EMG1 revealed a rod-shaped localization in yeast, indicating a possible spindle and microtubule association[114].

The pattern of *EMG1* expression following heat shock is similar to that of many ribosome protein genes in yeast, and it was found to be essential for the processing of precursor 35S rRNA to mature 18S rRNA. A reduction of *EMG1* levels leads to a corresponding reduction in levels of the mature small ribosome subunit[113], and deletion of Emg1 specifically abolishes 18S rRNA maturation without affecting 28S rRNA[59]. Additionally, expression of S-adenosyl methionine (SAM) synthase or medium supplementation with SAM rescues a temperature-sensitive EMG1 defect[114]. Taken together, these studies of EMG1 provide a picture of an essential, highly conserved protein required for processing of the 18S rRNA, possibly via mediation of a methylation reaction.

EMG1 is not associated with the mature ribosome, but is a component of the small subunit processome[59]. The small subunit processome is a large complex rivalling the small ribosomal subunit itself in size, composed of the U3 snoRNA and multiple ribosomal and non-ribosomal proteins, which aids in the processing of the full-length precursor rRNA into the mature 18S rRNA[60]. A yeast two-hybrid screen identified only one binding partner, Nucleolar Protein 14 (Nop14)[113], another small subunit processome component. Co-immunoprecipitations showed that, while the interaction between the two proteins is specific, only a fraction of Emg1 and Nop14 seem to bind each other. Furthermore, depletion of Nop14 eliminates the nuclear fraction of Emg1 without affecting total levels, suggesting that Nop14 is responsible for the sub-cellular localization of Emg1.

Genetic interactions between *EMG1* and *snR57*, *RPS19*, and fungi-specific *Tma23* and *Nop6* have also been shown[115,116]. These studies indicate that a growth defect in *Emg1* mutants can be compensated by a loss of snR57, the snoRNA responsible for sequence-specific 2′-OH methylation of G1570 on the 18S rRNA, or by the overexpression of small-subunit protein RPS19. This led to a proposed model whereby Emg1 binds 18S rRNA on helix 47, allowing the release of snR57 and promoting association of RPS19 to the nascent ribosome. Significantly, RPS19 is mutated in 25% of Diamond-Blackfan anemia cases[107], suggesting a possible link between BCS and Diamond-Blackfan anemia.

1.4.2 EMG1 crystal structure

Although analysis of EMG1 protein sequence showed strong conservation in eukaryotes and archaea[113,114], none of the homologues had defined functions. The resolution of the EMG1 X-ray crystal structure, however, allowed its inclusion in a well-studied superfamily of proteins. The *S. cerevisiae* Emg1 monomer crystal structure at 2Å[117] revealed an α/β sandwich fold with a central 5-stranded parallel β-sheet, and a deep knot in the amino acid backbone at the C-terminus which binds SAM. The presence of this characteristic fold assigns EMG1 to the superfamily of the α/β knot fold methyltransferases, also known as SpoU and TrmD (SPOUT) methyltransferases. Members of this superfamily do not necessarily share primary sequence similarities, but their secondary structures contain distinct folds. All α/β knot methyltransferases characterized to date depend on SAM as a methyl donor, function as dimers in solution, and are involved in posttranscriptional methylation of either tRNA or rRNA[118]. The crystal structure also revealed Emg1-specific insertions including a surface loop and a β-

α-β element, meaning that it represents the first member of a new subfamily of α/β knot fold methyltransferases.

The crystal structure at 2.2Å of the archaebacterial *Methanocaldococcus jannaschii* EMG1 homologue was published shortly after, this time showing the structure of the dimer[119]. Dimerization is achieved by hydrophobic interactions at the interface of the two monomers, between a pair of parallel α-helices from each subunit, forming a four-helix bundle with each pair of helices positioned perpendicular to the other. Several hydrogen bonds also contribute to the dimer stabilization. As is common with SPOUT methyltransferases, dimerization forms a continuous, positively charged groove at the interface between the subunits, proposed to be the binding site for nucleic acids.

In 2009, Wurm *et al*[120] used a combination of fluorescence- and nuclear magnetic resonance- spectrometry and a yeast-three-hybrid assay, to determine the specific 18S rRNA sequence which constituted the EMG1 substrate in *M. jannaschii*. RNA methylation assays showed that the consensus sequence was methylated at position 914 only in the presence of a pseudouridine (ψ) nucleotide, and reversed-phase high-performance liquid chromatography (HPLC) combined with ^{1}H-nuclear magnetic resonance spectroscopy identified N1 as the methylated atom. Thus, EMG1 is the first known N1-specific pseudouridine methyltransferase. This specific reaction was confirmed for human EMG1 using a human 18S rRNA sequence, identifying ψ 1248 as the target nucleotide. In a paper which constituted the first published structure of a

SPOUT methyltransferase bound to RNA, Thomas *et al*[121] determined the X-ray crystal structures of the *S. cerevisiae* EMG1 dimer in complex with one molecule of RNA at 1.9Å, and with two molecules of RNA at 3.0Å, using the RNA substrate analog sequence 5′-GGGCUUCAACGCCC-3′. The bound RNA forms a stem-loop structure, with EMG1 specifically recognizing and contacting the central UUCAAC in the RNA sequence. *In vivo,* the second uridine is replaced by pseudouridine. When bound to EMG1, the uridine at this position is flipped out from the RNA loop and bound in a pocket formed between the two monomers, near the methyl donor SAM. Meyer *et al*[72] later confirmed the EMG1 substrate *in vivo* using reversed-phase HPLC and electrospray ionisation mass spectroscopy.

Based on the crystal structure of EMG1 in complex with RNA, it was proposed that EMG1 may play a role in rRNA folding in addition to its function as a methyltransferase[121]. This model proposes that EMG1 contacts the 18S rRNA at two sites, the first of which is a loop containing the methylated pseudouridine (ψ1248 in humans). Contact with the second proposed site, the stem of helix 42, could break apart the helix upon EMG1 binding, changing interactions with helix 41 and promoting binding of ribosomal protein RPS19. This is in agreement with a previous model proposing that EMG1 binding to 18S rRNA promotes the association of small subunit ribosomal protein RPS19[116].

1.4.3 Importance of aspartate 86 in EMG1

Upon EMG1dimerization an extended basic surface is formed, which was proposed to be the rRNA binding site. This positively charged area is centered on a group of arginines from both monomers, one of which belongs to an Arginine-Proline-Aspartate-Isoleucine (RPDI) sequence which is completely conserved among EMG1 homologues[119] (asterisks in Figure 3.1 C). In both *S. cerevisiae* and *M. janaschii*, the RNA binding site is disrupted by mutation of the arginine residue (R84 in human EMG1)[117,119]. Crystal structures showed that the substrate uridine in the RNA analog sequence is held and stabilized by hydrogen bonds formed with this arginine residue[121], which in turn is oriented and held by strong hydrogen bonds with the side chain of the aspartate (D86 in human EMG1)[119], similar to the conformation found at the catalytic arginine in TrmD, another SPOUT-class methyltransferase. Substitution of the aspartate could alter RNA specificity or affinity, thereby affecting the methylation activity of the protein[120].

Meyer *et al*[72] studied the effects of mutating the aspartate in the conserved catalytic loop RPDI sequence in yeast, showing by analytical gel filtration that a D90G substitution (D86G in human EMG1) increases dimerization and affinity for the 18S rRNA substrate. Intriguingly, matrix-assisted laser desorption/ionization mass spectrometry results suggest that the mutated EMG1 methylates RNA as efficiently as the wild type protein, pointing to a structural, rather than a catalytic role of the D86G substitution in causing BCS. Additionally, the mutated protein mislocalizes to the nucleoplasm and partially to the cytoplasm in this system.

1.5 Animal models of BCS

1.5.1 Emg1 null mouse

An *Emg1* null mouse was generated in collaboration with Dr. Hao Ding, which replaced exons 2-6 of the mouse *Emg1* gene with a cassette containing a splice acceptor, an internal ribosome entry site, and a β-galactosidase-neomycin resistance fusion gene (SA-IRES-βgeo). Knockout mice are not viable and in fact, embryos arrest growth at the morula stage between E2.5 and E3.5, indicating that EMG1 is necessary for preimplantation development in mice[122]. Although no differences in rRNA levels between wild type or heterozygous embryos and *Emg1* null embryos could be found, this does not exclude the possibility of delayed rRNA processing, which was not examined.

Although the embryos did not survive long enough to allow both genotyping and detailed biochemical analyses, the presence of a β-galactosidase reporter under the control of the EMG1 promoter meant that examination of *Emg1* expression patterns was possible in heterozygotes. Using both X-gal staining and whole-mount RNA *in situ* hybridization, it was shown that *Emg1* is expressed in the inner cell mass of blastocysts, but not in the trophectoderm. In later embryos, *Emg1* is ubiquitously expressed with no clear tissue-specific patterns; however strong expression is found in the ventricular zone of the neuroepithelium, the neural layer of the retina, the follicles of vibrissae, the thymus, submandibular glands, brown adipose tissue, lung, nephric tubules, renal mesenchyme, and seminiferous tubules of the testes. Interestingly, expression in the adult, although

found in most tissues upon Northern blot, exhibits cell-specific localization by RNA *in situ* hybridization. In the brain, expression is restricted to the granular cells of the cerebellum and to the hippocampus, and in intestine is found in the crypts. In the testes, expression is high in spermatogonia but not in late stage spermatocytes[122]. This pattern indicates specific expression of *Emg1* in highly proliferative cell populations in the adult, while expression in embryos is more widespread.

1.5.2 Trp53 knockout attempted rescue of Emg1 null mouse

Previous mouse models of ribosome biogenesis defects could be rescued by ablating *Trp53*, removing a crucial cell-cycle checkpoint and allowing embryos to continue developing[95,105,123,124]. *Emg1* heterozygous mice were therefore crossed with *Trp53* knockout mice in an effort to rescue the early lethality. However, no *Emg1* null mice were born and embryos did not develop beyond E3.5, indicating that the *Emg1* defect is not completely mediated by the TRP53 pathway under these conditions. Shwachman-Diamond syndrome also seems to be TP53-independent[109], though it remains to be determined how growth arrest and cell death are mediated in these models.

1.6 Rationale, hypotheses and objectives

Bowen-Conradi syndrome is a significant burden on the Hutterite population, with an estimated carrier frequency of 1 in 10, and prevention or treatment require a detailed understanding of the molecular causes of the disorder. A study by Lamont *et al*[16,125] determined that the mutation responsible for BCS lay in a 1.9 megabase interval on the short arm of chromosome 12. This candidate interval was the basis of the search for the

mutation causing BCS, as described in Chapter 3. Evidence showed that BCS belongs to a diverse group of disorders with a similar molecular basis, known as the ribosomopathies. While it is well established that these diseases all stem from mutations affecting ribosome assembly and function, many aspects of ribosome biogenesis disorders remain obscure. Research to uncover the specific contributions of genes associated with ribosomopathies is shedding light on the manifold roles of ribosomes in the cell cycle, cell proliferation, and development of different tissues, tightly regulated in a temporal and spatial manner. Insights into ribosome biogenesis in the context of the etiology of BCS will aid not only the Hutterite population, but the wider spectrum of ribosomopathies as well.

The hypotheses explored herein are:

- That a single nucleotide mutation in the gene *EMG1* is responsible for Bowen-Conradi syndrome.
- That the mutation in *EMG1* results in EMG1 protein instability, leading to severely reduced levels of the protein in BCS patient cells.
- That EMG1 protein instability alters its turnover rate and localization, and negatively impacts cell proliferation due to impaired ribosome biogenesis in BCS patient cells.
- That the mutation in *EMG1* leads to reduced cell proliferation during development, resulting in growth retardation.

To examine these hypotheses, the objectives of the research discussed in this thesis are as follows:

(i) To determine the mutation responsible for causing Bowen-Conradi syndrome, to confirm its impact in BCS patient cells, and to infer how it might cause disease based on the known properties of the affected gene.

(ii) To establish the impact of the BCS-causing A>G mutation in *EMG1* on EMG1 protein stability, localization, and association with other proteins.

(iii) To determine the effect of reduced levels of EMG1 protein on cell proliferation and cell cycle progression, protein synthesis rates, and ribosome biogenesis in BCS patient cells.

(iv) To generate and characterize a mouse model of BCS, and to examine the evolution of the disease throughout development and in multiple tissues.

Chapter 2 : Materials and Methods

2.1 Study subjects

Lymphoblast or fibroblast cell lines from patients were generated as part of a previous study to map the BCS gene[16]. Signed informed consent was obtained for all participants. This study was approved by the Health Research Ethics Board at the University of Manitoba. Anonymous control DNA samples were from non-Hutterite Manitobans and individuals from the Eastern United States, and control lymphoblasts and fibroblasts were from children of a similar age as BCS patients.

2.2 Bioinformatics

Protein orthologs of EMG1 from various species were identified using the yeast Emg1 protein sequence (NP_013287.1) as a query for a basic local alignment search tool (BLAST) search[126] (NCBI). Sequences from a broad range of species were selected from the resulting homologues and a multiple sequence alignment was performed using Clustal W[127].

2.3 Cell culture

Baby hamster kidney (BHK), HeLa human cervical cancer, and fibroblast cells were maintained in minimal essential medium (alpha modification) or Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Sigma, or Lonza). Lymphoblasts were maintained in Roswell Park Memorial Institute (RPMI) medium 1640 (Invitrogen, Sigma). All media was supplemented with 10% fetal bovine serum, 100 U penicillin per ml, and 100 μg streptomycin per ml (Invitrogen). Mouse embryonic stem cell cultures were grown in high-glucose DMEM supplemented with 2 mM glutamine, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol, 1 mM sodium pyruvate, 50 U/ml penicillin, 50 ug/ml streptomycin, 15% fetal bovine serum, and 1000 units/ml leukemia inhibitory factor, on

mitotically inactive mouse embryonic fibroblast feeder cells[128]. Cells were maintained in a humidified chamber at 37°C in 5% CO_2. To passage adherent cells, tissue culture plates were first rinsed with 1 x phosphate buffered saline (PBS, 1.9 mM NaH_2PO_4, 8.1 mM Na_2HPO_4, 150 mM NaCl, pH 7.4), then incubated with 0.25% trypsin (Invitrogen) at 37°C for several minutes until the cells were visibly detached from the plate. Cells were resuspended in the appropriate culture medium at the required dilution before re-plating.

2.4 Cell proliferation assay

Cell proliferation rates were determined using a CellTiter 96 Non-Radioactive Cell Proliferation Assay (MTT) kit (Promega) and following the manufacturer's instructions. Briefly, equal numbers of cells were seeded into the wells of a 24-well plate and allowed to grow at 37°C for 0, 24, 48 or 72 hours. At each time point, the cells were incubated at 37°C for four hours in 150 µl MTT reagent per well, followed by 1 ml stop solution for one hour. The absorbance of each well was read by spectrophotometer at 570 nm.

The doubling time of a cell population was calculated using GraphPad Prism software, from the following equation:

$$T_d = (t_2 - t_1) * \frac{\log(2)}{\log(\frac{q_2}{q_1})}.$$

where T_d = doubling time, t = time, and q = quantity.

2.5 DNA isolation

2.5.1 Plasmid DNA isolation from bacterial cultures

Plasmid DNA was isolated from bacterial cultures using either a Qiagen plasmid preparation kit, or a quick preparation method.

Using the Qiagen Midi or Maxi kit, bacteria from 100-250 ml overnight cultures were harvested by centrifugation at 6000 x *g* for 15 minutes at 4°C. The pellet was resuspended in 10 ml of buffer P1 (50 mM TrisCl pH 8.0, 10 mM EDTA, 100 µg/ml RNase A), and then lysed by inverting with ten ml of buffer P2 (200 mM NaOH, 1% SDS). After a five minute incubation at room temperature, ten ml of ice-cold buffer P3 (3 M potassium acetate, pH 5.0) was added, mixed by inversion and loaded on a QIAGEN-tip pre-equilibrated with 10 ml buffer QBT (750 mM NaCl; 50 mM MOPS, pH 7.0; 15% isopropanol; 0.15% Triton X-100). The column was allowed to empty by gravity flow. The QIAGEN-tip was washed twice with 30 ml of buffer QC (1 M NaCl; 50 mM MOPS, pH 7.0; 15% isopropanol), and the DNA was eluted with 15 ml of buffer QN (1.6 M NaCl; 50 mM MOPS, pH7.0; 15% isopropanol). DNA was precipitated by adding 0.7 volumes of isopropanol, and pelleted by centrifugation at 15 000 x *g* for 30 min at 4°C. After decanting the isopropanol, the DNA pellet was washed with 5 ml of 70% ethanol, followed by centrifugation at 15 000 x *g* for 10 minutes at 4°C. The DNA was air-dried for 10 minutes and dissolved in TE (10 mM TrisCl, pH 8.0; 1 mM EDTA) or water.

For the quick preparation, 1 ml of an overnight bacterial culture was transferred into a 1.5 ml microfuge tube, and the cells were pelleted by centrifugation. The supernatant was discarded and the pellet resuspended in 200 µl P1 buffer. An equal volume of P2 buffer

was added and mixed by inversion. After a two minute incubation at room temperature, 200 µl of P3 buffer was added, the tubes were mixed, and placed on ice for 15 minutes. The cell debris was pelleted for 15 minutes in a microcentrifuge at maximum speed, and the supernatant was decanted into a new microfuge tube and the DNA was precipitated with 0.8 volumes of isopropanol. The DNA was then collected by centrifugation for 15 minutes in a microcentrifuge at maximum speed. The isopropanol was decanted, and the pellet was washed with ice-cold 70% ethanol. After a final two-minute centrifugation, the DNA pellet was air dried and dissolved in 50 µl water.

2.5.2 Genomic DNA isolation from mouse tails

Tail clips collected from 21 day-old mice were incubated overnight at 55°C in 0.5 ml digestion buffer (100 mM Tris pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 µg/ml proteinase K). The following morning, 0.5 ml buffered phenol/chloroform/isoamyl alcohol (25:24:1) was added, and the tubes were rotated for 1 hour at room temperature. The tubes were spun at 13 000 x *g* for 10 minutes, and the upper aqueous phase transferred to a second tube. After adding 0.5 ml isopropanol, the tubes were inverted several times. The DNA was pelleted with a 5 minute centrifugation, and the supernatant was discarded. The pellet was washed with 0.75 ml of 70% ethanol, followed by a second 5-minute centrifugation. The ethanol supernatant was removed and the DNA pellet was air dried for 30 minutes. The DNA was resuspended in 250-500 µl of RNase/DNase-free water.

Alternatively, the phenol/chloroform/isoamyl alcohol extraction step was omitted for faster DNA isolation. DNA was instead precipitated using isopropanol directly following digestion, and washed in 70% ethanol as above.

2.5.3 DNA isolation from agarose gels

DNA was isolated from fragments separated by agarose gel electrophoresis using a QIAquick gel extraction kit (Qiagen). The band was visualised under ultraviolet light and excised using a clean scalpel. The gel fragment was weighed, 3 volumes of buffer QG was added per volume of gel, and the gel fragment was dissolved by heating to 50°C in a water bath for ten minutes with occasional mixing. An equal volume of isopropanol was added, the sample was applied to a QIAquick column, and the liquid was eluted from the column by centrifugation at 13000 rpm for one minute; the eluate was discarded. After applying 0.5 ml buffer QG to the column and centrifugation for one minute, the eluate was again discarded. The column was washed with 0.75 ml buffer PE, the liquid was eluted by centrifugation for one minute, and the eluate was discarded. The DNA was eluted by adding 50 µl of water to the centre of the column, letting it incubate at room temperature for one minute, and then collecting the DNA by centrifugation for one minute in a DNase-free microcentrifuge tube.

2.6 RNA isolation

RNA was isolated from cells using TRIzol reagent (Invitrogen) and following the manufacturer's instructions. Cells grown in suspension were pelleted and the supernatant removed before lysis. Adherent cells were lysed directly in the culture vessel by addition of 1 ml TRIzol, followed by repeated up and down pipetting. The lysates were stored at -80°C until the RNA was needed. After thawing, 0.2 ml of chloroform was added to lysates, these were mixed by hand for 15 seconds, and incubated at room temperature for 3 minutes. The phases were separated by centrifugation at 12 000 x *g* for 15 minutes at 4°C, and the upper aqueous phase was then transferred to a new tube. The RNA was

precipitated by the addition of 0.5 ml isopropanol per tube, followed by mixing by rotation for ten minutes at room temperature. The RNA was then collected by centrifugation at 12 000 x *g* for 10 min at 4°C, and the supernatant decanted. The RNA pellet was then washed with 1 ml 75% ethanol, mixed with a vortex, and collected by centrifugation at 7 500 x *g* for 5 minutes at 4°C. The ethanol was removed and the pellet was allowed to air dry for 10 minutes, followed by resuspension in diethylpyrocarbonate (DEPC)-treated water by pipetting up and down several times. The RNA solution was then incubated for 10 minutes in a 60°C dry bath, and then stored at -80°C.

2.7 Agarose gel electrophoresis

2.7.1 DNA separation

Agarose gels were prepared at concentrations of 0.7% to 2.5% agarose (Invitrogen) depending on the size of the DNA fragments to be separated. Agarose was added to an appropriate volume of 1x Tris Borate EDTA (TBE) (90 mM Tris-borate; 2 mM EDTA pH 8.3) or Tris Acetate EDTA (TAE) (40 mM Tris-acetate; 1 mM EDTA pH 8.3) and heated to dissolve the agarose. The solution was cooled to approximately 60°C, ethidium bromide was added to a final concentration of 0.5 μg/ml and the solution was poured into a gel tray containing the appropriate combs. After the gel solidified, it was placed in a horizontal electrophoresis apparatus and covered to a depth of roughly 1 mm above the gel surface with 1x TBE or TAE.

DNA was mixed 1 in 6 with 6x loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol) and was pipetted into the wells of the agarose gel. A DNA ladder covering a similar range in size as the DNA fragments to be separated was loaded into the

outside lanes of the gel (Invitrogen, New England Biolabs, Frogga Biosystems). A voltage of 5V per cm was applied, and the migration of the dye markers was monitored to estimate the distance that the DNA fragments had migrated.

2.7.2 RNA separation

To prepare RNA for separation by gel electrophoresis, equal volumes of RNA and NorthernMax-Gly Sample Loading Dye (Ambion) were mixed in a small tube and incubated at 50°C for 30 minutes to denature secondary structures. RNA samples were then separated at 5V per cm in a 0.8 - 1.0% agarose BPTE gel (100 mM PIPES, 300 mM Bis-Tris, 10 mM EDTA pH 8.0) gel in 1 x BPTE buffer.

2.8 Capillary electrophoresis of RNA

Relative levels of 18S and 28S rRNA were quantified using the RNA 6000 Nano LabChip Kit in the Agilent 2100 Bioanalyzer and following the manufacturer's instructions. One µl of total RNA was added to duplicate wells of the Nano Chip along with 1 µl of RNA ladder, and RNA species were separated by capillary electrophoresis. The quality of the RNA separation was assessed by resolution of the individual peaks in the ladder. The area of each peak was determined by the Agilent Bioanalyer software, and the ratios of 28S/18S rRNA were calculated.

2.9 Quantification of nucleic acid concentrations

DNA concentration was estimated by comparison with a High or Low DNA Mass Ladder (Invitrogen). Each band in these ladders contains a specified amount of DNA, and by comparing the intensity of the band of interest with a ladder band of a similar size, the amount of DNA in the band of interest can be estimated. Alternatively, DNA or RNA

concentrations were determined by reading the absorbance of a sample at 260 nm using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The absorbance at 280 nm was also read, and the 260/280 ratio was used as an estimate of nucleic acid purity.

2.10 Polymerase chain reaction (PCR)

PCR amplification of DNA was performed in either a Bio-Rad MyCycler or an Eppendorf thermal cycler. A negative control without DNA was included in each experiment, as well as a positive and a negative control for the experiment whenever possible. Cycling conditions, including annealing temperatures and elongation times, were optimized for each amplicon. For each PCR reaction, 49 µl of reaction mixture (100 ng forward primer, 100 ng reverse primer, 0.2 mM dNTPs, 2.5 units Taq DNA polymerase [New England Biolabs], 1 x ThermoPol buffer [New England Biolabs]) was added to 1 µl of DNA template. Where necessary, high-fidelity polymerase (Phusion Taq, New England Biolabs) was employed to minimize errors in the PCR product. A typical reaction consisted of an initial 5 minute denaturation step at 95°C, followed by 30-35 cycles of 1 minute at 95°C, 1 minute at the optimized annealing temperature, and 2 minutes at 72°C. The final cycle was followed by a 5 minute extension at 72°C, and the products were held at 4°C. See Table 2.1 for PCR primers used throughout this thesis.

ID	Sequence (5 to 3')	Purpose
WPG 662	5' CGG GAT CCA TGG CGA AGG CGA AGA AG 3'	Human NOP14 w/5'BamHI site F
WPG 663	5' CGG GAT ATC TAT AGC ATT ATT ACT CTA AA 3'	Human NOP14 w/3'EcoRV site R
WPG 707	5'TCA CCA CAG CTT TTG AAG AAG TAT GG 3'	Mouse EMG1 3' arm amplification F

WPG 708	5' GGC GCG CCG GTC TGC ACC AGC AGG CCT AGC TCA TA 3'	Mouse EMG1 3' arm amplification R
WPG 709	5' CTG GCA GCT TGT TCT GTA TTT ACC 3'	Mouse EMG1 5' arm amplification F
WPG 710	5' GGC TAT CAC TCA AGG TCA CCT AAG 3'	Mouse EMG1 5' arm amplification R
WPG 717	5' GTG GGC TCT ATG GCT TCT GA 3'	EMG1 exon 2-6_DTG: seq NotI site F
WPG 718	5' GTT TTC CCA GTC ACG ACG TT 3'	EMG1 exon 2-6_DTG: seq NotI site R
WPG 719	5' CAA ACT CTT CGC GGT CTT TC 3'	Mouse EMG1 sequencing
WPG 720	5' TAC GCT TGA GGA GAG CCA TT 3'	Mouse EMG1 sequencing
WPG 721	5' ACC CCA TTG TAT GGG ATC TG 3'	Mouse EMG1 sequencing
WPG 722	5' GTT GCA CCA CAG ATG AAA CG 3'	Mouse EMG1 sequencing
WPG 723	5' GTT TCA TCT GTG GTG CAA CG 3'	Mouse EMG1 sequencing
WPG 724	5' CAG CAG CAG ACC ATT TTC AA 3'	Mouse EMG1 sequencing
WPG 725	5' GTG CAG ATT GAA AAT GGT CT 3'	Mouse EMG1 sequencing
WPG 726	5' TCT CTC CAG GTA GCG AAA GC 3'	Mouse EMG1 sequencing
WPG 727	5' CCG ATA TTA TTT GCC CGA TG 3'	Mouse EMG1 sequencing
WPG 728	5' TAC GCG TAC TGT GAG CCA GA 3'	Mouse EMG1 sequencing

WPG 729	5' TTG AAC TGC CTG AAC TAC CG 3'	Mouse EMG1 sequencing
WPG 730	5' CTC GCC ACT TCA ACA TCA AC 3'	Mouse EMG1 sequencing
WPG 731	5' ATG GCG ATT ACC GTT GAT GT 3'	Mouse EMG1 sequencing
WPG 732	5' TGT CTG TTG TGC CCA GTC AT 3'	Mouse EMG1 sequencing
WPG 733	5' AGA CAA TCG GCT GCT CTG AT 3'	Mouse EMG1 sequencing
WPG 734	5' AGC CAA CGC TAT GTC CTG AT 3'	Mouse EMG1 sequencing
WPG 735	5' CGT TGG CTA CCC GTG ATA TT 3'	Mouse EMG1 sequencing
WPG 736	5' AAG GCC TAT GGC TAT CAC TCA A 3'	Mouse EMG1 sequencing
WPG 737	5' TGT GGC ATG ACT TGT CTC CT 3'	Mouse EMG1 sequencing
WPG 738	5' GTC ACT CCC CAC TGA AGG AA 3'	Mouse EMG1 sequencing
WPG 739	5' CTT CAG TGG GGA GTG ACT GG 3'	Mouse EMG1 sequencing
WPG 740	5' CAG CTA CCG CCT GTC TTA CC 3'	Mouse EMG1 sequencing
WPG 741	5' CAG GAG CAG TTG GCA TTT TA 3'	Mouse EMG1 sequencing
WPG 742	5' TCA GGA ACG CCT ACT GTC AA 3'	Mouse EMG1 sequencing
WPG 743	5' CCC CTT AGA TTC TGG TGT TCC 3'	Mouse EMG1 sequencing

WPG 744	5' GGG GAC ACT GTA GGT TCT GG 3'	Mouse EMG1 sequencing
WPG 745	5' TCC CCT TTT CTC AGC TTG TC 3'	Mouse EMG1 sequencing
WPG 746	5' AAG GAGA CCA AGC TGC TAT TG 3'	Mouse EMG1 sequencing
WPG 747	5' CAA TAG CAG CTT GGC TCC TT 3'	Mouse EMG1 sequencing
WPG 748	5' CAT CCC CGA ACC TTT TGA TA 3'	Mouse EMG1 sequencing
WPG 749	5' GGA GCA AGT CGG AAC AGA AG 3'	Mouse EMG1 sequencing
WPG 750	5' CCC CCA GAA TAG AAT GAC ACC 3'	Mouse EMG1 sequencing
WPG 751	5' GGG AGG ATT GGG AAG ACA AT 3'	Mouse EMG1 sequencing
WPG 752	5' TTG AGT GTT GTT CCA GTT TGG 3'	Mouse EMG1 sequencing
WPG 753	5' TCT TGT CTG ATT CTA GAA CT AGG TG 3'	Mouse EMG1 sequencing
WPG 754	5' GCT CTG CCC ACA CAG ATA GCT 3'	Mouse EMG1 sequencing
WPG 755	5' GCT AGC TTG GCT GGA CGT AA 3'	Mouse EMG1 sequencing
WPG 756	5' CGA TTG TAT ATG TGC CGA TAC C 3'	Mouse EMG1 sequencing
WPG 757	5' ACT GGC ATG TTC CTC AGC AT 3'	Mouse EMG1 sequencing
WPG 758	5' GAG CTT GCG GAA CCC TTA AT 3'	Mouse EMG1 sequencing

WPG 759	5' GAA GGA AGT TGG CCG TGA G 3'	Mouse EMG1 sequencing/ 5' external probe F
WPG 760	5' CCA GTC AAG CCC CAT TTA AC 3'	Mouse EMG1 sequencing/ 5' external probe R
WPG 761	5' GTT TGG CAT CTT CCC TGG TA 3'	Mouse EMG1 sequencing/ 3' external probe F
WPG 762	5' CCC CAC AGA GCC AAA ATA AA 3'	Mouse EMG1 sequencing/ 3' external probe R
WPG 763	5' GAG AAT CCG ACG GGT TGT TA 3'	Mouse EMG1 sequencing/ Bgeo probe F
WPG 764	5' CAG CAG CAG ACC ATT TTC AA 3'	Mouse EMG1 sequencing/ Bgeo probe R
WPG 765	5' TGA CCT TTA GCA AGC ATG GAG 3'	BCS Mouse 2nd lox p site detection F alternative
WPG 766	5' CGA TTG TAT ATG TGC CGA TAC C 3'	BCS Mouse 2nd lox p site detection R alternative
WPG 767	5' TGA CCT TTA GCA AGC ATG GAG 3'	BCS Mouse 2nd lox p site detection F #2
WPG 768	5' CTA TCT TGG AAC TAT ACC AG 3'	BCS Mouse 2nd lox p site detection R #2
WPG 769	5' CAC TGA TAT TGT AAG TAG TTT GC 3'	Flp allele at ROSA26 in B6.12954 mice F
WPG 770	5'-CTA GTG CGA AGT AGT GAT CAG G3'	Flp allele at ROSA26 in B6.12954 mice R
WPG 771	5' TGT TTT GGA GGC AGG AAG CAC TTG	WT ROSA26 allele F
WPG 772	5' AAA TAC TCC GAG GCG GAT CAC AAG 3'	WT ROSA26 allele R
WPG 773	5'ACT GGC ATG TTC CTC AGC AT 3'	BCS mouse single Frt site after Bgeo removal F

WPG 774	5' AAG CTT GCG GAA CCC TTA AT 3'	BCS mouse single Frt site after Bgeo removal R
WPG 800	5' CAT TTC CTG GCC CAC TTT C 3'	Human EMG1 sequencing exon 1 F
WPG 801	5' CCT CTC TGC TTT CTA CAT CC 3'	Human EMG1 sequencing exon 1 R
WPG 802	5' GGA CCA CAC TTG AGA AGA ACC AC 3'	Human EMG1 sequencing exon 2-3 F
WPG 803	5' TTC TCC CTT TTC CTC TTC CTC 3'	Human EMG1 sequencing exon 2-3 R
WPG 804	5' CCT TGT TCG ATG ACT GGA CAG 3'	Human EMG1 sequencing exon 4-5 F
WPG 805	5' GCA TTG TTA AAA TGC CTA CTG C 3'	Human EMG1 sequencing exon 4-5 R
WPG 806	5' ACT GGC ATG TTC CTC AGC AT 3'	BCS mouse bgeo cassette detection F
WPG 807	5' GGA CAG GGA TAA GTA TGA CAT C 3'	BCS mouse bgeo cassette detection R
WPG 808	5' CAT GTT CTG TTC TGC AGG TAG 3'	BCS mouse A>G mutation detection F
WPG 809	5' GCT CTG CCC ACA CAG ATA GCT 3'	BCS mouse A>G mutation detection R
WPG 838	5' GAA GAC ATA TGA GCT ACT CAA CTG 3'	Human EMG1 257 A>G mutation exon 2/intron 2 F
WPG 839	5' GAG TTA CCT GGT GGG TG**G** TA 3'	Human EMG1 257 A>G mutation exon 2/intron 2 R
WPG A	5′ CTG GCA TGT TCCTCA GCA TA 3′	Mouse EMG1 mutation introduction segment 1 F
WPG B	5′ GTG GGT GAT GTC TGG TCT GAC 3′	Mouse EMG1 mutation introduction segment 1 R

WPGC	5′ GTC AGA CCA GAC ATC ACC CAC 3′	Mouse EMG1 mutation introduction segment 2 F
WPGD	5′ CCA GAT AAA ATC ACT TCT GTT TCT TC 3′	Mouse EMG1 mutation introduction segment 2 R
WPGE	5' GTT TTC CTG CCC TAA AGA AGG 3'	Human EMG1 sequencing exon 6 F
WPGF	5' CTC TCC CCA TCT AGC CAG G 3'	Human EMG1 sequencing exon 6 R

Table 2.1 Oligonucleotides used as PCR primers

2.11 Sequencing

To screen for the BCS-causing mutation, the coding region of candidate genes was amplified by PCR with primer pairs placed at least 30 base pairs away from each intron-exon junction. Exons were amplified together if they were separated by only a small intron. PCR products to be sequenced were extracted from agarose gels as above, and submitted to The Centre for Applied Genomics in Toronto for sequencing. Sequencing was performed with the same primers used for the amplification using the Sanger method and reaction products were analyzed on an ABI 3730XL instrument.

2.12 Vector construction

2.12.1 Restriction enzyme digestion

DNA digests were performed according to the manufacturer's instructions (New England Biolabs or Fermentas Life Sciences). Generally, 1-10 units of restriction enzyme were mixed with 1 µg DNA in 1x enzyme reaction buffer and incubated from one hour to overnight under the supplier-recommended conditions. The enzyme buffer varied according to the enzyme being used, and double digests were performed in a single buffer

that was compatible with both enzymes (see Table 2.2). When required, DNA was purified from the reaction components by gel extraction.

Enzyme	Site (5' to 3')	Source	T°	BSA	Buffer
Acc65I	G^GTACC	NEB	37	y	3
AscI	GG^CGCGCC	NEB or Fermentas	37	n	4 or Fast Digest
BamHI	G^GATCC	NEB	37	y	3
BlpI	GC^TNAGC	NEB	37	n	4
BspHI	T^CATGA	NEB or Fermentas	37	n	4 or Fast Digest
BsrG1	T^GTACA	NEB	37	y	2
EagI	C^GGCCG	NEB	37	n	3
HindIII	A^AGCTT	NEB	37	n	2
KpnI	GGTAC^C	NEB or Fermentas	37	y	1 or Fast Digest
NcoI	C^CATGG	NEB	37	n	3
NheI	G^CTAGC	NEB	37	y	2
NotI	GC^GGCCGC	NEB	37	y	3
NspI	RCATG^Y	NEB	37	y	2
PacI	TTAAT^TAA	NEB	37	y	1
PmlI	CAC^GTG	NEB	37	y	1
SacII	CCGC^GG	NEB	37	n	4
SalI	G^TCGAC	NEB or Fermentas	37	y	3 or Fast Digest
SbfI	CCTGCA^GG	NEB	37	n	4
ScaI	AGT^ACT	NEB or Fermentas	37	n	3 or Fast Digest
ScrF1	CC^NGG	NEB	37	n	4
XhoI	C^TCGAG	NEB	37	y	4

ZraI	GAC^GTC	NEB	37	n	4

Table 2.2 Restriction enzymes

2.12.2 Creation of blunt ends

To prepare DNA for blunt-end ligation, the 5′ overhang was filled in or the 3′ overhang was removed using the Klenow fragment of DNA polymerase I (New England Biolabs). One unit of enzyme per µg of DNA and 33 µM of each dNTP in 1x buffer 2 (New England Biolabs) was incubated at room temperature for 15 minutes. EDTA was added to a final concentration of 10 nM and the enzyme was heat inactivated at 75°C for 20 minutes.

2.12.3 Dephosphorylation of vector ends

To prevent restriction enzyme-digested vector from self-ligating, the 5′ terminal phosphate was removed using calf intestinal alkaline phosphatase (New England Biolabs) or Antarctic phosphatase (New England Biolabs). One unit of enzyme was added per µg of restriction enzyme-digested DNA and incubated at 37°C for 30 minutes. Antarctic phosphatase was heat inactivated at 65°C for 5 minutes.

2.12.4 Ligation

Using a molar ratio between 3:1 and 6:1 (insert:vector), DNA was ligated using T4 DNA ligase (New England Biolabs). One unit of enzyme was added per µg of the insert and vector mixture in 1x T4 ligase buffer and incubated at room temperature overnight. Alternatively, ligations were performed using a DNA ligation kit (Takara) at 16°C for 30 minutes.

2.12.5 HA-tagged human EMG1

Mammalian expression vectors containing the wild type and A>G-mutated *EMG1* cDNA sequences were constructed using a PCR-based strategy. Briefly, total RNA isolated from normal or patient fibroblasts was reverse transcribed with Superscript (Invitrogen). The coding sequence was amplified using oligonucleotide primers that added BamHI and XhoI sites onto the 5′ and 3′ ends, respectively. The PCR products were isolated from agarose gels, and digested with BamHI and XhoI to generate the 755 bp insert. The fragments were ligated with BamHI/XhoI digested pcDNA3.1 (Invitrogen) to generate pcDNAEmg1 and pcDNAEmg1/D86G constructs. To generate constructs that had an HA tag added to the C-terminus of EMG1, these vectors were used as the templates for PCR-amplification using Phusion polymerase. The products were digested with KpnI and NotI and cloned into KpnI/NotI- digested pMH plasmid (Roche Applied Science) to generate pMH-EMG1 and pMH-EMG1/D86G plasmids. The full EMG1 coding region was sequenced to confirm the presence of only the desired A to G substitution in the pMH-EMG1/D86G vector. The HA-tagged EMG1 was expected to produce a 28 kDa protein. A β-galactosidase expression plasmid for a transfection control, pRCMVβ-gal, was purchased from Invitrogen.

2.12.6 Flag-tagged human NOP14

The human NOP14 coding sequence (cDNA clone MGC: 39284 IMAGE: 4878317) was purchased from Open Biosystems. Because an antibody toward NOP14 was not commercially available, a FLAG-tag was added to the N-terminus of NOP14 to allow its detection. Briefly, NOP14 cDNA was PCR amplified using primers WPG662 and WPG663 to add a 5′ BamHI site and a 3′ EcoRV site, respectively. The PCR product was

gel purified and digested with BamHI and EcoRV, generating a 2427 bp NOP14 insert. The pCMVTag2B (Stratagene) plasmid containing the FLAG-tag coding sequence was digested using BamHI and EcoRV, gel purified, and ligated with the NOP14 insert. After transformation into *E. coli*, colonies were selected on 50 µg/ml kanamycin. Plasmids were prepared from overnight cultures of individual colonies, and separated by gel electrophoresis. Plasmids of the expected size were digested with BamHI and EcoRV to verify the presence of the NOP14 insert, and the final construct was sequenced to confirm that there were no coding sequence changes. The FLAG-tagged NOP14 was expected to produce a protein of 97 kDa.

2.13 Transfection and transformation

2.13.1 Transformation of bacterial cells

Electrocompetent *Escherichia coli* (DH10β, New England Biolabs) were used for transformations. One µl of ligation mixture was added to 40 µl of bacteria in a chilled electroporation cuvette with a 1 mm gap. Electroporation was performed by subjecting the mixture to 1.3 kV, 129 Ω, 50 µF using a BTX Electro Cell Manipulator 600. Following electroporation, the bacteria were immediately mixed with 960 µl of prewarmed SOC broth (Sigma) and incubated at 37°C for one hour. The transformed mixture was plated onto Miller's Lysogeny Broth (LB) agar containing the appropriate antibiotic and incubated at 37°C. The following day, individual colonies were subcultured into 3-10 ml LB broth containing the appropriate antibiotic and incubated with shaking at 37°C for 8 hours. The cultures served as pre-culture for larger cultures.

Alternatively, chemically competent cells (One Shot, Invitrogen) were employed. One µl ligation mixture was added to a vial of One Shot cells, mixed gently, and incubated on ice for 30 minutes. The cells were heat shocked by incubation at 42°C for 30 seconds, then placed on ice for two minutes. After adding 250 µl of pre-warmed SOC broth to the vial, cells were incubated with shaking at 37°C for one hour. The bacteria were spread onto agar plates as above.

2.13.2 Transfection of mammalian cells

BHK cells were cultured overnight in DMEM without antibiotics, and were approximately 90% confluent on the day of the transfection. Cells were washed, and then incubated in 1.5 ml of serum-free Opti-MEM (Invitrogen) for 45 minutes. During the incubation, samples for transfection were prepared as follows for each well: 4.5 µg of DNA (4 µg pMH vector, pMHEMG1, or pMHEMG1/D86G, plus 0.5 µg pRCMVβ-gal plasmid) was diluted in 250 µl of Opti-MEM and mixed gently. Ten µl of Lipofectamine 2000 (Invitrogen) was incubated in 250 µl Opti-MEM for 5 minutes at room temperature in a separate tube. The diluted DNA was mixed gently with the diluted Lipofectamine 2000 by pipetting up and down three times, and incubated at room temperature for 20 minutes. The DNA – Lipofectamine 2000 complexes were then added to each well and mixed by rocking the plate. After a 5 hour incubation at 37°C, the media was replaced with complete DMEM. The transfections were allowed to proceed for 24 hours. HeLa cells stably overexpressing the wild type or mutated EMG1 were obtained by transfection as above, and G418-resistant clones were selected.

2.14 Southern blot analysis

2.14.1 Gel electrophoresis of restriction enzyme-digested DNA

Sounthern blotting was performed using a modified protocol[129]. DNA that had been restriction-enzyme digested (see section 2.12.1) was mixed 1:6 with 6 x loading dye, and 35 µl was loaded in a 0.7% agarose gel in 1 x TBE, next to a HindIII λ ladder (New England Biolabs). DNA was separated by gel electrophoresis at 100V for 1 hour to run the DNA into the gel, followed by 60V overnight.

2.14.2 DNA transfer

Following DNA separation by gel electrophoresis, the gel was photographed under ultraviolet light with a ruler in the photograph to mark the position of the ladder in relation to the wells. Unnecessary portions of the gel were removed using a clean scalpel, and the top right corner was removed to mark the orientation. The gel was rinsed for five minutes with shaking in double-distilled H_2O. The gel was depurinated in 0.25 M HCl for 10 minutes, rinsed twice in double-distilled H_2O, and soaked in each of denaturation buffer (1.5 M NaCl, 0.5 M NaOH), followed by neutralization buffer (1 M NaCl, 0.5 M Tris, pH 7.4), for 30 minutes.

The gel was placed on a wick of filter paper that was soaked in 10 x SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0). A Hybond N^+ nylon membrane, cut to the same size as the gel and previously soaked in 10 x SSC, was placed on top. A glass pipette was used to remove all air bubbles between the gel and the membrane. Two sheets of filter paper were placed on top, followed by a layer of paper towels to act as a wick. Pieces of plastic

wrap surrounding the gel prevented the paper towel layer from contacting the filter paper wick. The DNA was allowed to transfer overnight at room temperature.

2.14.3 Labelling

The membrane was removed and rinsed in 5 x SSC for five minutes, then blotted on filter paper to dry. The DNA was crosslinked to the membrane in a Spectrolinker XL-1000 ultraviolet crosslinker (Spectronics Corporation) set on automatic "Optimal Crosslink"(120mJ/ cm^2), and the membrane was placed in a glass tube with 10 ml of PerfectHyb (Sigma). The tube was rotated in a hybridization oven at 65°C for four hours.

The DNA to be used as a probe was labelled using a Rediprime DNA labelling vial (GE Healthcare) and 5 µl of Easy Tides Deoxycytidine 5′ Triphosphate (Perkin Elmer) and following the manufacturer's instructions. Unincorporated nucleotides were removed using a Sephadex spin column (Illustra Probe Quant G-50 Micro Column, GE Healthcare). The purified probe was denatured by boiling for 5 min followed by cooling on ice, added to the PerfectHyb, and incubated overnight at 65°C with rotation. The next day, the membrane was washed in 0.2 x SSC; 0.1% SDS three times for 15 minutes each. The hybridized membrane was heat-sealed in a polyethylene bag and exposed to Kodak Biomax MS film at -80°C with two intensifying screens.

2.15 Protein isolation

Soluble and insoluble proteins were isolated from cultured cells by lysis in a modified radio immunoprecipitation assay (RIPA) buffer (1% tert-Octylphenoxy (poly)oxyethylene ethanol [IGEPAL]; 0.5% sodium deoxycholate; 0.1% SDS; 1.9 mM NaH_2PO_4; 8.1 mM Na_2HPO_4; 150 mM NaCl). RIPA-insoluble protein was separated by

centrifugation for 10 minutes in a microcentrifuge at 10000 x *g*, leaving the RIPA-soluble proteins in the supernatant. The pellet (insoluble fraction) was resuspended in RIPA buffer by sonication for 5 seconds. For isolation of cytoplasmic and nuclear fractions, cultured cells were lysed in Triton buffer (50 mM Tris pH 7.4; 1 mM EDTA; 150 mM NaCl; 1% Triton) and the fractions were separated by centrifugation at 2000 x *g* for 10 minutes in a microcentrifuge. The supernatant was removed and the pellet was resuspended by sonication for 5 seconds in nuclear lysis buffer (50 mM Tris pH 7.4; 10 mM EDTA; 1% SDS). All lysis buffers were supplemented 1 in 500 with a protease inhibitor cocktail (Sigma).

2.16 Determination of protein concentration

Protein concentration was determined using either the Bradford assay or the bicinchoninic acid (BCA) assay. In both cases, known quantities of bovine gamma globulin (Sigma) were used to create a standard curve, and each sample was read in duplicate or triplicate. One sample containing water and an appropriate volume of lysis buffer was used as a control. For the Bradford assay, 2 µl of the sample was added to 200 µl of Bradford reagent (Bio-Rad) and 798 µl water, mixed, incubated at room temperature for five minutes, and the absorbance was read at 595 nm using an Ultrospec 1000 spectrophotometer (Pharmacia Biotech).

The BCA assay kit (Thermo Scientific) was employed when protein lysates contained high concentrations of detergent. The working reagent was prepared by mixing appropriate volumes of reagent A and reagent B at a 50:1 ratio. After adding 10 µl of each sample to the wells of a 96-well flat-bottom microplate, 200 µl of the working reagent was added to each well. The covered plate was incubated at 37°C for 30 minutes,

cooled at 4°C for five minutes, and the absorbance was read at 562 nm using a plate reader (SpectraMax M2e, Molecular Devices)

2.17 Immunoblot analysis

2.17.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Polyacrylamide gels were prepared by casting a resolving gel at a concentration of acrylamide appropriate to resolve the size of the protein of interest (40 mM Tris pH 8.8, 0.1% SDS, 7.5% to 10% polyacrylamide [29:1 acrylamide:bisacrylamide], 0.075% ammonium persulfate, 0.1% TEMED). After allowing the gel to polymerize, a 4% stacking gel (60 mM Tris pH 6.8, 0.1% SDS, 4% polyacrylamide, 0.1% ammonium persulfate, 0.01% TEMED) was overlaid and a comb was inserted. Equal volumes of the cell lysates (20-25 μg of protein) were mixed with sample prep buffer (8% SDS, 250 mM Tris pH 6.8, 40% glycerol, 0.01 % bromophenol blue, 400 mM dithrioethiol), boiled for five minutes and cooled on ice. Samples and a protein ladder (Fisher Scientific) were pipetted into each well and the tank was filled with running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). The gel was run at 160V for approximately one hour, or until the bromophenol blue marker in the loading buffer reached the bottom of the gel.

2.17.2 Transfer of proteins to a nitrocellulose membrane

Following separation of proteins by SDS-PAGE, the gel was equilibrated in ice-cold *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS) buffer (0.01 M CAPS pH 11, 10% methanol) or ice-cold Tris-glycine transfer buffer (80% running buffer, 20% methanol) for 15 minutes. The proteins were transferred to a nitrocellulose membrane (Bio-Rad) at

100 V for 1 hour at 4°C using the modified method of Towbin *et al*[130]. Total protein was visualised by rinsing the membrane with Ponceau S (0.1% Ponceau S in 5% acetic acid).

2.17.3 Immunodetection of specific proteins

Membranes were blocked in 5% skim milk in Tris buffered saline Tween (TBST) (20 mM Tris pH 7.4; 0.15 M NaCl; 0.1% Tween) for one hour. Proteins of interest were detected using a primary antibody diluted in 5% skim milk powder in TBST overnight at 4°C. The following day, blots were washed three times with TBST, incubated with horseradish peroxidase (HRP)-conjugated secondary donkey anti-rabbit or donkey anti-mouse antibody at a dilution of 1:20000 for one hour, washed, and visualized using Immobilon western chemiluminescent HRP substrate (Millipore). The blot was exposed to Clear Blue X-ray film (Thermo Scientific) for 5 seconds to 10 minutes and then developed to reveal the bands of interest. See Table 2.3 for a list of antibodies used.

PRIMARY ANTIBODIES			
Identity	**Source**	**Company**	**Application**
anti-β-actin	mouse monoclonal	Sigma	1 in 2000
anti-β-galactosidase	chicken polyclonal	Aves	1 in 500
anti-BrdU	mouse monoclonal	Roche	1 in 3
anti-cleaved caspase 3	rabbit polyclonal	Cell Signaling Technology	1 in 500
anti-coilin	mouse monoclonal	Novus	1 in 50
anti-EMG1	rabbit polyclonal	Aviva Systems Biology	1 in 2000
anti-EMG1	rabbit polyclonal	Proteintech Group	1 in 500-1 in 1000

anti-fibrillarin	rabbit polyclonal	Aviva Systems Biology	1 in 2000
anti-FLAG Tag	rabbit polyclonal	Genscript	1 in 1000
Anti-HA Tag (HA.C5)	mouse monoclonal	Applied Biological Materials	1 in 2000
anti-HA Tag	rabbit polyclonal	Beckton Dickinson	1 in 3000
anti-NOP14	mouse polyclonal	Novus	1 in 100
anti-p53 (2B2.68)	mouse monoclonal	Santa Cruz	1 in 2000
anti-RPL7	rabbit polyclonal	Bethyl Labs	1 in 5000
anti-RPL26	rabbit polyclonal	Bethyl Labs	1 in 3000
anti-RPS6	rabbit polyclonal	Bethyl Labs	1 in 1000
anti-RPS19	rabbit polyclonal	Proteintech Group	1 in 3000
anti-SMN	goat polyclonal	Santa Cruz	1 in 250
anti-ubiquitin	rabbit polyclonal	StressMarq	1 in 1000-1 in 1500
SECONDARY ANTIBODIES			
Donkey anti-mouse HRP	donkey	Jackson Immunoresearch	1:10 000
Donkey anti-rabbit HRP	donkey	Jackson Immunoresearch	1: 10 000
Goat anti rabbit biotinylated	goat	Dako or Vector	1 in 1000
Rabbit anti mouse biotinylated	rabbit	Dako or Vector	1 in 1000 or 1 in250
Goat anti chicken biotinylated	goat	Aves	1 in 1000
Donkey anti-mouse Alexa Fluor 488	donkey	Molecular Probes	1 in 500
Donkey anti-mouse Alexa	donkey	Molecular Probes	1 in 500

Fluor 555			
Donkey anti-goat Alexa Fluor 488	donkey	Molecular Probes	1 in 500

Table 2.3 Antibodies

2.17.4 EMG1 antibody competition

To confirm the specificity of the anti-EMG1 antibody from Aviva, an aliquot of the antibody was pre-incubated with purified maltose binding protein-conjugated EMG1 (courtesy of Dr. Steven Pind) before being used in an immunoblot. The same blot was then re-probed with non-competed anti-EMG1 antibody to reveal specific EMG1 bands.

2.18 Immunoprecipitation

Protein G beads (GE Healthcare) were prepared by washing three times in PBS to remove ethanol, then resuspended in PBS to form a slurry. To prevent bacterial growth, 0.02% sodium azide was added. On the day of the experiment, protein G beads were incubated with rotation in PBS containing 0.1% bovine serum albumin for one hour. Cells were lysed in either RIPA or Triton lysis buffer (1 mM EDTA, 150 mM NaCl, 1% Triton X-100, protease inhibitor cocktail 1:500) supplemented with 0.1% bovine serum albumin, scraped, and collected in 1.5 ml microfuge tubes. The lysates were incubated on ice for 20 minutes and then cell debris was removed by centrifugation at full speed for 10 minutes at 4°C. The supernatants were precleared by rotation with 25 µl of protein G beads for 30 minutes, followed by centrifugation at full speed for 10 minutes at 4°C, and the supernatants were changed to new tubes containing 25 µl of protein G beads and a mixture of 2 µg monoclonal and 1 µg polyclonal anti-HA antibody. In general, 1 µg of

antibody was added to 0.5 mg – 1 mg of protein from a whole-cell extract. The slurry was incubated overnight with rotation. The following day, protein G beads were pelleted by centrifugation, and the supernatant was removed and replaced with 0.5 ml wash buffer (0.5 x RIPA or Triton buffer). The wash step was repeated four times in wash buffer and once in PBS. The resulting protein G beads were boiled in 30 µl of 2x Laemmli sample prep buffer (4x: 8% SDS, 250 mM Tris 6.8, 40% glycerol, bromophenol blue, 400 mM dithiothreitol) to solubilize bound protein. The supernatant was removed and separated by SDS-PAGE in a 7.5 or 10% polyacrylamide gel[131].

2.19 Pulse-chase metabolic labeling of protein

Transiently transfected BHK cells were starved in cysteine- (Cys) and methionine- (Met) free DMEM (Invitrogen) supplemented 1 in 100 with Glutamax (Invitrogen). Following starvation, nascent protein was labelled by the addition of 200 µCi/ml ^{35}S Cys/Met EasyTag Expre^{35}S (Perkin Elmer) for a pulse period of 10 minutes. The label was then removed, the cells were washed twice in prewarmed unlabeled chase media (DMEM supplemented with 5 mM each of Cys and Met), then incubated in chase media for the indicated times[132]. At the end of each chase period, cells were washed once with PBS, and lysed in RIPA buffer (see section 2.15). EMG1 was immunoprecipitated from the supernatant using an anti-HA antibody as described in section 2.18. Sample volumes for SDS-PAGE were normalized to protein concentration, and β-galactosidase activity was used to correct for transfection efficiency. Following gel electrophoresis to separate the proteins, the gel was incubated for 15 minutes in 10% acetic acid/ 30% methanol (fix solution), then incubated with shaking overnight in the same solution containing dilute Coomassie Blue to visualize protein. The following morning, the gel was washed in

several changes of fix solution to remove excess Coomassie Blue and rinsed three times with distilled water. The gel was incubated for 30 minutes in 1 M sodium salicylate/ 1% glycerol, and dried for two hours at 70°C under vacuum on a gel dryer. The dried gel was exposed to Kodak Biomax XAR or MS film overnight at -80°C between two intensifying screens (Kodak Biomax MS), and the film was developed to visualize bands.

2.20 Colorimetric β-galactosidase activity assay

β-galactosidase activity was determined by incubating the cell lysate with the substrate o-nitrophenyl-β-D-galactopyranoside (ONPG), producing o-nitrophenyl, a yellow product. Protein extracts were diluted with water to 150 µl total in 1.5 ml microcentrifuge tubes, 150 µl of assay 2x buffer (120 mM Na_2HPO_4, 50 mM NaH_2PO_4, 2 mM $MgCl_2$, 100 mM β-mercaptoethanol, 1.33 mg/ml ONPG) was added, and the tubes were mixed by vortexing. The samples were incubated at 37°C for 30 minutes, during which time a visible yellow colour was produced in the presence of β-galactosidase. The reaction was stopped by adding 500 µl of 1 M sodium carbonate, and the absorbance was read at 420 nm using a spectrophotometer.

2.21 Proteasome inhibition

The activity of the proteasome was inhibited by treating cells either with 10 µM of Z-Ile-Glu (OtBu)-Ala-Leucinal (Proteasome Inhibitor I, PI1), or 15 µM Cbz-Leu-Leu-Leucinal (MG-132) dissolved in dimethylsulfoxide (DMSO) (all Calbiochem) overnight. As a negative control, cells were treated with DMSO alone.

2.22 Assessment of protein synthesis rates

Cells were incubated for four hours at 37°C with media containing 150 μCi ^{35}S Cys/Met (Perkin Elmer), washed with ice-cold PBS, and lysed in 500 μl RIPA buffer supplemented with protease inhibitors (Sigma). The lysate was separated into supernatant and pellet by centrifugation at 13 200 rpm for 10 minutes, and 20 μl of supernatant from each sample was mixed with 100 μl of 40 mg/ml bovine serum albumin and 1 ml of ice-cold 10% trichloroacetic acid (TCA), then incubated on ice for 45 minutes to precipitate the protein. Each sample was then passed through a 1.2 μm glass microfibre filter (GF/C, Whatman) using a vacuum manifold, the filters were washed twice with 10% TCA and once with 100% ice-cold ethanol, and radioactivity of filter-bound protein was quantified in counts per minute (CPM) with a liquid scintillation counter (Beckman Coulter)[132]. The protein concentration was determined for each sample using the bicinchoninic acid assay (Pierce). Protein synthesis rate was then determined by the amount of incorporated radioactivity in CPM per μg of protein. Samples treated with cycloheximide (Sigma) to inhibit protein synthesis were used as a negative control.

2.23 Microscopy

Light and fluorescence microscopy were performed using a Zeiss AxioCam A1 equipped with a colour AxioCam MRc camera and 5x to 100x objectives. Photomicrographs were processed using AxioVision software.

2.24 Tissue processing and embedding

Mouse tissues and embryos were dehydrated and infiltrated with paraffin wax according to the following schedule: 70% ethanol for 1 hour, 2 changes of 95% ethanol for 1.5 hours each, 4 changes of 100% ethanol for 1.5 hours each, xylene for 30 minutes, two

changes of xylene for 1 hour each, paraffin for 2 hours, and paraffin for 4 hours (modified from the method of Prophet *et al*[133]). Very small tissues and embryos were first immobilized in Histogel (Thermo Scientific Richard-Allan Scientific) to prevent damage or loss. Processed tissues were embedded in paraffin blocks and cut into 5 µm sections using a Leica RM2245 microtome. Sections were floated on a warm water bath, placed on positively-charged glass slides (Leica), and left to dry at room temperature overnight.

2.25 Hematoxylin and eosin staining

Slides of formaldehyde-fixed, paraffin-embedded tissue were deparaffinized and rehydrated in several changes of xylene and ethanol as follows: two changes of xylene for five minutes each, two changes of 100% ethanol for five minutes each, two changes of 95% ethanol for three minutes each, and incubations of three minutes each in 70% ethanol, 50% ethanol, and distilled water. Slides were stained in Mayer's hematoxylin (Sigma) for three minutes, followed by washing in running tap water for 15 minutes. The slides were transferred to 80% ethanol for one minute, and stained in eosin Y (Sigma) for 30 seconds. The slides were then dehydrated using one, three minute incubation in 70% ethanol, two changes of 95% ethanol for three minutes each, and five minute incubations in two changes of 100% ethanol and two changes of xylene. Glass coverslips were affixed with Permount (Fisher Scientific). The mounting medium was allowed to cure overnight, and the slides were visualized using light microscopy.

2.26 Immunohistochemistry

2.26.1 Antigen retrieval

Formaldehyde fixed, paraffin-embedded tissues mounted on slides were deparaffinized and rehydrated as in section 2.25. Antigen retrieval was performed for 20 minutes in preheated 10 mM sodium citrate buffer (pH 6.0) heated to 95°C above a boiling water bath. The slides were removed from the steam and allowed to cool in the citrate buffer for 20 minutes at room temperature.

2.26.2 Blocking

To quench endogenous peroxidases, slides were washed twice in distilled water for five minutes each, and treated with 3% H_2O_2 for 10 minutes. The slides were washed in TBST for five minutes, a circle was drawn around the section using a hydrophobic slide marker (Research Products International), and sections were blocked in 3% BSA in TBST for 60 minutes at room temperature. Endogenous biotin was blocked using the Avidin-Biotin blocking kit (Vector Labs) following the manufacturer's instructions.

2.26.3 Antigen detection

Slides were washed in TBST for two minutes, and the sections were incubated with a primary antibody (see Table 2.3 for a list) in blocking solution overnight at 4°C in a humidified chamber. The following day, slides were washed in three changes of TBST for five minutes each, and incubated with biotinylated secondary antibody for one hour at room temperature. Meanwhile, the ABC complex containing horseradish peroxidase was prepared according to the manufacturer's instructions (Vector Labs). The slides were washed three times in TBST, and incubated in a drop of the prepared ABC complex for

one hour, followed by three washes of two minutes each in TBST. One drop of DAB substrate (Vector Labs) was added to each section, and incubated for ten minutes at room temperature to allow a brown precipitate to form. To counterstain nuclei, the slides were washed in distilled water for five minutes, and incubated in methyl green solution (0.5% methyl green, 0.1M sodium acetate pH 4.2) for 30 seconds to five minutes at room temperature. Slides were rinsed in water, then dehydrated and mounted as in section 2.25.

2.27 Immunocytochemistry

Adherent cells were grown to sub-confluency on sterilized glass coverslips, rinsed with PBS, and fixed in 4% paraformaldehyde for 15 minutes at room temperature. The cover slips were washed in PBS 3 times for 5 minutes each, then quenched using 50 mM ammonium chloride in PBS for 10 minutes, followed by three washes in PBS. The cell membrane was permeabilised with 0.1% Triton X-100 in PBS for 10 minutes, the coverslips were washed three times in PBS, and the cells were incubated in a blocking solution (3% bovine serum albumen, 0.05% Tween in PBS) for 60 minutes by placing 100 µl drops on parafilm and inverting coverslips on top. To prevent evaporation, coverslips were covered with culture plate lids and moist paper towels.

Primary antibodies were diluted in blocking solution and 100 µl was placed on parafilm with the coverslip inverted on top for one hour. Fluor-conjugated secondary antibodies (Alexa Fluor, Molecular Probes) were incubated on the coverslip for 30 minutes. The coverslips were rinsed twice with water, and incubated for 30 seconds in Hoechst 33342 diluted 1 µl in 100 ml of water to stain nuclei. The coverslips were washed twice in water and the excess liquid was removed by holding a tissue to a corner of the coverslip. Prolong Gold Antifade Reagent (Molecular Probes) was applied to a glass slide, and the

coverslip inverted on top. The Prolong Gold was allowed to cure overnight, and the slides were sealed the following morning by painting the edges of the coverslip with clear nail polish. The slides were imaged using fluorescence microscopy.

2.28 Detection of small and large subunit of the ribosome

To detect ribosomal subunits, a modified method was followed[134]. Cells were cultured in 150 mM plates and harvested at subconfluency. Adherent cells were washed with sterile PBS, scraped, and harvested into 15 ml centrifuge tubes. Suspension cells were washed by centrifugation and resuspended in PBS. Equal numbers of cells were pelleted by centrifugation, and ribosome extraction buffer (10 mM Tris pH 7.4, 1% IGEPAL, 0.5% sodium deoxycholate, 2 mM [low] or 15 mM [high] $MgCl_2$, 0.1M KCl, 0.5 mg/ml heparin, 80 U/ml RNAseOut [Invitrogen], 1 mM DTT, 1/250 protease inhibitor stock [Sigma]) was added to the pellets to lyse the cells. Lysates were incubated on ice for 15 min and then nuclei and cell debris were pelleted by centrifugation at 12 000 x g for 2 min. A portion of the lysate was saved as an input fraction. Sucrose prepared at a concentration of 10-45% in sucrose gradient buffer (Tris pH 7.4, 0.1 M KCl, $MgCl_2$ equal to ribosome extraction buffer, 2 mM EDTA [low Mg gradients only]) were prepared in polyallomer ultracentrifuge tubes using a gradient maker. Approximately 300 µl of the ribosome lysate was applied to the top of the gradient tubes and separated by centrifugation at 36 000 rpm for 3h at 4°C with the brake off. Gradient tubes were then pierced from the bottom with an ISCO gradient collector and heavier sucrose (60%) was pumped in to move the gradient solutions up and out of the tube. The solution was continuously monitored for absorbance at 254 nM and 0.5 ml fractions were collected. The tracing produced was then scanned, and the area under the curve was determined

using the Measure function in Axiovision software.Average 60S/40S ratios for wild type and BCS cells were compared using Student's *t*-test.

2.29 Pulse-chase metabolic labeling of rRNA

Ribosomal RNA pulse-chase experiments were performed following the method of Pestov *et al*[134]. For rRNA analyses, water and solutions were treated with 0.1% DEPC at room temperature overnight, followed by autoclaving. Equipment that would come into contact with the RNA was treated with 3% H_2O_2 for 10 minutes and rinsed in DEPC water, and benchtops were treated with Eliminase (Decon Labs, Inc) or RNase Killer (5 Prime). Glass and metalware were baked at 300°C for four hours, and RNase/DNase-free plasticware was used wherever possible.

For labelling, equal numbers of lymphoblasts were counted using a hemacytometer and aliquoted in RPMI media on the same day as the experiment, whereas fibroblasts were counted the day before the experiment and allowed to grow overnight in DMEM to subconfluence. Depending on the labelling method, cells were washed twice and starved in either phosphate-free or methionine-free media (Invitrogen) for 30 minutes to one hour. After removing the starvation media, media containing the label (either 10 µCi/ml $^{32}P_i$ (Perkin Elmer) or 50-75 µCi/ml L-[methyl-^{3}H] methionine (Perkin Elmer)) was added for a 20-30 minute pulse. Following the pulse, cells were washed twice in complete DMEM or DMEM supplemented with 5 mM methionine, and chased in the same media for the time periods indicated. RNA was isolated as described in section 2.6.

RNA samples were separated by agarose gel electrophoresis as described in section 2.7.2. To load equal amounts of label, incorporated radioactivity in 2 µl of RNA was quantified

in CPM using a liquid scintillation counter, and samples were diluted in DEPC water to ensure an equal number of counts in each sample.

For $^{32}P_i$ labelling, the gel was photographed under UV light in order to visualise the total RNA, then dried on a piece of filter paper. The dried gel was exposed to a phosphor storage screen (screen K, Kodak) or to MS film (Kodak) using intensifying screens (MS screen, Kodak) at -80°C.

For L-[methyl-^{3}H] methionine labelling, RNA was transferred from the gel to a positively-charged nylon membrane (N+, GE Healthcare) at 12V for 60 minutes in 1x BPTE using the Genie transfer apparatus (Idea Scientific). The total RNA was stained using methylene blue (0.02% methylene blue in 0.3 M sodium acetate pH 5.5) for approximately 3 minutes and photographed under white light. The membrane was destained in 0.2% SSC/ 1% SDS 3X 5 minutes and the membrane was air-dried on filter paper, and exposed to a phosphor storage screen (screen K/Tritium, Kodak) or to MS film using an intensifying screen (LE screen, Kodak) at -80°C.

To visualise the labelled RNA, the film was developed and imaged using the ChemiDoc system (Bio-Rad), or the phosphor storage screen was scanned using a phosphorimager (Personal Imager FX, Bio-Rad). The intensities of individual bands were analysed using either Quantity One or Image Lab software (Bio-Rad).

2.30 Cell cycle analysis

Cells were harvested and washed twice with wash buffer (0.1% fetal bovine serum in PBS) to derive a single cell suspension. Equal numbers of cells were suspended in 1 ml of buffer, and 3 ml of -20°C absolute ethanol were added dropwise with light vortexing

to fix cells while minimizing clumping. Cells were fixed a minimum of 3 hours at 4°C and stored at room temperature until the day of analysis. Fixed suspensions were washed twice with PBS, suspended in 1 ml of propidium iodide staining solution (40 µg/ml propidium iodide, 0.2 ml RNase A, 0.1% TritonX-100 in PBS), and incubated at room temperature for 30 minutes. Samples were placed on ice until analysis by flow cytometry, using a MoFlo XDP cell sorter (Beckman Coulter) and Modfit LT software for cell cycle analysis (Verity Software House).

2.31 Detection of LacZ reporter gene expression

To evaluate expression of the wild-type EMG1 gene in mice, a mouse expressing the LacZ gene under the control of the endogenous EMG1 promoter was employed[122]. β-galactosidase activity was detected following the method of Nagy *et al*[128]. Tissues isolated from heterozygous mice were pre-fixed in 0.4% formaldehyde (titrated from paraformaldehyde) and rinsed in detergent rinse (0.1 M phosphate buffer pH 7.3, 2 mM $MgCl_2$, 0.01% sodium deoxycholate, 0.02% Nonidet P-40). Tissues were stained in X-gal staining solution (0.1 M phosphate buffer, 2 mM $MgCl_2$, 0.01% sodium deoxycholate, 0.02% Nonidet P-40, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 1 mg/ml X-gal [5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside], 20 mM Tris pH 7.3) at room temperature overnight. Tissues were washed extensively in PBS and fixed overnight at 4°C in 4% formaldehyde titrated from paraformaldehyde. To prevent dissolution of the product during processing, tissues were processed using 1-butanol instead of xylene.

2.32 Assembly of EMG1 targeting construct

The strategy for cloning the *Emg1* targeting construct used three pre-existing vectors (a generous gift of Dr. Hao Ding): pSA-IRESβgeopA containing a splice acceptor, an internal ribosomal entry site, and a β-galactosidase-neomycin phosphotransferase fusion gene; pKS-2*loxP*Tie2, containing two *loxP* sites in the same orientation; and pGKDTApA, which contained the negative selection cassette encoding diphtheria toxin A.

Figure 2.1 summarizes the cloning strategy to generate the final targeting construct. As a first step, a region of mouse *Emg1* containing within it exons 2-6, was PCR-amplified from mouse *C129* genomic DNA. The BCS A>G mutation was introduced by amplifying the region as two fragments using complementary forward and reverse primers harbouring the mutation. The 386 bp 5′ fragment (amplified by WPGA and WPGB) and 1629 bp 3′ fragment (amplified by WPGC and WPGD) were used as templates to produce a 1994 bp PCR product using just the outermost primers (WPGA and WPGD). In order to facilitate the subsequent cloning step, the PCR product was cloned using a Perfectly Blunt cloning kit (Novagen) into a pST-Blue1 vector, creating pSTBlue1EMG1exons 2-6, which was amplified in *E. coli*. To verify that only the desired mutation was introduced during the PCR amplification, the coding sequence of the cloned fragment was verified by sequencing. To introduce exons 2-6 of Emg1 between two *loxP* sites, the entire region as well as some adjacent noncoding sequence, was excised from pSTBlue1 with BspHI (filled in) and NheI. pKS-2*loxP*-Tie2 was cut with HindIII (filled in) and then with NheI. The vector and insert were ligated to generate pKS-2*loxP*-EMG1exons 2-6.

To generate the 3′ homologous arm of the targeting construct, PCR primers WPG707 and WPG708 were used to amplify a 4794 bp region and introduce an AscI site at the 3′ end. During cloning, the 3′ arm was truncated and the AscI site was lost; therefore an alternate PmlI restriction site was used. The fragment was cloned into pSTBlue-1 as above, and one clone, pSTBlue1-3′EMG1, was identified and sequenced. The 2460 bp insert was released with SalI (filled in) and PmlI digestion, and inserted into XhoI (filled in) /AscI-digested pGKDTApA to generate pGKDTApA3′EMG1, containing the 3′ arm of homology and the negative selection marker diphtheria toxin A (DTA). To join exons 2-6 with the 3′ arm and negative selection marker, KS-2*loxP*-EMG1exons2-6 was opened with SalI (filled in) and NotI restriction digestion, and ligated with Nhe1 (filled in) /NotI-digested pGKDTApA3'EMG1 fragment to generate KS-2*loxP*EMG1exons2-6_3′EMG1-DTA.

To generate the 5′ homologous arm of the targeting construct, PCR primers WPG709 and WPG708 were used to amplify a 5505 bp region upstream of exon 2 of Emg1. This fragment was inserted into the pSTBlue-1 vector to form pSTBlue-5'EMG1, and the sequence was verified. The insert was released from clones using SalI and BspHI (filled in) restriction digestion, and ligated with XhoI/SbfI (filled in) -digested SA-IRESβgeopA to generate 5′EMG1SA-IRESβgeopA.

To assemble the final construct, the 5′EMG1SA-IRESβgeopA construct containing the 5′ arm of homology and SA-IRESβgeo was opened by SalI (filled in) and NotI

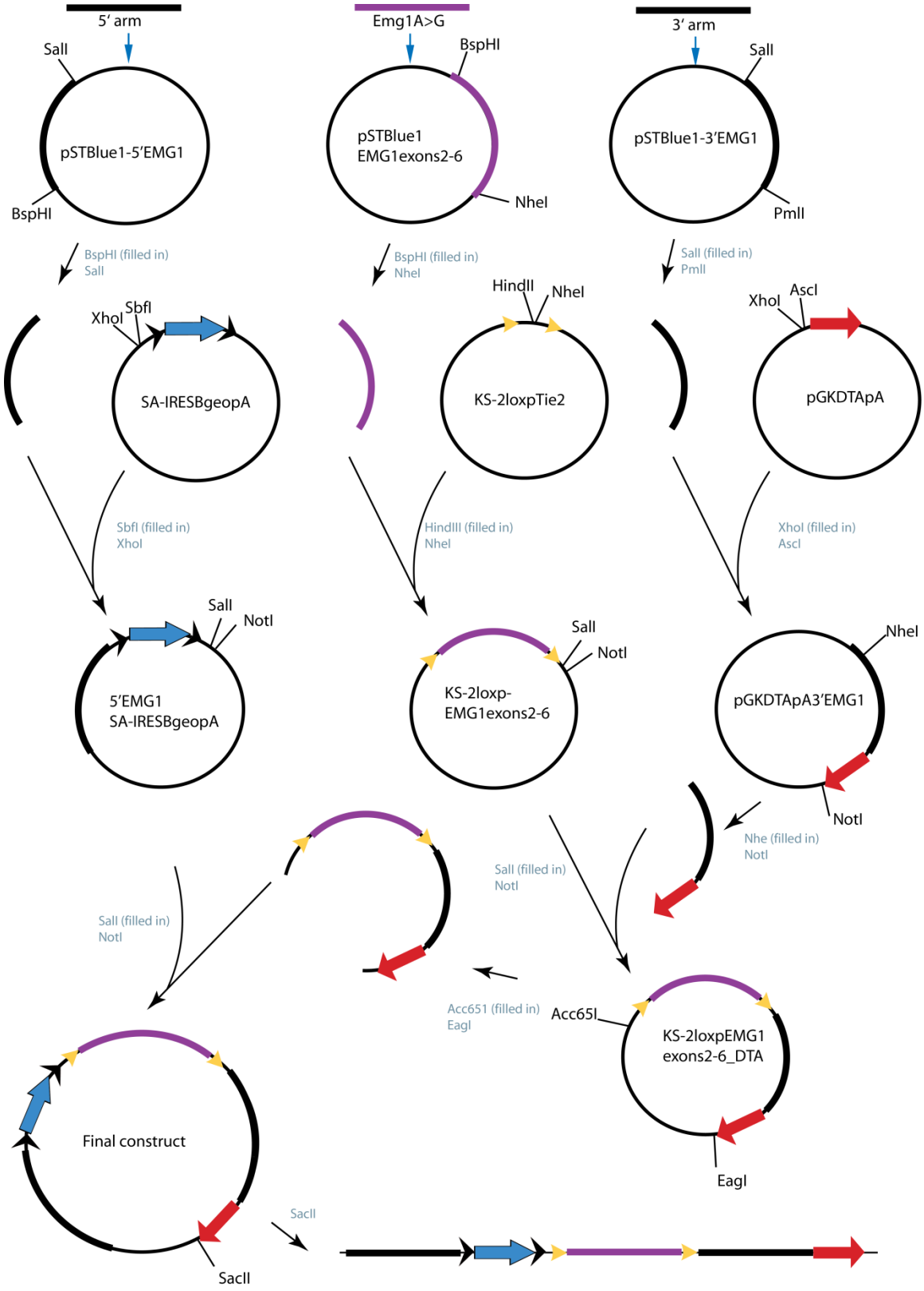

5' arm
SalI
pSTBlue1-5'EMG1
BspHI
Emg1A>G
BspHI
pSTBlue1
EMG1exons2-6
NheI
3' arm
SalI
pSTBlue1-3'EMG1
PmlI
BspHI (filled in)
SalI
BspHI (filled in)
NheI
SalI (filled in)
PmlI
XhoI
SbfI
SA-IRESBgeopA
HindIII
NheI
KS-2loxpTie2
XhoI
AscI
pGKDTApA
SbfI (filled in)
XhoI
HindIII (filled in)
NheI
XhoI (filled in)
AscI
SalI
NotI
5'EMG1
SA-IRESBgeopA
SalI
NotI
KS-2loxp-
EMG1exons2-6
NheI
pGKDTApA3'EMG1
NotI
Nhe (filled in)
NotI
SalI (filled in)
NotI
SalI (filled in)
NotI
Acc651 (filled in)
EagI
Acc65I
KS-2loxpEMG1
exons2-6_DTA
EagI
Final construct
SacII
SacII

Figure 2.1 Cloning and restriction enzyme strategy for the EMG1 targeting construct.

The 5’ and 3’ arms of homology (black bars) and the exons 2-6 of EMG1 including the A>G mutation (purple bar) were amplified and inserted into separate pSTBlue1 vectors, forming pSTBlue1-5’EMG1, pSTBlue1EMG1exons2-6, and pSTBlue1-3’EMG1. The plasmids were amplified in *E. coli*, the inserts were released by restriction digestion as indicated in grey, and gel purified. Fragments containing the 5’ arm, exons 2-6 and the 3’ arm were ligated into restriction-digested destination vectors containing the SA-IRES cassette (blue arrow) flanked by *FRT* sites (black arrowheads), the two *loxP* sites (yellow arrowheads), or the Diptheria toxin A cassette (red arrow). The fragment containing the 3’ arm and the DTA cassette was released by restriction enzyme digestion and ligated downstream of floxed exons 2-6. Finally, the fragment containing floxed exons 2-6, the 3’ arm, and the DTA cassette was released by restriction digestion and ligated downstream of the βgeo cassette flanked by *FRT* sites to form the final construct. The final construct was sequenced prior to linearization by SacII digestion and electroporation into mouse embryonic stem cells. Diagram is not to scale.

restriction digestion, then ligated with the Acc65I (filled in)/ EagI-digested KS-2*loxP*EMG1exons2-6_3′EMG1-DTA fragment. The correct orientation of the insert was verified by digestion with ZraI and SalI, and the final targeting construct was sequenced to verify that there were no coding sequence errors.

2.33 Integration of the vector into mouse ES cells

Electroporation of the final vector into mouse embryonic stem (ES) cells was performed by Dr. Hao Ding's laboratory following standard protocols[128]. The targeting vector was purified, linearized by digestion with SacII, and 20-40 μg was electroporated into R1 ES cells derived from *C129* mice[128] using a single pulse at 250V and 500 μF. Beginning 24 hours following electroporation, selection for neomycin resistance was performed by incubation in media containing 250 μg/ ml G418. Surviving colonies were picked in single-cell suspensions into 96-well plates with selective medium, and grown to 70-80% confluency. The cells were passaged into replica 96-well plates for DNA isolation and screening, while maintaining frozen stocks for later recovery of correctly-targeted colonies.

DNA was isolated and restriction-digested directly in the 96-well plate. Cells were washed in PBS, lysed in ES cell lysis buffer (10 mM Tris-HCl, pH 7.5; 10 mM EDTA; 10 mM NaCl; 0.5% sarcosyl; 1 mg/ml proteinase K), and incubated at 55°C in a humidified chamber overnight. To precipitate DNA, 100 μl of NaCl/ethanol (150 μl of 5M NaCl per 10 ml of cold 100% ethanol) was added to each well, and the plate was incubated at room temperature until the DNA was visible, approximately one hour. The

plate was inverted on a paper towel, and the wells were rinsed three times with 70% ethanol, allowing the ethanol to evaporate completely following the final wash. For restriction digestion, 30-40 µl of the restriction digest mix was added to each well, the plate was sealed and incubated overnight at 37°C in a humidified chamber. From 192 colonies, two independent correctly targeted clones were selected to generate chimeras, following screening for the presence of the 5′ arm, the 3′ arm, and the βgeo cassette by Southern blot, and for the 3′ *loxP* site and the A to G mutation by PCR (see Table 6.1). The selected ES cell lines, denoted 2B10 and 2C5, were thawed from replicate 96-well plates, expanded in selection medium up to 100 mm culture plates, and frozen in 3-5 cryovials. The cell lines were thawed, expanded, and re-screened to verify correct targeting prior to aggregation.

2.34 Aggregation of correctly targeted ES cells

Aggregations were performed by Dr. Hao Ding's laboratory following standard protocols[128]. One to two days before aggregation, subconfluent ES cells from the 2B10 and 2C5 clones were passaged onto gelatin-coated plates at a high dilution in order to produce colonies of 8-15 cells. On the day of the aggregation, the small colonies of ES cells were detached by gentle trypsinization. Wild-type morulae (embryonic day 2.5) from *CD1* mice were prepared for aggregation by removing the zona pellucida using acidified Tyrode's solution. ES cells and morulae were aggregated under mineral oil in drops of KSOM supplemented with amino acids, and cultured for 24h to produce early-stage blastocysts. Pseudopregnant *CD1* females were prepared by breeding with vasectomized males, and blastocysts were transferred into the uterus. Chimeric pups were identified by their reticulated coat, with the white coat colour originating from the *CD1*

morulae, and brown coat colour from the targeted *C129* ES cells. Chimeric males exhibiting a predominantly brown coat were crossed with wild type *CD1* females to test for transmission of the mutated *Emg1* gene to the germ line. Heterozygous males were crossed with *C57Bl/6* females expressing a thermostable Flp recombinase (Flpe) from the ROSA 26 locus (Gt(ROSA)26Sor$^{tm1(FLP1)Dym}$,Jackson Labs) to remove the βgeo cassette[135]. Pups were screened by PCR and Southern blot for the removal of the selection cassette, and heterozygous mutants without the βgeo cassette were crossed to generate homozygous mutants.

2.35 Phenotypic analysis of homozygous mutants

To determine the stage at which death was occurring in homozygous mutants, heterozygous female mice were crossed with heterozygous males and monitored for vaginal plugs. The morning a plug was discovered was considered embryonic day (E) 0.5. Timed pregnant females were sacrificed by cervical dislocation at E6.5-15.5, and the uterus was removed and placed in ice-cold PBS. The uterus was examined under a dissecting microscope for implantation sites, and embryos were removed and fixed in 4% formaldehyde (titrated from paraformaldehyde) in PBS for one hour to overnight depending on their size. To genotype embryos, DNA was isolated from the yolk sac or a portion of the tail. Alternatively, embryos were genotyped using DNA extracted from formaldehyde-fixed, paraffin embedded sections scraped off the slide using a scalpel blade. BCS embryos were compared with heterozygous littermates except where stated.

2.35 Bromodeoxyuridine labelling

To label cells in S phase of the cell cycle, timed pregnant females at E6.5 - E9.5 were injected intraperitoneally with 50 mg/kg bromodeoxyuridine (Roche) in PBS following

the manufacturer's instructions, and sacrificed two hours post-injection. Formaldehyde-fixed, paraffin-embedded sections from labelled embryos were treated with 2N HCl for 30 minutes at 37°C to denature DNA, and neutralized in 0.1 M borate buffer pH 8.5. Incorporated bromodeoxyuridine was detected using an anti-BrdU antibody (Roche) with the ABC/DAB system as in section 2.26.

2.37 Statistical analyses

Statistical significance was assessed using Student's t-test or Chi-square (χ^2) analysis calculated by GraphPad Prism software. A *p* value of less than 0.05 was considered statistically significant.

Chapter 3 : Identification of the mutation causing BCS

Acknowledgements: RT-PCR was performed by Shaungbo Liu. Northern blots were performed by Edward Nylen. Human EMG1 was modeled by Dr. Brian Mark. Nehal Patel performed semi-quantitative PCR analysis. Jill Wiechert performed PCR genotyping assays. DNA samples were collected by Dr. Cheryl Rockman-Greenberg and archival brain tissue was provided by Dr. Marc Del Bigio. Dr. Sunita Khatkar and Edward Nylen sequenced candidate genes and tested segregation of variants.

The majority of this work was published as: Armistead J, Khatkar S, Meyer B, Mark BL, Patel N, Coghlan G, Lamont RE, Liu S, Wiechert J, Cattini PA, Koetter P, Wrogemann K, Greenberg CR, Entian KD, Zelinski T, Triggs-Raine B. Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome. The American Journal of Human Genetics 2009; 84 (6):728-39

3.1 Introduction

The autosomal recessive disorder Bowen-Conradi syndrome is a significant problem in the Hutterite population, but studies of the affected patients were limited and afforded no clues as to the etiology of the disease. Where they were performed, karyotypic, hematological, and biochemical tests such as urinary and plasma amino acids were normal in affected infants[2,3]. Of the seven autopsies reported by Lowry *et al* in 2003, no consistent abnormalities were found although five showed very small brains. Of the two reported cases which included computerized axial tomogram results, one patient had a Dandy-Walker malformation and another had hypoplasia of the structures of the posterior fossa, suggesting cerebellum involvement[2]. Nevertheless, the progression of the disease was similar: affected infants had extreme growth failure, difficulty feeding requiring gavage, frequent bouts of pneumonia, and most died within the first two years of age[2,3].

The current Hutterite population in North America is likely descended from fewer than 100 ancestors who emigrated in the 1870s[11,12]. The endogamous nature of the population pointed towards the likelihood that BCS is due to a single mutation inherited from a common ancestor[2]. Thus, homozygosity by descent was used as the basis for finding the responsible gene. This approach posits that the mutation arose once early on in the population, and that because the population is derived from a small group of founders, the mutation will be the same in all affected individuals; further, the flanking haplotype will also be homozygous[136,137]. This technique has been used previously to identify the genes responsible for a number of other diseases in the Hutterite population (Reviewed in[13]).

DNA was collected from 42 Hutterite individuals, including 18 parents of BCS patients, 8 BCS patients from 6 sibships, 15 unaffected siblings and one grandparent. Following a genome-wide scan using 389 microsatellite markers and linkage analysis, linkage was detected for regions in chromosomes 12 and 16. Subsequently, 7 additional DNA samples from two families with one BCS patient each were incorporated into the study, as well as additional microsatellite markers and informative single nucleotide polymorphisms (SNPs). The disease-causing locus was thus narrowed down to a region spanning 1.9 Megabases (Mbp) on chromosome 12p13.3 between markers *F8VWF* and *D12S397*[16]. This region contained 59 known or predicted genes at the time of assessment. Due to the neurological features of BCS such as microcephaly and severe developmental delay, we hypothesized that the disease-causing gene would have a fundamental function in the brain. All of the genes in the candidate interval were therefore assessed for their potential to cause a developmental disease of the central nervous system.

3.2 Ranking and sequencing of candidate genes

3.2.1 Strategy

The search for the mutation causing BCS was based on a previous study[16] which mapped the disease-causing locus to a 1.9 Mbp region on 12p13.3. At the time of analysis, the candidate interval bound by *F8VWF* and *D12S397* contained 59 known and predicted genes (Table 3.1). As a preliminary step to determining the gene responsible, mRNA levels of 48 of these genes were assessed using cDNA reverse-transcribed from total RNA from patient and control lymphoblasts (data not shown). Sizes and levels of mRNA

were similar between patient and control cells, and consequently mRNA assessment was uninformative for evaluating candidate genes. The 59 genes were therefore ranked to determine which would be the best candidates for sequencing, based on available information regarding the gene function, expression profile, and level of sequence conservation. An ideal candidate was not already associated with a disease, had an essential function, and was expressed in the brain, considering the neurological features of BCS. A gene without a known function was automatically ranked as a good candidate so that it would be included in sequencing.

3.2.2 Identification of BCS-causing mutation in EMG1

Following ranking, each of the 20 good candidates, 13 medium candidates, and two of the poor candidates were sequenced. The exons and exon-intron junctions from the DNA of a single patient were sequenced, using NCBI reference sequences NT_009759.15 or NT_009714.16 (www.ncbi.nlm.nih.gov) for comparison. From the sequenced genes, 59 known and six novel SNPs were identified (see Table 3.1).

GENE	MIM	Rank	RT/ PCR	Sequence	SNPs Detected	
					Known	Novel[a]
CD9	*143030	Poor	Yes	No		
PLEKHG6	*611743	Medium	Yes	Yes	rs740842[b]; rs1468603; rs4764497; rs3217075	
TNFRSF1A	*191190	Poor	Yes	No		
SCNN1A	*600228	Poor	Yes	No		

LTBR	*600979	Poor	Yes	No		
CD27	*186711	Poor	Yes	No		
TAPBPL	*607081	Poor	Yes	No		
VAMP1	*185880	Good	Yes	Yes	rs2532489; rs2534720; rs2534718; rs2072375; rs1017101; rs7390	
MRPL51	*611855	Medium	Yes	Yes	ND	
NCAPD2		Good	Yes	Yes		g.6558820G>T rs78834955
GAPDH	*138400	Medium	Yes	Yes	ND	
IFFO1	*610495	Medium	NT	Yes		g.6597090G>A rs76915198
NOP2	*164031	Good	Yes	Yes	ND	
CHD4	*603277	Good	Yes	Yes	rs34863790	
LPAR5	*606926	Medium	Yes	Yes	rs3741924	
ACRBP	*608352	Poor	Yes	No		
ING4	*608524	Good	Yes	Yes	rs4764506	
ZNF384	*609951	Good	Yes	Yes	ND	
C12orf53		Good	Yes	Yes	ND	
COPS7A		Medium	Yes	Yes	rs2286731; rs3168600	g.6780121C>T rs77365023
MLF2	*601401	Poor	Yes	No		
PTMS		Good	Yes	Yes	ND	
LAG3	*153337	Poor	Yes	No		
CD4	*186940	Poor	Yes	No		
GPR162		Good	NT	Yes	rs1051409; rs2071082	
LEPREL2	*610342	Good	NT	Yes	rs11064423; rs4963512; rs4963518;	g.6877245_6877246insCA rs75267272; g.6877532C>T rs73055793

					rs4963510; rs1047776; rs57050687[c]	
GNB3	*139130	Medium	Yes	Yes	ND	
CDCA3	*607749	Medium	Yes	Yes		
USP5	*601447	Good	Yes	Yes	ND	
TPI1	+190450	Poor	Yes	No		
SPSB2	*611658	Good	Yes	Yes	ND	
LRRC23		Good	Yes	Yes	ND	
ENO2	*131360	Medium	Yes	Yes	ND	
ATN1	*607462	Poor	Yes	Yes	rs2071075; rs33940553	
C12orf57		Good	Yes	Yes	rs915997; rs2019743	
PTPN6	*176883	Medium	Yes	Yes	rs2301262; rs10744724; rs7963446; rs2110071; rs2110072; rs759052; rs759053	
PHB2	*610704	Medium	Yes	Yes	rs2159887	
EMG1	*611531	Good	Yes	Yes	rs17857448[d] ; rs11428482	g.7023589A>G rs74435397
LPCAT3	*611950	Good	Yes	Yes	ND	
C1S	*120580	Poor	Yes	No		
C1R	%216950	Poor	Yes	No	ND	
LOC727728		Poor	NA	No		
C1RL	*608974	Poor	Yes	No		
LOC283314		Poor	NA	No		
RBP5	*611866	Poor	Yes	No		

CLSTN3	*611324	Medium	Yes	Yes	rs2278822; rs3816423	
PEX5	*600414	Poor	Yes	No		
ACSM4		Poor	NA	No	rs7962994	
CD163L1		Poor	Yes	No		
LOC727815		Poor	NA	No		
CD163	*605545	Poor	Yes	No		
LOC643739		Poor	NA	No		
APOBEC1	*600130	Poor	NT	Yes	rs34275479[e]	
GDF3	*606522	Good	NT	Yes	ND	
DPPA3	*608408	Medium	Yes	Yes	rs2024320[f]; rs34619457	
CLEC4C	*606677	Poor	Yes	No		
NANOG	*607937	Good	Yes	Yes	rs2889552[g]; rs4354764; rs4242903; rs4294629; rs1860931; rs4012939; rs4012937; rs11831829; rs58791167	
SLC2A14	*611039	Good	NT	Yes	rs3919534; rs2215713; rs4402376	
SLC2A3	*138170	Good	Yes	Yes	rs2541281; rs2541279; rs1043712; rs3931701	

[a]Numbering is based on NT_009759.16

[b]NM_018173.2:c.103G>A, p.A35T

[c]NM_014262.3:c.419_421insG

[d]NM_006331.6:c.126G>T, p.R42S

[e]NM_001644.3:c.62G>A, p.W21X

[f]NM_199286.2:c.151G>C, p.E51Q

[g]NM_024865.2:c.47C>A, p.A16E

Table 3.1 List of genes in the minimal interval, in telomeric to centromeric order (NCBI build 36.2). The prefix of a MIM number (Mendelian Inheritance in Man, www.omim.org) indicates the type of entry: * describes a gene with known sequence, + describes a gene with known sequence and its associated phenotype, and % describes a Mendelian phenotype or locus for which the molecular basis is unknown. NT, no transcript detected; NA, not analyzed; ND, none detected. Superscripts b-g represent the six known SNPs causing nonsynonymous changes.

Each SNP was evaluated for its likelihood to cause disease, based on its location in the gene and information available from the SNP database (www.ncbi.nlm.nih.gov/SNP). Of the 59 known SNPs, 53 were located within introns or caused synonymous changes, and were not evaluated further. The remaining 6 known SNPs caused nonsynonymous changes, however they were already present in the SNP database, were frequently found outside the Hutterite population (*PLEKHG6* c.103G>A, *DPPA3* c.151G>C), were located in nonconserved regions of the corresponding protein (*EMG1* c.126G>T, *NANOG* c.47C>A), were likely errors in the reference sequence (*LEPREL2* g.6878024_6878025insG), or were considered unlikely to cause BCS based on established mouse models[138] (*APOBEC1* c.62G>A). The six known SNPs causing nonsynonymous changes are indicated in Table 3.1. Of the six novel SNPs that were identified, five were located within introns. The remaining SNP was c.257A>G, p.D86G in *EMG1* (Essential for Mitotic Growth 1) (Figure 3.1 A). To screen for the mutation in

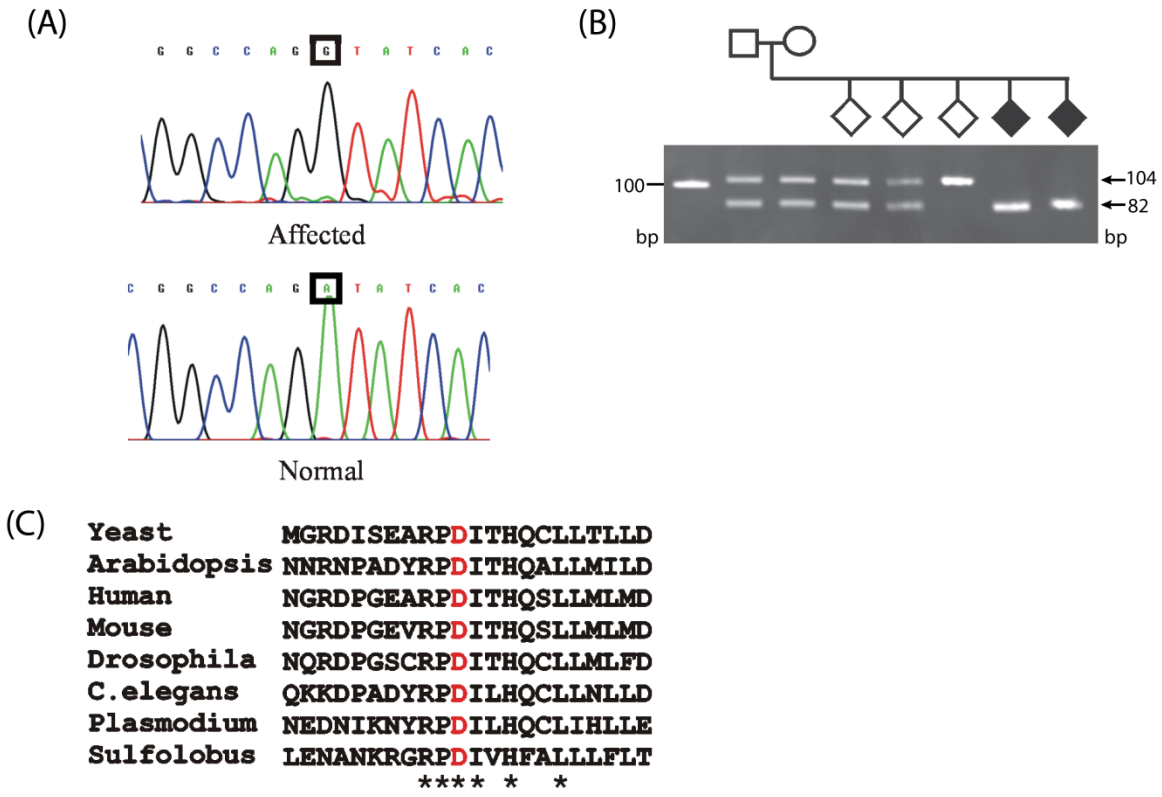


Figure 3.1 The c.257 A>G, p.D86G mutation in *EMG1*.

(A) Sequence chromatograms of a BCS-affected patient (top) and a normal control (bottom). The position of the affected nucleotide is indicated by a box, and the top chromatogram shows the A>G mutation. The coding sequence of *EMG1* was PCR-amplified as four fragments using primers specific for exon 1 (WPG 800-801) exons 2 and 3 (WPG 802-803), exons 4-5 (WPG 804-805) and exon 6 (WPG E and WPG F) (see Table 2.1). Sequencing was performed using the same primers. **(B)** Detection of the c.257A>G mutation in a Hutterite family. The region of *EMG1* containing the c.257A>G mutation was amplified by PCR from the DNA samples of a family with BCS-affected children using primers WPG 838 and WPG 839 (Table 2.1). The reverse primer created a KpnI site only in the presence of the mutation. Following KpnI digestion at 37°C for two hours, the 82 and 22 bp products generated in the presence of the mutation were separated from the the undigested 104 bp product by gel electrophoresis in an 8% polyacrylamide gel. The affected children (black diamonds) have only the 82 base pair fragment, whereas the parents and two of the siblings are heterozygous for the mutation, as indicated by the presence of both 104 bp and 82 bp fragments, and one child is homozygous for normal EMG1, as indicated by the presence of only the 104 bp fragment. The 22 bp product is not visible on this gel. **(C)** Protein sequence alignment via Clustal W of the region of the EMG1 protein containing the D86G substitution. The residues that are completely conserved in all orthologs are indicated with an asterisk, and the Asp (D) that is mutated in BCS is shown in red.

EMG1, a PCR-based assay was established in which a KpnI site was created only in the presence of the mutation. KpnI digestion of the PCR product resulted in an 82 and a 22 bp band in the presence of the mutation (Figure 3.1 B). The c.257A>G mutation segregated completely with disease and was not found in any of 414 non-Hutterite alleles.

3.3 EMG1 protein structure

3.3.1 Conservation of EMG1 sequence

To assess the potential of the c.257A>G, p.D86G mutation in *EMG1* to cause disease, orthologs were found in the NCBI database using BLAST[126] (blast.ncbi.nlm.nih.gov). A multiple protein sequence alignment created using Clustal W[127] (www.ebi.ac.uk/Tools/msa/clustalw2) showed not only that the region of EMG1 surrounding the mutated residue was highly conserved (asterisks in Figure 3.1 C), but also that the aspartate residue mutated in BCS was completely conserved (highlighted in red in Figure 3.1 C). Conservation even in Archaea (represented by *Sulfolobus*) was a good indication that the D86 residue was essential for EMG1 protein function and that an amino acid substitution at this location could be detrimental to its function.

3.3.2 Model of EMG1 homodimer structure

To understand the potential impact of the D86G substitution on protein structure and function, a model of the normal human EMG1 protein was created. The crystal structures of EMG1 homologues in *Saccharomyces cerevisiae* and the *Methanocaldococcus jannaschii* were already available[117,119] and served as templates for modeling the human

protein. The human EMG1 amino acid sequence is 52% identical and 65% similar to yeast Emg1 according to BLAST, and 29% identical and 49% similar to *M. jannaschii* Emg1. Moreover, the structural folds of the protein are very highly conserved and allowed for the construction of a human model. The model showed that EMG1, as a SPOUT class methyltransferase, was a homodimer[117,119] with an extended positively charged surface which putatively bound nucleic acids (Figure 3.2). A pocket in each subunit of the homodimer bound SAM, the methyl donor. The model showed that D86, which is substituted by glycine in BCS, formed two hydrogen bonds with R84 and stabilized it. R84 was found at the N terminus of α-helix 2 in the positively-charged RNA-binding groove of EMG1. The bundles formed by α-helices 2 and 7 of each monomer were positioned perpendicular to each other at the interface of the two subunits of the homodimer.

3.3.3 Prediction of impact of D86G on EMG1 structure

The model of human EMG1 presented multiple ways by which D86G could cause a disruption of EMG1 function. First, the yeast equivalent of R84 has been implicated in the RNA-binding capacity of Emg1[117], and the D86G substitution could disrupt the rigidly held conformation of R84, reducing RNA binding. Second, the introduction of a glycine residue at the N terminus of α-helix 2 could result in local misfolding in the protein. Since α2 and α7 participate in the formation of the dimer interface, misfolding of α2 could deform the dimerization interface and possibly expose core hydrophobic residues, leading to aggregation. All SPOUT-class methyltransferases studied to date are found in dimer form, and dimerization is crucial for protein function[118,139]. In this family

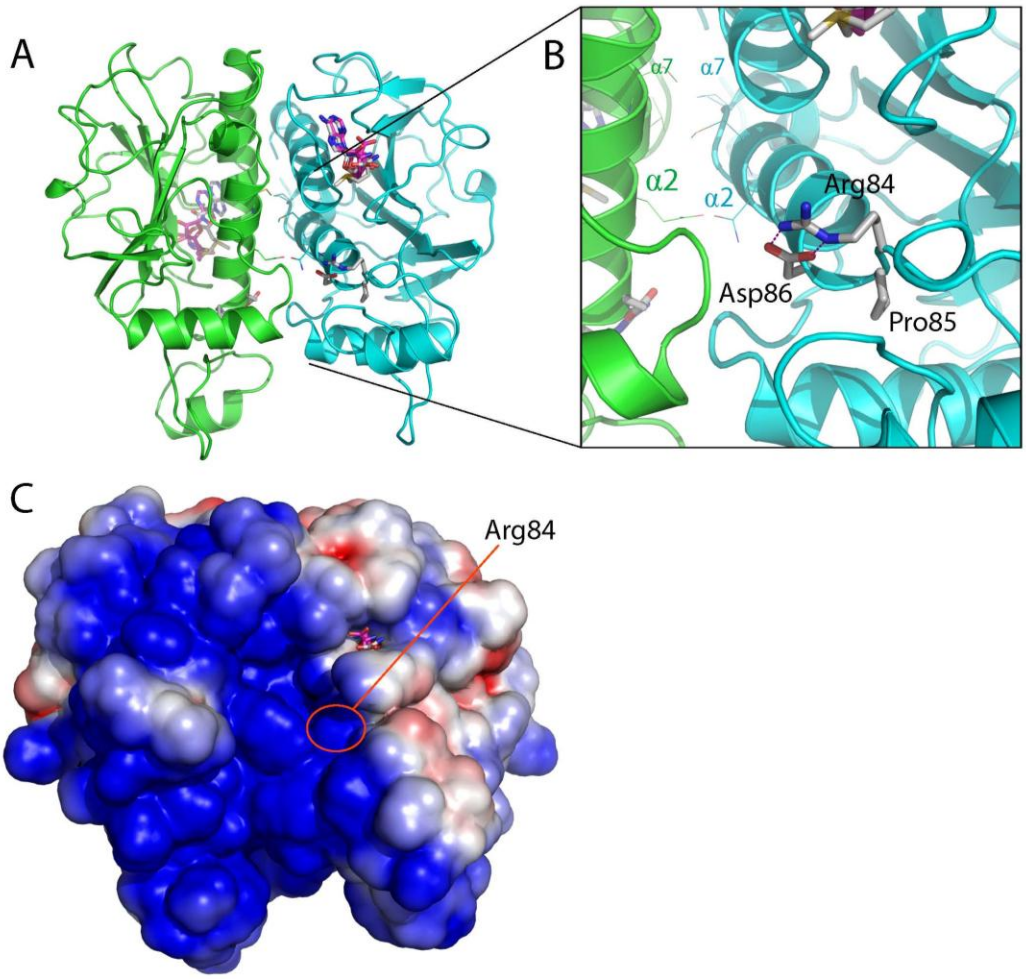


Figure 3.2 EMG1 protein model.

(A) Ribbon diagram of the modeled homodimer, with the individual monomers colored green and blue. Methyl donors are drawn as sticks. **(B)** Close-up view, showing the hydrogen bonding interactions of D86 with R84 occurring at the N terminus of α helix 2. The two parallel α helices from each monomer (helices 2 and 7) that form the core of the dimer interface are labeled α2 and α7 for each monomer. **(C)** Electrostatic potential map of the solvent-accessible surface of the human EMG1 homodimer model. The map shows that the model predicts a positively charged region (blue) that is homologous to the RNA-binding groove of the yeast EMG1 homologue Nep1, and R84 is located in the centre of this groove. Red indicates a negative charge.

of enzymes, the active site is formed by residues from both monomers, and in fact the catalytic residue is provided by the opposite subunit. Additionally, the loop that binds the methyl donor SAM is stabilized by interactions with the opposing monomer, and the positively-charged RNA-binding groove is also formed upon dimerization. Thus, the D86G substitution could reduce both RNA and SAM binding, abrogate catalytic activity, and could likely reduce protein solubility and stability.

3.4 *EMG1* expression in humans

3.4.1 Expression of EMG1 message in multiple adult and fetal tissues

To better understand the role of EMG1 in human tissues, we compared the expression of *EMG1* in a range of fetal and adult tissues by semiquantitative PCR analysis (Figure 3.3). Using the Human Multiple Tissue cDNA panel (Clontech Laboratories), cDNA from various tissues was PCR amplified using *EMG1*-specific primers. Each panel is pre-normalized to the expression levels of four housekeeping genes (glyceraldehyde 3-phosphate dehydrogenase [GAPDH], β-actin, α-tubulin and phospholipase A2). Generally, *EMG1* expression was higher in adult tissues than in the fetus. However, the highest level of expression was in fetal spleen which could suggest a role in hematopoiesis. Several other ribosome biogenesis genes associated with disease have been implicated in hematopoiesis, as evidenced in Diamond-Blackfan anemia, 5q-syndrome, Shwachman-Diamond syndrome, cartilage-hair hypoplasia, and dyskeratosis congenita (Reviewed in[90,140,141]). Interestingly, considering BCS primarily affects the

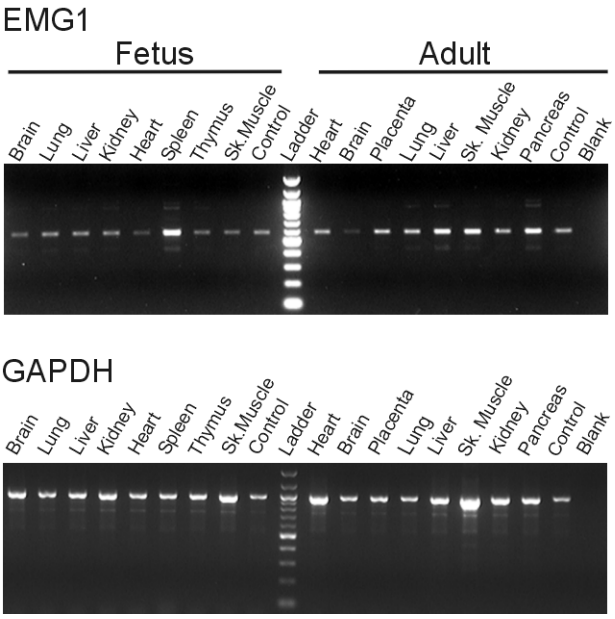


Figure 3.3 *EMG1* expression in human fetal and adult tissues.

Five nanograms of cDNA from various tissues was amplified by PCR with the use of EMG1- (top panel) or GAPDH-specific primers (bottom panel). Samples were taken at different cycle numbers, separated on a 2% agarose gel, and stained with ethidium bromide. The results for the samples taken at 26 cycles are shown for both EMG1 and GAPDH. Control cDNA was provided with each panel, and the PCR was performed in the absence of cDNA for the blank. Each panel was pre-normalized by the manufacturer to the expression levels of four housekeeping genes: GAPDH, beta-actin, alpha-tubulin, and phospholipase A2.

central nervous system during development, expression was higher in fetal brain than in the adult. This may indicate an increased need for EMG1 during brain development.

3.4.2 Expression of EMG1 in normal brain

To examine the possible role of EMG1 in brain development, immunohistochemistry was employed to visualize EMG1 in normal human brain sections of 15 weeks gestation, newborn, and adult (55 years old) (Figure 3.4). All brain tissue was taken upon autopsy from patients who died from causes unrelated to brain disease or trauma, and were presumed to have healthy brain tissue. Expression was strong and broad at the earliest developmental time point examined, although EMG1 was detected most strongly in the ventricular zone and the cortical plate. In the newborn brain, expression was still strong, although it was somewhat more restricted. The highest levels of expression were found adjacent to the ventricles, where EMG1 appeared to be cytoplasmic as well as nuclear, and in the Purkinje cell layer of the cerebellum. In the 55 year-old adult brain, expression was strongest in nuclei in the dentate gyrus of the hippocampus, and in the granular layer of the cerebellum. Strong staining was observed in some pyramidal neurons, indicating that EMG1 is also found in post-mitotic cells. Taken together with the cDNA panel data, this pattern of expression strongly suggests an essential function for EMG1 during brain development, and a more restricted role postnatally and in adulthood.

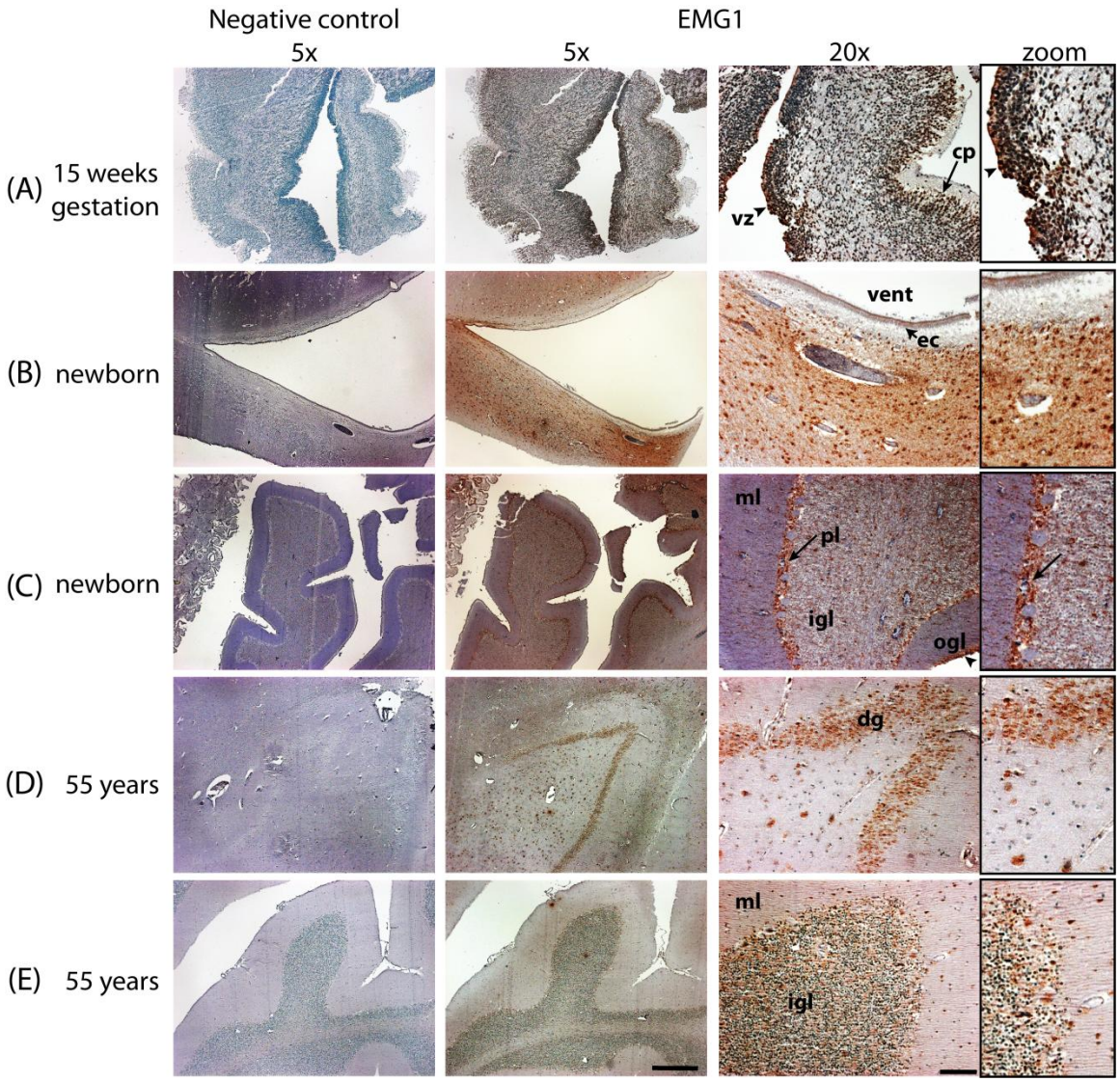


Figure 3.4 EMG1 expression in the human brain.

EMG1 was detected in archival, formalin-fixed, paraffin-embedded brain sections using an anti-EMG1 antibody (brown precipitate), and counterstained with methyl green. At 15 weeks gestation **(A)**, EMG1 expression was very strong in the nuclei of cells in the ventricular zone (arrowhead) and the cortical plate (arrow), with reduced expression in the intermediate layer. In the newborn, EMG1 expression was found in the nuclei and cytoplasm of most areas of the brain, but most strongly in the areas surrounding ventricles **(B)** and in the cerebellum **(C)** within the Purkinje cell layer (arrow) and the outer granular cell layer (arrowhead). In the adult, EMG1 was strong in the dentate gyrus of the hippocampus **(D)**, and was found in many neurons, including in the granular cell layer of the cerebellum **(E)**. Scale bar at 5x = 500 µm, 20x = 100 µm. **cp**, cortical plate; **dg**, dentate gyrus; **ec**, ependymal cells; **igl**, inner granular layer; **ml**, molecular layer; **pl**, Purkinje layer; **ogl**, outer granular layer; **vent**, ventricle; **vz**, ventricular zone.

3.4.3 EMG1 expression in control and BCS patient cells

3.4.3.1 mRNA expression

As a first step in understanding the effects of the c.257A>G, p.D86G mutation, northern blot analysis was used to examine the expression levels of *EMG1* mRNA from BCS-affected and normal control fibroblasts. Using an *EMG1*-specific probe, a band was detected at the expected size of approximately 1000 bp. Comparison with *GAPDH* as a loading control showed no obvious differences between patient and control mRNA levels (Figure 3.5 A), indicating that the mutation in *EMG1* has no effect at the mRNA level.

3.4.3.2 Protein expression

Since there did not appear to be a defect at the mRNA level, endogenous EMG1 protein levels in BCS patient and normal control fibroblasts were examined using a polyclonal anti-EMG1 antibody (Aviva). To validate the specificity of the antibody for immunoblotting, an antibody competition assay was performed. Using uncompeted antibody, bands were observed at the apparent molecular weight of 28 kDa for HA-tagged EMG1, and at 27 kDa for endogenous wild type EMG1 (Figure 3.5 B, top left panel). In parallel, an EMG1 antibody was incubated with purified, MBP-tagged EMG1 prior to immunoblotting. No bands at the predicted molecular weights were visible using this competed antibody (top right panel). The same blots were then stripped and re-probed with uncompeted EMG1 antibody, and bands were revealed at the predicted molecular weights (Figure 3.5 B, bottom panels). This indicated that the polyclonal antibody from Aviva was indeed specific for EMG1.

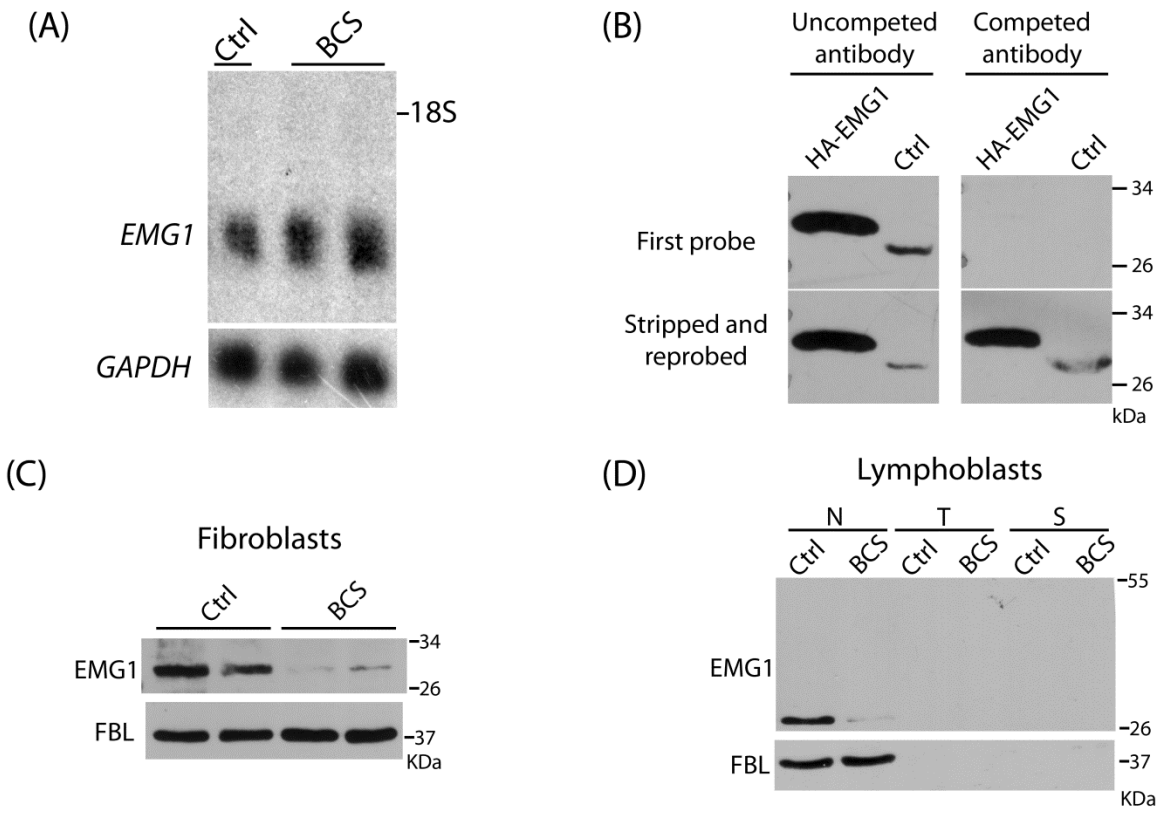


Figure 3.5 EMG1 in unaffected control and BCS patient cells.

(A) RNA analysis. Seven micrograms of total RNA isolated from unaffected control (Ctrl) or BCS patient fibroblasts (BCS) were separated on a 1.2% agarose gel and transferred to a nylon membrane. The membrane was UV cross-linked and probed with labelled EMG1 (top panel) and GAPDH (bottom panel) cDNA probes. Film was exposed to the membrane for 7 hours at -80°C (EMG1) or for 2 hours at room temperature (GAPDH). EMG1 is 1068 bp and GAPDH is 1310 bp; 18S rRNA position is marked as a size reference. **(B)** Validation of EMG1 antibody by antigen competition. Nuclear lysates from BHK cells overexpressing HA-tagged EMG1 (HA-EMG1), and from wild type lymphoblasts (Ctrl), were separated on a 10% gel by SDS-PAGE. Proteins were then transferred to a nitrocellulose membrane and probed with either an untreated anti-EMG1 antibody (top left panel) or an antibody that had previously been incubated with purified, MBP-conjugated EMG1 protein (top right panel). The membranes were then stripped and re-probed with untreated anti-EMG1 antibody (bottom panels). Antibody competition with EMG1 protein abolished the putative EMG1 bands (top right panel), while reprobing with uncompeted antibody revealed bands at the expected molecular weight (bottom right panel), validating its specificity for EMG1. **(C)** Immunoblot analysis of EMG1 protein levels. Twenty-five micrograms of nuclear lysates from unaffected control (Ctrl) and BCS patient (BCS) fibroblasts were separated on a 10% gel by SDS-PAGE, and protein was detected by immunoblot with an EMG1 antibody. The blot was then stripped and reprobed for fibrillarin (FBL), a nuclear protein, so that equal loading was ensured. **(D)** Nuclear (N), total cell (T) or post-nuclear supernatant (S) fractions were isolated from unaffected control and BCS patient lymphoblasts by centrifugation, and probed for EMG1 or fibrillarin as above.

EMG1 was then detected in control and BCS patient fibroblasts and lymphoblasts. Since it was difficult to detect EMG1 in whole-cell lysates, nuclear isolates were used in order to enrich the protein. Fibrillarin, a nucleolar protein, was used as a loading control and was not detected in whole-cell lysates or post-nuclear supernatant fractions (Figure 3.5 D, bottom panel). Unlike mRNA levels, EMG1 protein levels were dramatically reduced in both BCS patient fibroblasts (Figure 3.5 C) and lymphoblasts (Figure 3.5 D). Importantly, patients did not have a complete EMG1 deficiency as some protein was still detectable. In agreement with our model of the mutated EMG1 protein (section 3.3.3), a drastic reduction in EMG1 protein levels suggests that the D86G substitution causes instability of the protein, resulting in EMG1 degradation or aggregation in patient cells. The remaining protein in BCS patient cells likely retains its normal function, as a complete deficiency is lethal in both yeast[112,113] and mouse cells[122].

3.5 Discussion

Bowen-Conradi syndrome is a disease that is devastating for Hutterite families, but biochemical tests on affected children offered no clues to the molecular cause of the disease. The history of the Hutterite population led to the hypothesis that the mutation causing BCS arose only once in the population, resulting in a genome-wide scan approach and linkage analysis assuming homozygosity by descent. Of the six novel SNPs that were identified by sequencing genes in the 1.9 Mb candidate interval, the only polymorphism in a coding region was a c.257A>G, p.D86G mutation in *EMG1*. This mutation segregated completely with disease, as all patients shared the c.257A>G change, and it was not found in any of 414 non-Hutterite control alleles. These data support our

original hypothesis of homozygosity by descent. Additionally, *Emg1* is essential for growth and mitosis in yeast[113], confirming our assumption that the gene responsible for BCS would have a fundamental function.

Several lines of evidence indicated that the *EMG1* mutation was capable of causing a disease like BCS. First, a multiple sequence alignment showed that the EMG1 protein sequence was highly conserved, even in Archaea, indicating that it likely has an essential function. Significantly, D86 was invariant in all of the orthologs examined, suggesting a critical role for this residue in EMG1 structure or function. Because human and *S. cerevisiae* EMG1 protein sequences are 51% identical, and human EMG1 complements a yeast Emg1 deficiency[113,114], they likely share functions as well. Studies of yeast Emg1 function will be consequently applicable to understanding the human protein. Moreover, the alignment also showed that human and mouse EMG1 protein sequences are 89% identical, raising the possibility of generating a mouse model of BCS.

The second line of evidence supporting EMG1 D86G as the causative mutation is our model of human EMG1 based on previously published yeast and Archaea crystal structures, which showed that D86 is positioned at the interface of the two EMG1 subunits and stabilizes R84. R84 is known to play a role in RNA binding in yeast, and an alanine substitution at this position abolishes the ability of yeast Emg1 to bind RNA[117,119]. Although D86 does not participate directly in the dimer interface, its positioning indicates that its disruption could cause misfolding of the N-terminus of α-

helix 2, and reduce or eliminate proper homodimerization of the monomers. Dimerization is crucial to the function of SPOUT-class methyltransferases[118,139] such as EMG1, and is necessary for the formation of the RNA binding groove. Therefore, the D86G alteration could not only disrupt RNA and SAM binding, but could also lead to misfolding of the protein and expose hydrophobic residues. An increase in surface hydrophobic residues could lower protein solubility and stability, leading to the degradation or aggregation of the mutated protein.

Third, an examination of *EMG1* expression in fetal and adult tissues showed that, while *EMG1* was expressed in most tissues, levels were higher in fetal brain than in adult brain. This was supported by EMG1 protein expression in normal human brain, which was broad early in development, and more restricted to areas such as the granular cell layer of the cerebellum and the dentate gyrus of the hippocampus in the adult brain. Since the defects in BCS are primarily of a neurological nature and appear to take effect during development, these results indicate an important role for EMG1 during central nervous system development in humans. A reduction of EMG1 levels in the developing fetus could therefore result in the central nervous system defects seen in BCS-affected infants. The cerebellum plays an important role in motor control, and an EMG1 deficiency leading to reduced cell proliferation could contribute to the cerebellar hypoplasia and psychomotor retardation reported in BCS patients[2,3]. Further investigation is necessary to confirm this hypothesis, and to determine if cerebellar involvement is common to all cases of BCS. Histological analysis of BCS-affected brain tissue has not been performed to date.

Finally, although similar levels of *EMG1* mRNA were found in BCS patient and normal control fibroblasts, EMG1 protein levels were severely reduced in BCS patient cells. Since the c.257A>G mutation did not result in a stop or a nonsense codon, we did not expect an alteration in mRNA levels due to nonstop or nonsense-mediated decay. However, our EMG1 model predicted that the D86G substitution would result in an unstable protein. Thus, mRNA was normal in BCS patients, but the translated protein was likely unstable and was therefore rapidly segregated in an insoluble fraction or degraded in the cell. Intriguingly, although we showed a reduction in EMG1 protein levels, it was not a complete deficiency as some EMG1 was still detectable in the nuclear fraction of BCS patient fibroblasts and lymphoblasts. The remaining EMG1 protein is likely functional, since a complete knockout of EMG1 orthologs is lethal in yeast[112,113] and embryonic lethal in a mouse model[122]. It may be possible to stabilize D86G EMG1 before it is degraded, either by introducing or inducing molecular chaperones. Such chaperones are already in use as treatments for conditions including neurodegenerative disease and lysosomal storage disorders[142,143].

Taken together, this extensive genetic and biochemical evidence supports the conclusion that the c.257A>G, pD86G mutation in *EMG1* is responsible for Bowen-Conradi syndrome. The mutation caused a nonsynonymous change, segregated completely with disease, and was not found in any non-Hutterite alleles examined. EMG1 is a highly conserved protein with an essential function, as evidenced by studies in yeast and in

mice. Furthermore, D86 is a highly conserved residue, and the D86G substitution likely disrupts EMG1 dimerization and the conformation of RNA-binding residue R84, resulting in aggregation or degradation of the mutated EMG1. Based on these data, testing for the *EMG1* mutation was offered to individuals in the Hutterite population. In addition, cases of suspected BCS can now be accurately diagnosed. Disorders such as trisomy 18 and cerebrooculofacioskeletal syndrome (OMIM #214150) are clinically very similar to BCS[3,4,16,144] as all share growth and psychomotor retardation, microcephaly, micrognathia, limb contractures, and failure to thrive leading to early death. Cases of BCS outside the Hutterite population have been reported and disputed[10,145,146], and *EMG1* testing can assist in the workup of patients with comparable phenotypes.

Chapter 4 : Biochemical characterization and localization of wild type and D86G EMG1

Parts of this work were published as: Armistead J and Triggs-Raine B. EMG1 harbouring the Bowen-Conradi syndrome D86G substitution localizes normally but has reduced stability. In *Ribosomes: Molecular structure, role in biological functions and implications for genetic diseases*, Zhou Lin and Wang Liu, eds. Nova Science Publishers, New York. 2013; pp 147-160.

4.1 Introduction

The structure-function relationship dictates that protein function is directed by its structure. Accordingly, an understanding of the biochemical causes of BCS first requires a clear picture of EMG1 structure. The crystal structure of EMG1 orthologs in *Saccharomyces cerevisiae* and the archaebacterium *Methanocaldococcus janaschii*[117,119] revealed a distinctive core five-stranded β-sheet and a C-terminal deep trefoil knot in the amino acid backbone, hallmarks of members of the SPOUT-class methyltransferases[118]. All characterized members of the SPOUT family exist as homodimers in solution. In the EMG1 homodimer, two parallel α-helices form the surface of the interface in one subunit and sit perpendicular to the two parallel α-helices in the other, forming a core four-helix bundle in the dimer stabilized by hydrophobic interactions[119]. The enzyme thus has two active sites and two binding sites for SAM, the methyl donor. Dimerization creates an extended, positively charged groove, which is thought to bind RNA. The aspartate which is substituted by glycine in BCS patient cells, D86, is positioned near the end of alpha helix α2, one of the two parallel helices that form the interface between the two monomers. D86 binds and orients residue R84, which is necessary for RNA binding[117]. In addition, D86 is placed near the interface of the two dimer subunits. Substitution of aspartate for glycine is predicted to disrupt the alpha helix, collapsing the structure of the protein and revealing hydrophobic residues[147]. This could result in aggregation of the mutated protein to reduce the hydrophobic area exposed to the aqueous environment.

We hypothesized that the loss of structure in D86G EMG1 would lead to its degradation by the proteasome, one of the degradation pathways for misfolded proteins (reviewed

in[148,149]). Proteins to be degraded undergo post-translational modification by ubiquitin, an 8.5 kDa protein[149-151]. Briefly, ubiquitin is first covalently conjugated to a ubiquitin-activating enzyme (E1) in an ATP-dependent reaction. Ubiquitin is then transferred to a ubiquitin-conjugating enzyme (E2), and a ubiquitin-protein ligase (E3) transfers the ubiquitin moiety to a specific substrate protein. This monoubiquitin can be extended by the addition of further ubiquitin moieties to form polyubiquitin chains (Reviewed in[151]). The fate of the tagged protein depends on the polyubiquitination pattern. Chains extended using the internal lysine 48 residue of ubiquitin mark the protein to be sent to the 26S proteasome where it is deubiquitinated, unfolded, and degraded by proteolytic cleavage in an ATP-dependent manner[150,151]. We hypothesized that proteasome inhibition would block EMG1 degradation and restore D86G EMG1 levels in BCS patient cells[152]. Importantly, both proteasomes and ubiquitin have been localized to the nucleus and to pre-ribosomal complexes, where EMG1 is highly concentrated[153-155]. Proteasome inhibition has been shown to alter the mobility and localization of certain ribosomal proteins and processing factors[155], possibly due to the accumulation of large complexes which cannot cycle through nucleolar subcompartments at normal rates. Consequently we also examined the sub-cellular localization of EMG1 before and after proteasome inhibition, hypothesizing that the predicted aggregates formed in BCS patients would alter the localization of EMG1.

Ribosome biogenesis is a tightly regulated process, and is highly dependent on localization. The genes encoding ribosomal RNA are found in tandem repeats separated by intergenic spacers on five different chromosomes in humans[45,46], such that the human

genome contains approximately 400 copies of the rDNA[42-44]. These areas of multiple rDNA copies are considered nucleolar organizer regions (NORs), as they have the capacity to form nucleoli upon rDNA transcription[50]. The subsequent assembly of ribosomes is compartmentalized into nucleolar sub-structures visible by electron microscopy. The fibrillar centre (FC) is at the centre of the NOR, and the transition zone between the FC and the surrounding dense fibrillar component (DFC) is likely the site of rDNA transcription[156,157]. The DFC itself is found within the granular component (GC) (see Figure 1.1). Although the exact role of each of these compartments and what constitutes the boundaries between them are somewhat unclear, it is generally accepted that the nascent ribosome starts its life in the transition zone between the FC and the DFC, and moves outwards through the DFC and the GC as it is processed and modified, before being exported to the nucleoplasm and then the cytoplasm for final maturation[51,52]. Each compartment can therefore be associated with proteins that are representative of its function: for example the FC can be marked by upstream binding factor, a transcriptional protein, the DFC can be marked by fibrillarin, a methyltransferase which modifies precursor rRNA co-transcriptionally, and the GC can be marked by the late processing factor B23/ nucleophosmin[53]. Proper localization of each of these proteins is accordingly essential for its function, and mislocalization of ribosome-associated proteins has been connected with the ribosomopathies Treacher Collins syndrome and Diamond-Blackfan anemia[158-160]. Moreover, both yeast and human GFP-tagged EMG1 bearing the BCS-causing mutation are mislocalized to the cytoplasm when they are overexpressed[72]. Since ribosome biogenesis takes place exclusively in the nucleolus, proper localization of EMG1 is likely essential for its function.

The highly structured nature of ribosome biogenesis also means that proper ribosome assembly requires multiple interactions between specific protein partners. Protein- and snoRNA-containing particles called processomes form during rRNA transcription to carry out specific tasks including rRNA cleavage and modification[60,161,162]. Emg1 in yeast is a member of the small subunit processome, a 2.2 MDa ribonucleoprotein complex which contains the U3 snoRNA, 35 non-ribosomal proteins and 5 ribosomal proteins, and modifies the rRNA co-transcriptionally[59,60]. Members of the small subunit processome tend to co-purify with each other, making it difficult to tease out relevant interactions. Nevertheless, Emg1 in yeast has been shown to interact with Nop14, another member of the small subunit processome, and the nuclear localization of Emg1 is dependent upon binding with Nop14[113,161]. Nop14 is essential for yeast ribosome biogenesis and cell growth[113], and an ENU mutagenesis screen showed that mutations in *Nop14* in mice are pre-implantation lethal[163,164]. Although the function of NOP14 is undefined, it associates with the nuclear pore complex and may be involved in nuclear-cytoplasmic transport of the small ribosomal subunit[165,166]. We hypothesized that the D86G substitution in EMG1 would disrupt its association with protein partners, and therefore examined its binding and co-localization with NOP14 as an example.

4.2 Examination of the effects of the D86G substitution on EMG1 stability

4.2.1 Steady-state overexpression of EMG1

To study the effects of the D86G substitution on EMG1 function, we transiently overexpressed HA-tagged human wild type (WT) and D86G EMG1 in Baby Hamster Kidney (BHK) cells. Cell lysates were separated into RIPA-soluble and –insoluble

fractions, and EMG1 was revealed by immunoblotting using an anti-HA antibody. The HA-tagged EMG1 was detected at the expected apparent molecular mass of 28 kDa (Figure 4.1 A). In the soluble fraction, D86G EMG1 levels were reduced, reflecting the endogenous EMG1 results found in BCS patient fibroblasts and lymphoblasts (see section 3.4.3.2). However, in the RIPA-insoluble fraction the amounts of D86G and WT EMG1 were similar (Figure 4.1 A), indicating that D86G EMG1 preferentially associated with the insoluble fraction. These results were in accordance with the human EMG1 model described in Chapter 3, which predicted that the D86G substitution would result in misfolding and subsequent aggregation of the protein. The aggregated protein would therefore be more likely to associate with the RIPA-insoluble fraction, leaving less detectable soluble protein. However, given that the total WT EMG1 protein levels still exceeded those of the D86G EMG1, it seemed likely that the mutated protein was susceptible to degradation.

4.2.2 Pulse-chase metabolic labelling of overexpressed EMG1

To characterize the kinetics of D86G EMG1 reduction in the soluble fraction and association with the insoluble fraction, pulse-chase analysis of the EMG1 protein was performed. Nascent EMG1 protein was labelled using ^{35}S cysteine/methionine, and followed at intervals during a 24-hour chase period to observe its presence in soluble and insoluble fractions. In the soluble fraction, EMG1 was immunoprecipitated using an anti-HA antibody; however the high detergent concentration in the insoluble fraction made it impossible to perform an immunoprecipitation. Consequently, equal volumes of the

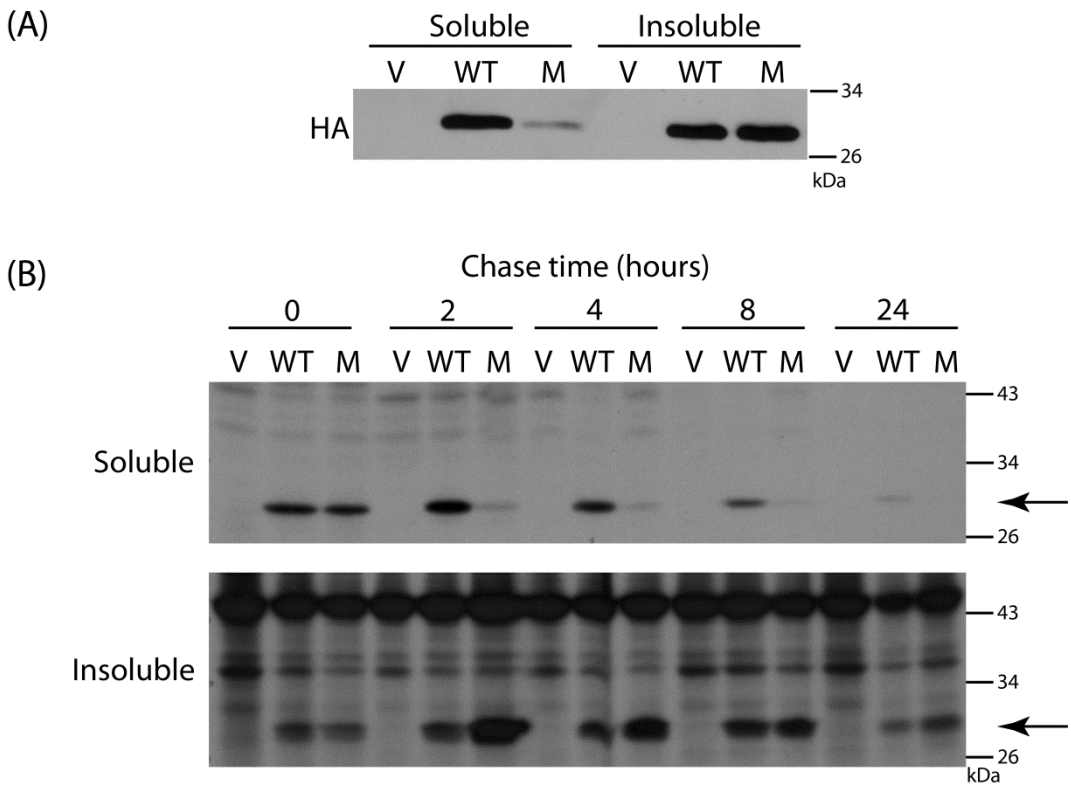


Figure 4.1 Transiently overexpressed HA-tagged EMG1 in BHK cells.

(A) HA-tagged mutant EMG1 protein transiently expressed in BHK cells is preferentially associated with the insoluble fraction at steady state. The RIPA-soluble fraction from BHK cells transiently expressing HA-tagged *EMG1* was separated from the insoluble fraction by centrifugation. The insoluble fraction was then resuspended in RIPA and sonicated before loading approximately 20 µg of protein, and separating in a 10% SDS-PAGE gel. Protein was then transferred to a nitrocellulose membrane and detected by western blot using an anti-HA antibody. Cells were co-transfected with pRCMVβ-gal to serve as a transfection control; loading was therefore corrected for both protein and β-galactosidase activity. V, vector control; WT, wild type EMG1 protein; M, mutant D86G EMG1 protein. **(B)** BHK cells were transiently transfected with empty vector (V), wild type (WT), or mutant (M) EMG1-expressing plasmids. Following a 10-minute labelling of nascent proteins with ^{35}S-cysteine and methionine, and incubation in unlabeled medium for 2 to 24 hours, cell lysates were separated into detergent-soluble and detergent-insoluble fractions. EMG1 was immunoprecipitated from the soluble fraction using an anti-HA antibody before gel loading, and the insoluble fraction was sonicated and loaded directly onto the 10% polyacrylamide gel. The loading was normalized to protein concentration and β-galactosidase activity. Immediately following labelling (0), WT and mutant EMG1 were present at equal levels in the soluble fraction, but levels of the mutant protein were quickly reduced compared to WT protein. V, vector control; WT, WT EMG1 protein; M, mutant D86G EMG1 protein. The arrows indicate the position of EMG1.

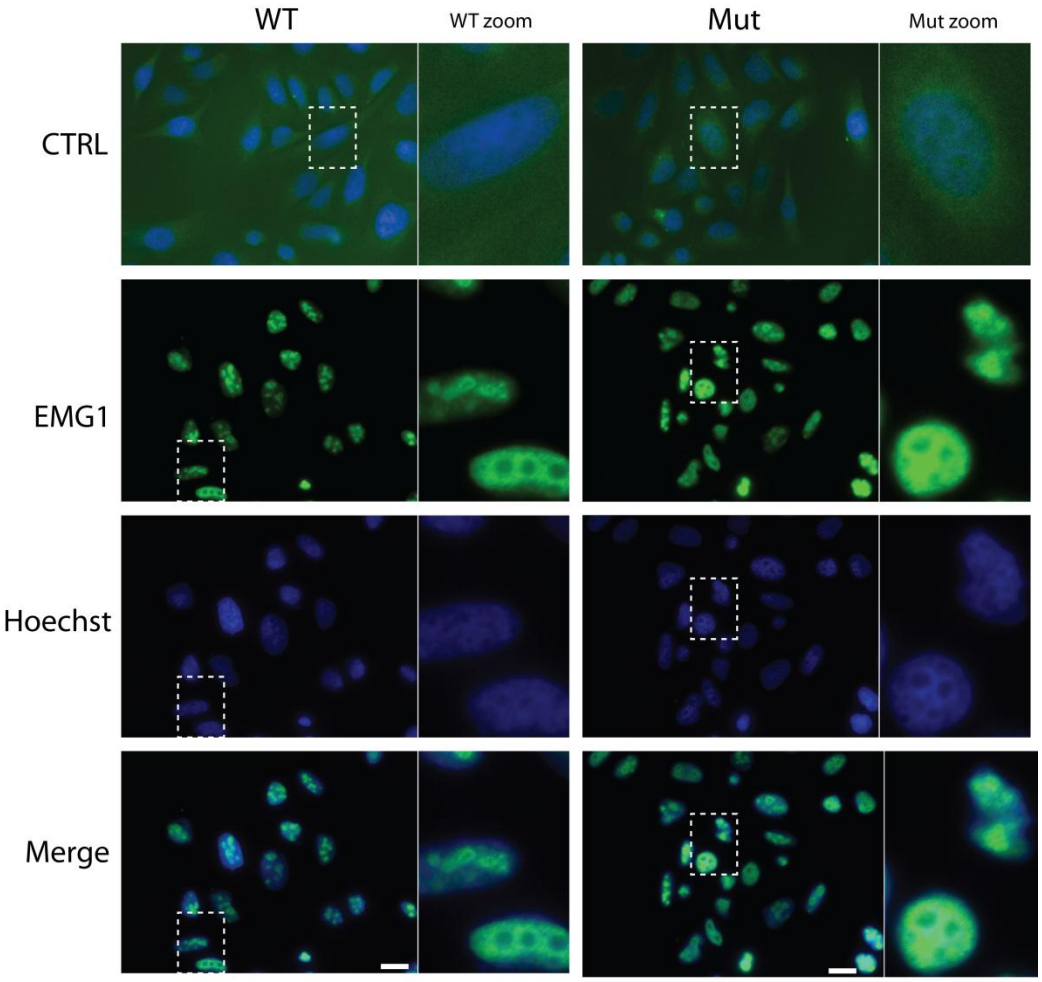
insoluble fraction were directly loaded onto the gel. Autoradiography showed strong bands at the expected molecular weight of 28 kDa that were not present in cells transfected with the empty vector (Figure 4.1 B). WT and D86G EMG1 started out at similar levels in the soluble fraction immediately following ^{35}S labelling (0 hours), but the D86G EMG1 levels diminished within two hours, in contrast to the WT EMG1 which was not reduced to a similar level until 24 hours. An examination of the insoluble fraction revealed the opposite results, with the mutant D86G EMG1 at much higher levels than WT at the 2 hour time point. Together, the detergent- soluble and –insoluble fractions show that D86G and WT EMG1 were translated at similar levels, but that the D86G EMG1 became rapidly insoluble, leaving very little soluble protein. These data support the hypothesis that the overexpressed D86G EMG1 aggregates quickly following translation.

4.3 Determination of the localization of the wild type and mutant EMG1 protein

4.3.1 Localization in HeLa cells

To determine if the changes in D86G EMG1 solubility were reflected in its sub-cellular localization, EMG1 was visualised in HeLa cells (a human cervical cancer cell line) using an anti-EMG1 antibody and immunofluorescent microscopy. The EMG1 signal was predominantly nuclear, but appeared to have discrete modes of subnuclear organization. In some cells, the EMG1 signal was detectable in nucleoli, and in others the signal was predominantly nucleoplasmic (Figure 4.2 A). In some cells the EMG1 signal was diffuse in the nucleus, with no specific area of localization. This pattern was apparent in cells stably overexpressing either WT or D86G EMG1. To determine the distribution of cells

(A)


WT
WT zoom
Mut
Mut zoom
CTRL
EMG1
Hoechst
Merge

(B)

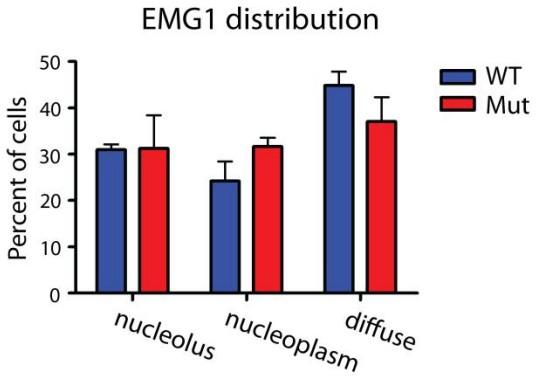

EMG1 distribution
Percent of cells
50
40
30
20
10
0
WT
Mut
nucleolus
nucleoplasm
diffuse

Figure 4.2 EMG1 localization in HeLa cells.

(A) EMG1 is localized to the nuclei in HeLa cells. HeLa cells stably expressing either WT or mutant EMG1 were examined by fluorescence microscopy following fixation using 4% formaldehyde. EMG1 was detected using an anti-EMG1 antibody (green), and nuclei were revealed using DNA-binding Hoechst 33342 (blue). Cells that were not exposed to the primary antibody were used as a control. EMG1 localization took three different forms: either primarily nucleolar, primarily nucleoplasmic, or diffuse throughout the nucleus. The areas boxed in white are magnified on the right side. Original magnification, 63 x; scale bar = 20 µm. **(B)** Distribution of HeLa cells with primarily nucleolar, primarily nucleoplasmic, or diffuse EMG1 staining. For each cell line, cells in five random fields in two independent experiments were counted. Columns show the mean percentage of cells displaying the indicated staining pattern of EMG1, and error bars indicate the standard deviation. No significant difference was found between cells expressing WT or mutant EMG1.

displaying nucleolar and nucleoplasmic staining, cells were counted in five random fields on each of two independently prepared slides, for a total of ten fields per cell line (Figure 4.2 B). No apparent difference in distribution was found in cells overexpressing D86G EMG1 when compared to WT EMG1. Thus, the aggregation detected by pulse-chase analysis of overexpressed D86G EMG1 was not visible by immunofluorescent microscopy using this method. A microscopic analysis of cells taken during the pulse-chase experiment might have revealed a difference in localization patterns between the WT and D86G EMG1-expressing HeLa cells.

4.3.2. Localization in fibroblasts

The sub-cellular localization of endogenous EMG1 was also examined in normal control and fibroblasts from BCS patients. In fibroblasts, EMG1 was generally localized to nuclei with strong signals in the nucleoli. Co-labelling with fibrillarin, a well-characterized nucleolar protein, confirmed nucleolar localization, likely in the dense fibrillar component consistent with a role in rRNA modification (Figure 4.3 A). Contrary to the HeLa cells, the evenly distributed nucleoplasmic signal was not observed. However, an EMG1 signal was present in small punctate structures throughout the nuclei, which did not co-localize with fibrillarin. WT and D86G EMG1 appeared to localize similarly although the intensity of the D86G signal was fainter (Figure 4.3 B), as determined by measuring the intensity of the green channel in ten random fields using Axiovision software. In the same cells, the signal intensity of fibrillarin was not significantly different between WT and D86G cells (Figure 4.3 C). The mean area of nucleoli was also measured to assess the effect of the D86G substitution in EMG1 on

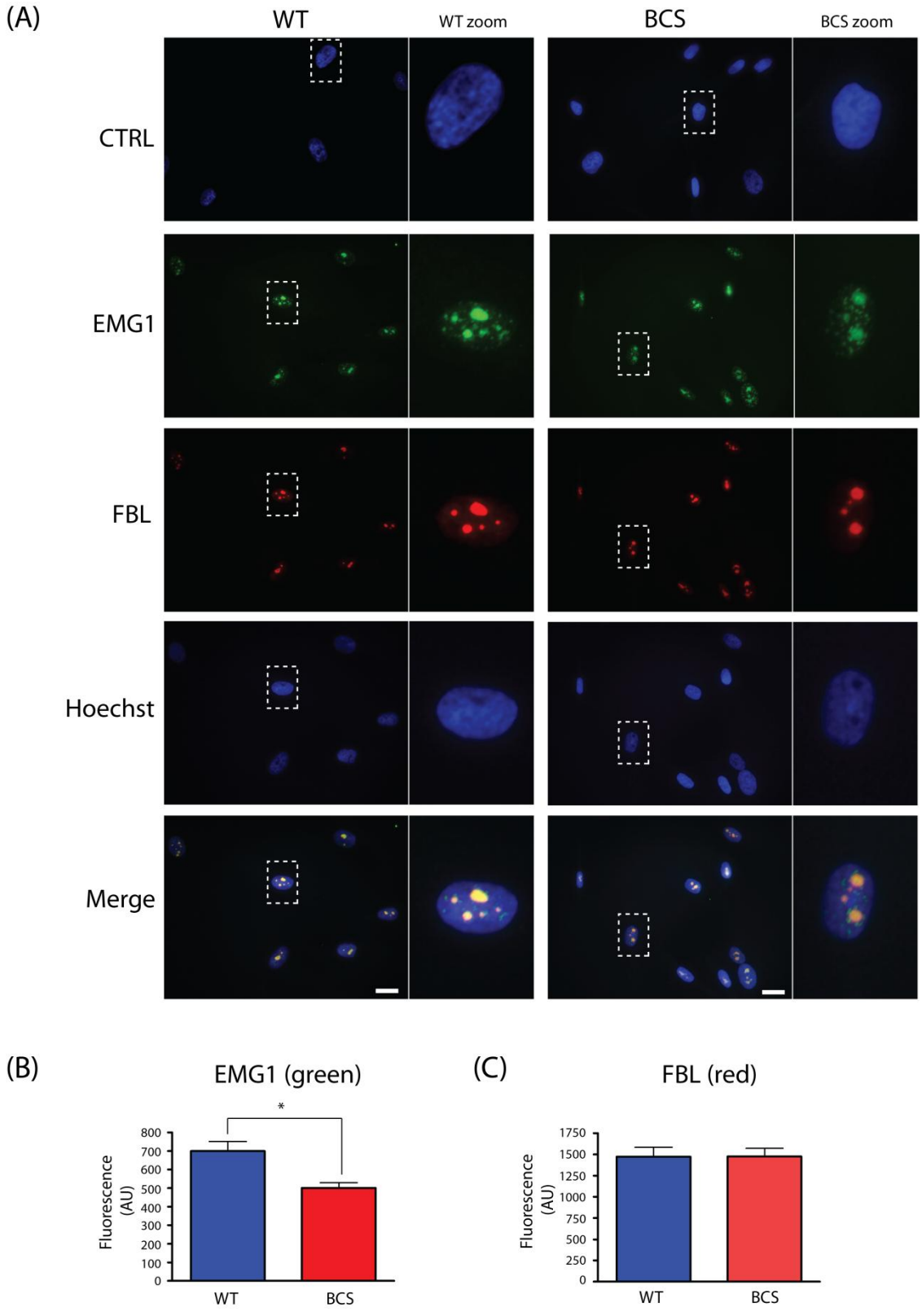

(A)
WT
WT zoom
BCS
BCS zoom
CTRL
EMG1
FBL
Hoechst
Merge
(B)
EMG1 (green)
*
Fluorescence (AU)
800
700
600
500
400
300
200
100
0
WT
BCS
(C)
FBL (red)
Fluorescence (AU)
1750
1500
1250
1000
750
500
250
0
WT
BCS

Figure 4.3 EMG1 and fibrillarin co-localization in fibroblasts.

(A) EMG1 is localized to nucleoli in fibroblasts. Sub-confluent wild type (WT) or patient (BCS) fibroblasts were visualized using fluorescence microscopy following fixation using 4% formaldehyde. EMG1 was detected using an anti-EMG1 antibody (green), while fibrillarin (FBL), a nucleolar protein, was detected using an anti-fibrillarin antibody (red). Nuclei were revealed using DNA-binding Hoechst 33342 (blue). Cells that were not exposed to the primary antibody were used as a control. The merged photo shows co-localization of EMG1 and fibrillarin in nucleoli of both WT and BCS cells. Original magnification, 63 x; scale bar = 20 µm. **(B)** The intensity of the EMG1 signal is lower in BCS patient fibroblasts than WT controls. In two independent experiments, two slides each of WT control and BCS patient fibroblasts were created, and photos were taken of five random fields for each slide, for a total of ten fields per cell line per experiment. In each field, individual nucleoli were isolated and the intensity of the green channel signal was measured using Axiovision software. The graph shows the mean fluorescence in arbitrary units of the green channel signal in WT and BCS fibroblast nucleoli; error bars indicate SEM. Results were compared using Student's *t* test, $p = 0.0034$. **(C)** The intensity of the FBL signal is similar in WT and BCS patient fibroblasts. Using the same slides as in (B), the intensity of the red channel was measured. The graph shows the mean fluorescence in arbitrary units of the red channel signal; error bars indicate SEM. Results were compared using Student's *t* test, $p = 0.9873$.

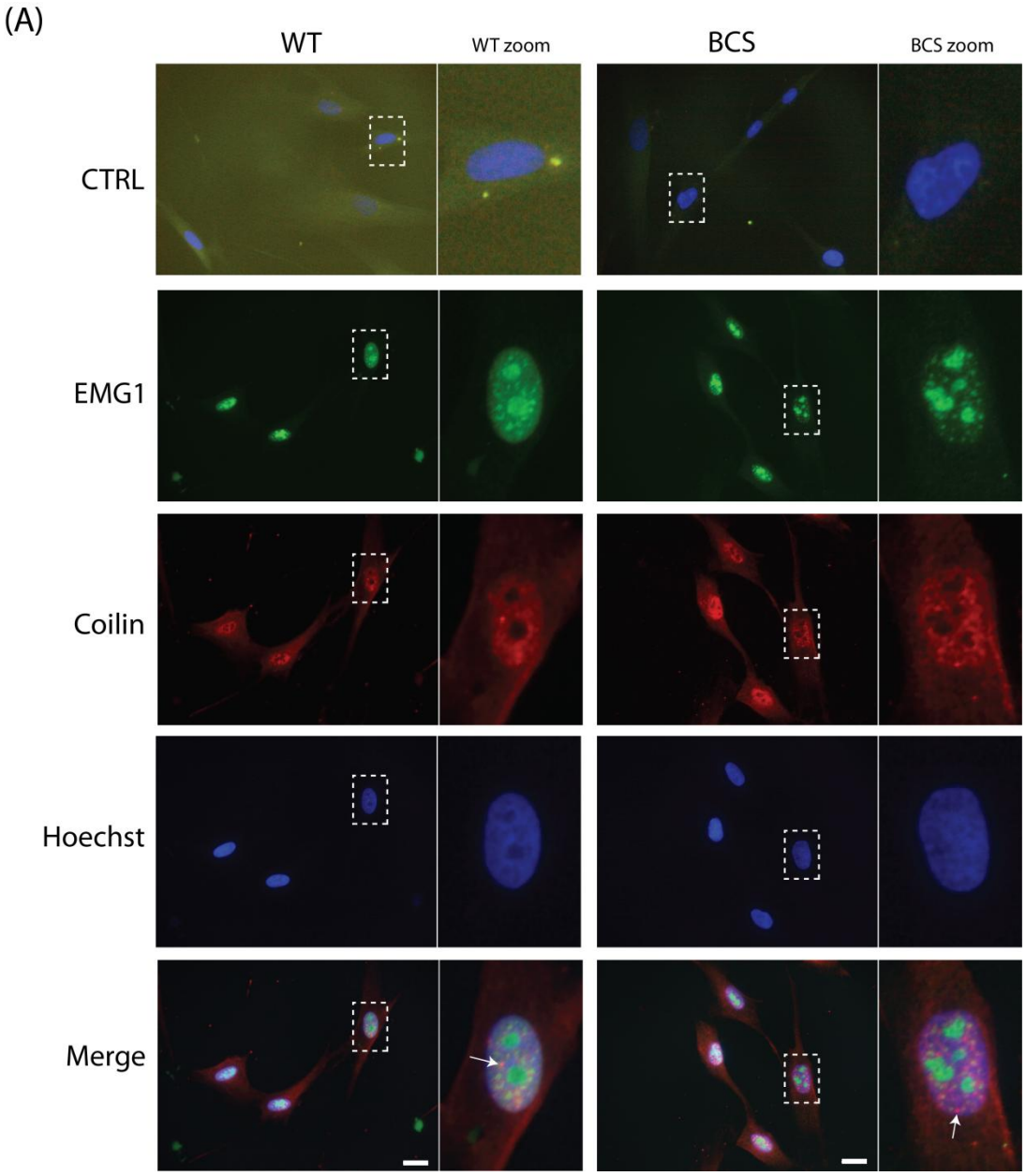
nucleolar stress, which resulted in no significant difference between the WT and the D86G cell lines (not shown). These data suggest that, while the sub-cellular localization of EMG1 is unchanged in BCS patient fibroblasts, the amount of EMG1 protein is reduced, in agreement with our immunoblot findings (see Figure 3.5 C).

The appearance of EMG1 in nuclear structures outside the nucleolus could indicate an extra-ribosomal role for EMG1. In an effort to identify the nature of the EMG1-containing puncta, cells were co-labelled with coilin or SMN to mark Cajal bodies[167], which also contain RNA-modifying enzymes. EMG1 did not co-localize with these proteins in fibroblast nuclei (Figure 4.4 A and 4.4 B), indicating that the nucleoplasmic EMG1-containing puncta do not correspond with Cajal bodies, and that EMG1 likely does not modify Cajal body snoRNAs.

4.4 Examination of the mechanism by which the EMG1 mutation causes BCS

4.4.1 Determination of the mechanism of EMG1 degradation

Since our EMG1 protein model predicted that the structure of the mutated protein would be compromised, it seemed likely that the D86G EMG1 would be quickly degraded by the cell[149,155]. To determine if EMG1 was degraded via the proteasome pathway, proteasome activity was inhibited using either MG132 or Proteasome Inhibitor I (PI1). Following overnight treatment with proteasome inhibitor, proteins were isolated from nuclear lysates and separated using SDS-PAGE. As expected, proteasome inhibition led to the accumulation of ubiquitinated proteins, detected using an anti-ubiquitin antibody (Figure 4.5 A, "MG132" lanes). Overnight treatment with MG132 consistently gave rise


(A)
WT
WT zoom
BCS
BCS zoom
CTRL
EMG1
Coilin
Hoechst
Merge

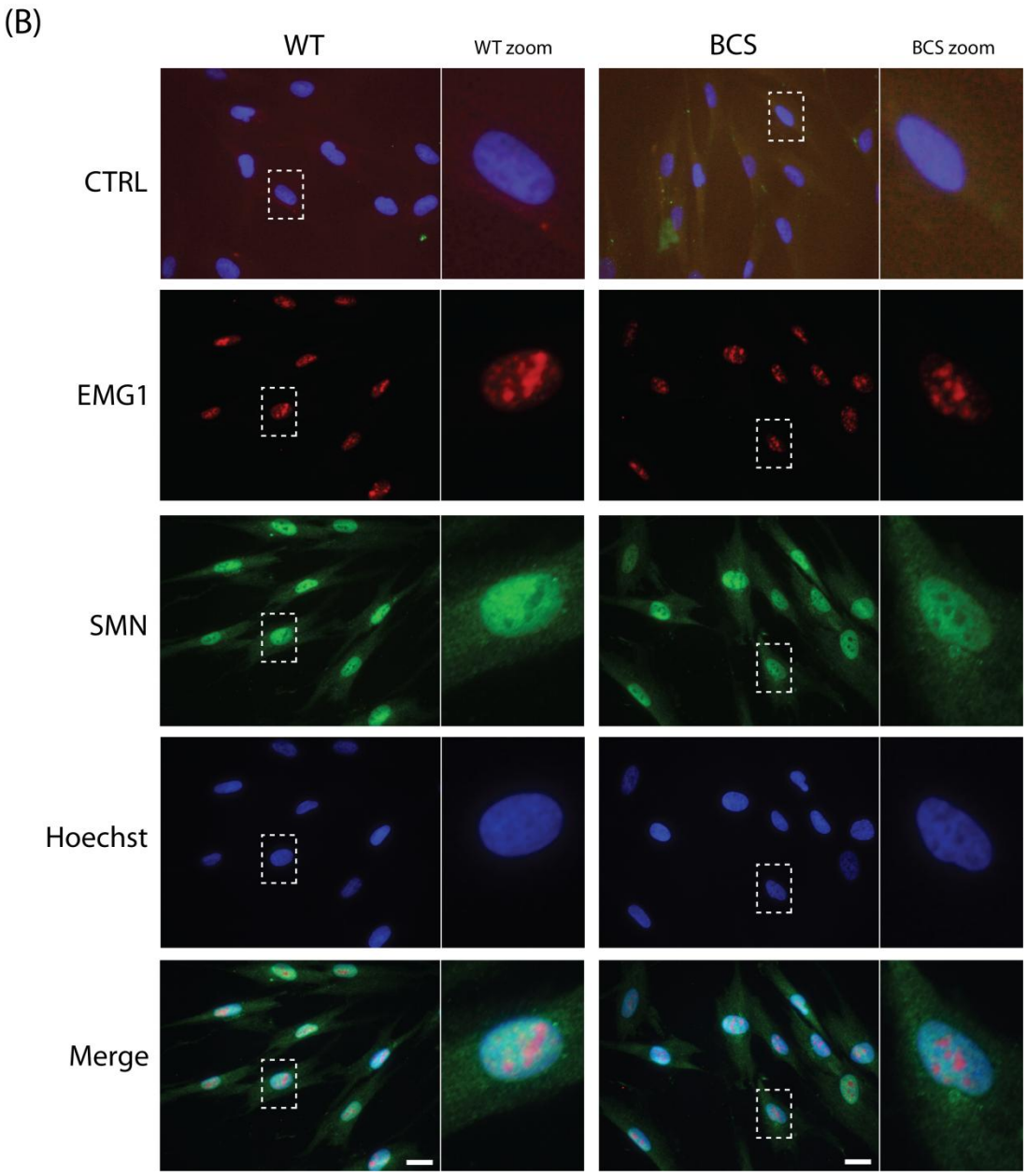


Figure 4.4 EMG1, coilin, and SMN localization in fibroblasts.

WT and BCS EMG1 do not co-localize with either coilin **(A)** or SMN **(B)**. EMG1 was detected using an anti-EMG1 antibody (green in A and red in B). Coilin was detected using an anti-coilin antibody (red in A), and SMN was labelled with an anti-SMN antibody (green in B). Nuclei were revealed using DNA-binding Hoechst 33342 (blue). Cells that were not exposed to the primary antibody were used as a control. In A, Cajal bodies that do not contain EMG1 are indicated with white arrows in the merged photo. Original magnification, 63 x; scale bar = 20 µm.

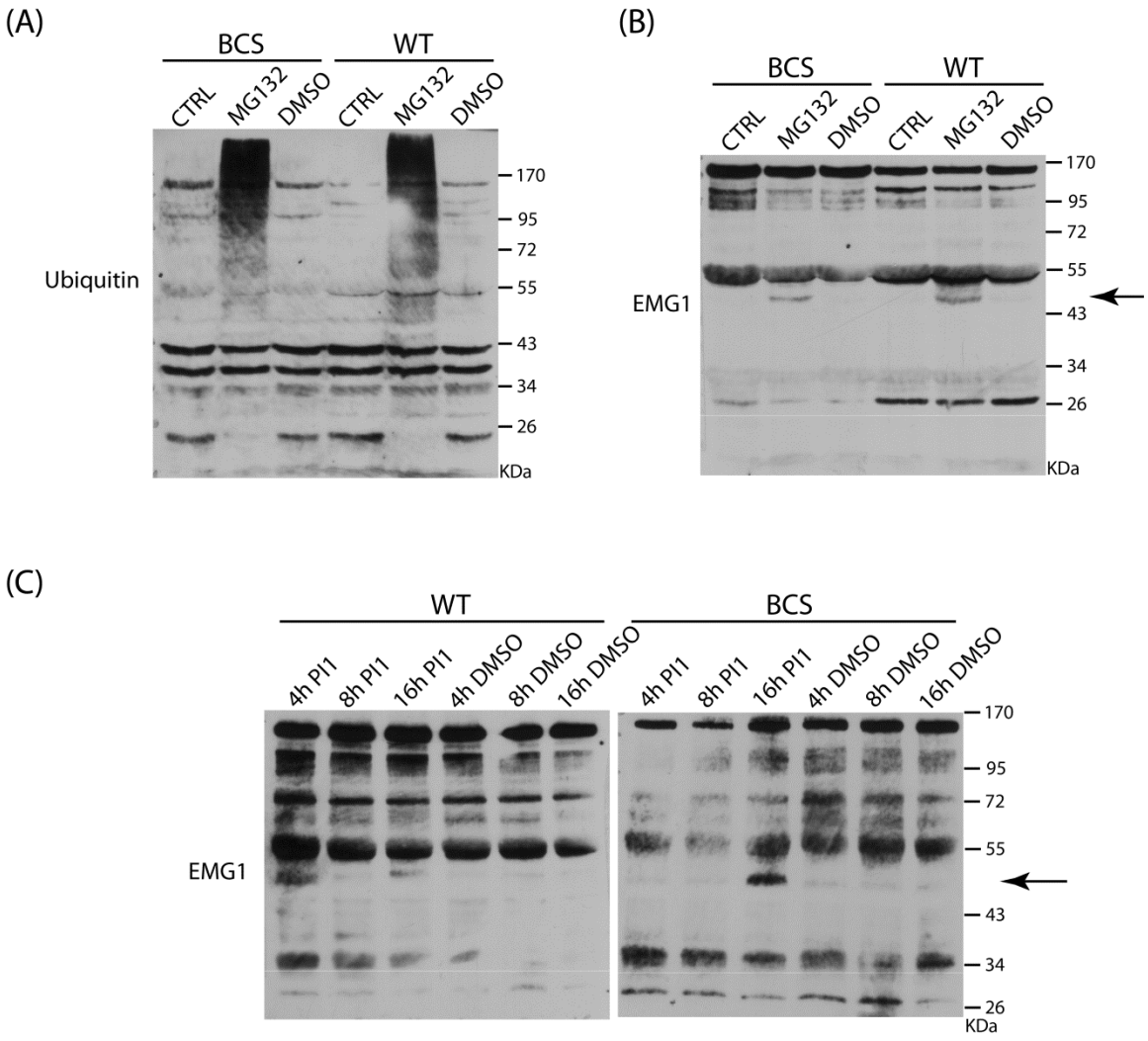


Figure 4.5 EMG1 following proteasome inhibition.

(A) Ubiquitinated proteins upon proteasome inhibition using MG132. Normal control (WT) or patient (BCS) fibroblasts were treated overnight with MG132 or vehicle (DMSO) alone, and cell lysates were separated in a 10% SDS-PAGE gel. Lysates from cells taken prior to treatment were used as a control (CTRL). Proteins were transferred to a nitrocellulose membrane and probed with an anti-ubiquitin antibody. The presence of high molecular weight species revealed by the ubiquitin antibody after MG132 treatment indicated the accumulation of ubiquitinated proteins upon inhibition of their degradation by the proteasome (dark bands in MG132 lanes). **(B)** EMG1 upon proteasome inhibition using MG132. Probing with anti-EMG1 antibody revealed an extra band between 43 and 55 kDa (arrow indicates) only in samples treated with MG132, which may indicate the accumulation of ubiquitinated EMG1. **(C)** EMG1 upon proteasome inhibition using PI1. Fibroblasts were treated for 4, 8, or 16 hours with either PI1 or DMSO alone, and cell lysates were separated in a 10% SDS-PAGE gel. The arrow indicates the accumulation of a similar molecular weight band as in (B) upon PI1 treatment and probing with an anti-EMG1 antibody.

to a band at approximately 50 kDa detected by an anti-EMG1 antibody in nuclear fractions, which was absent in controls treated with DMSO alone (Figure 4.5 B, arrow). This band may correspond to polyubiquitinated EMG1 as it is a similar molecular mass as EMG1 plus three ubiquitin moieties (52.5 kDa). The same band was present on immunoblots from cells treated overnight with PI1 (Figure 4.5 C, arrow). These data suggest that EMG1 is normally degraded via the proteasome pathway. However, although proteasome inhibition strongly increased the putative polyubiquitinated EMG1 band, it did not restore D86G EMG1 to levels similar to WT as expected. It is possible that inhibition of the proteasome pathway would cause aggregation of the protein, and an examination of the insoluble fraction would have revealed a significant increase in D86G EMG1.

To determine if proteasome inhibition altered sub-cellular localization, EMG1 was detected in fibroblasts by immunofluorescence following treatment with DMSO alone (Figure 4.6 A) or with MG132 (Figure 4.6 B). Upon overnight proteasome inhibition, EMG1 was re-localized from nucleoli and accumulated in nuclear puncta and in the cytoplasm (Figure 4.6 B). The nuclear puncta were very similar to those described in section 4.3.2, indicating that EMG1 in fibroblasts may normally localize to nuclear proteasomes. No differences in localization were observed between WT and BCS patient cells.

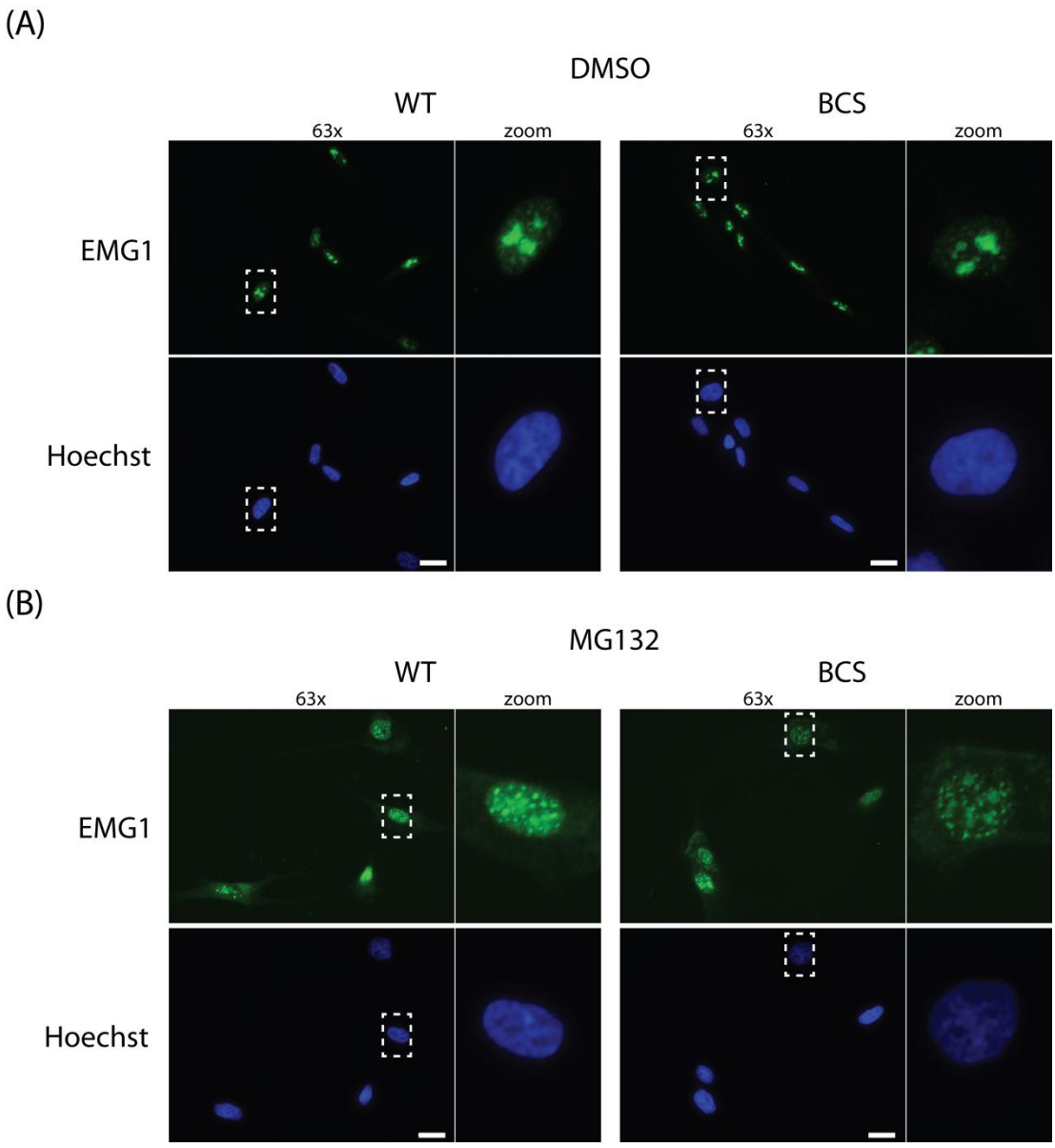


Figure 4.6 EMG1 localizes to nuclear foci upon proteasome inhibition.

(A) EMG1 was detected using an anti-EMG1 antibody (green) in fibroblast cells incubated overnight with vehicle (DMSO) alone. Nuclei were labeled with Hoechst 33342 (blue). **(B)** EMG1 re-localizes to nuclear puncta upon proteasome inhibition. Fibroblasts were treated with MG132 overnight to inhibit proteasome activity. EMG1 nucleolar localization was reduced upon MG132 treatment, and accumulation in nucleoplasmic puncta was increased. Original magnification, 63 x; scale bars, 20 µm.

4.4.2 Examination of the effect of the decreased stability of EMG1 on its interactions with NOP14

4.4.2.1 Co-immunoprecipitation of EMG1 and NOP14

Very little is known about EMG1 binding partners, although it co-purifies with members of the small subunit processome[113,161]. It has nonetheless been shown in yeast that Nop14, a stress-response protein, binds to Emg1. Emg1 is absent from the nuclear fraction of yeast cells depleted of Nop14, indicating that nuclear localization of Emg1 depends on Nop14 binding[113]. To elucidate this interaction further and examine the effect of the D86G substitution on NOP14 binding, co-immunoprecipitation studies were performed. HeLa cells stably expressing HA-tagged WT or D86G EMG1 were co-transfected with FLAG-tagged NOP14 and a β-galactosidase-expressing plasmid as a transfection control. Co-immunoprecipitation using anti-FLAG and blotting using anti-HA confirmed an interaction between NOP14 and WT EMG1 (Figure 4.7, arrow). The interaction was very weak when compared with total HA-tagged EMG1 in cell lysates, but previous results in yeast indicated that only a fraction of Nop14 and Emg1 associate with each other[113]. Co-immunoprecipitation with HA-tagged D86G EMG1 and FLAG-tagged NOP14 did not show any interaction (Figure 4.7). It is not clear whether the mutated EMG1 was unable to bind to NOP14 because of structural changes, or whether the low levels of D86G EMG1 precluded a detectable interaction. The inverse immunoprecipitation was unsuccessful.

4.4.2.2 Sub-cellular localization of EMG1 and NOP14

Co-localization of EMG1 and NOP14 was examined *in vitro* using immunofluorescence. Interestingly, in HeLa cells transiently transfected with FLAG-tagged NOP14, NOP14

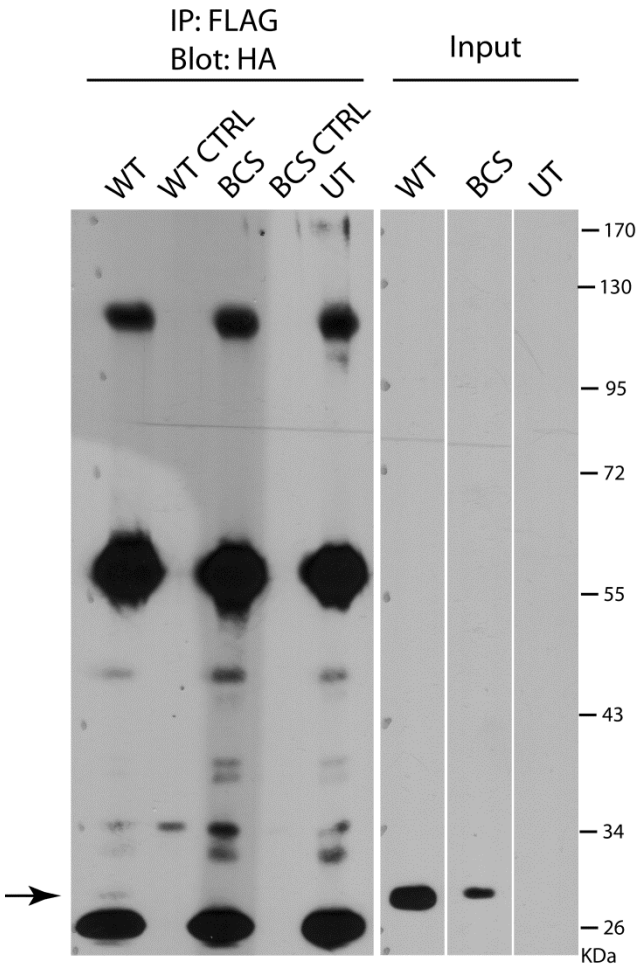
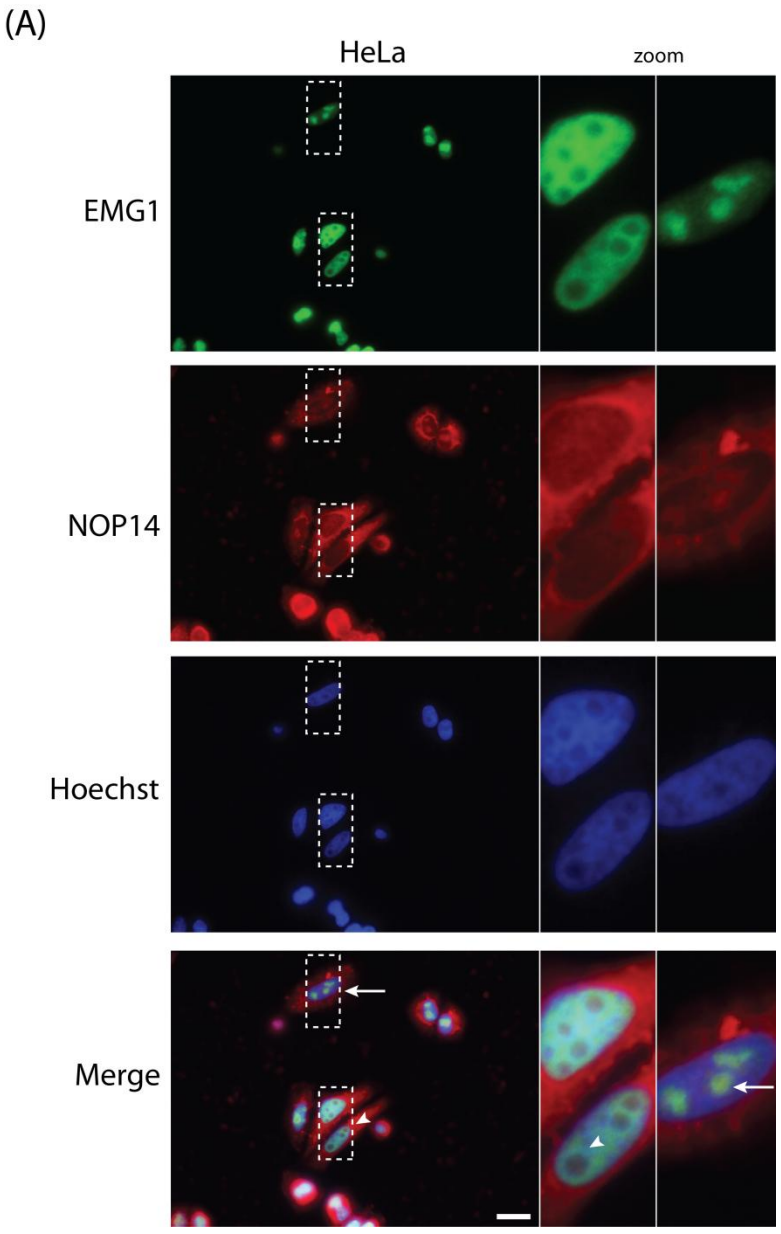


Figure 4.7 EMG1 interaction with NOP14.

Co-immunoprecipitation shows that a fraction of wild type EMG1 is associated with NOP14. HeLa cells transiently transfected with FLAG-tagged NOP14 and stably expressing wild type (WT) or mutant EMG1 (BCS) were used, and untransfected (UT) HeLa cells were used as a negative control. Immunoprecipitated protein using an anti-FLAG antibody was separated in a 10% SDS-PAGE gel, and transferred to a nitrocellulose membrane. The membrane was blotted with an anti-HA antibody to detect EMG1 (arrow). For control samples, no primary antibody was included in the immunoprecipitation reaction. EMG1 detected in cell lysates without immunoprecipitation (input) is shown for comparison. A representative image of three independent experiments is pictured.

had a biphasic localization wherein it was either concentrated in the nucleolus or more diffuse throughout the cell, predominantly in the cytoplasm. The localization of NOP14 reflected that of EMG1: when NOP14 was concentrated in the nucleolus, EMG1 co-localized with it. However, when NOP14 was predominantly cytoplasmic, EMG1 was excluded from the nucleolus (Figure 4.8 A). This again points to a relationship between NOP14 and EMG1, and a possible role for NOP14 in dictating the sub-cellular localization of EMG1 in HeLa cells. Confirming the co-immunoprecipitation results, only a very small fraction of NOP14 appeared to co-localize with EMG1. The majority of NOP14 was found in the cytoplasm, and only co-localized with EMG1 when it was in the nucleolus in a subset of cells, suggesting a dynamic relationship between EMG1 and NOP14. It remains to be determined why NOP14 cycles in and out of nucleoli, although it has been associated with nuclear pore complexes and may play a role in nuclear transport[165,166].

In contrast, when endogenous EMG1 and NOP14 co-localization was examined in fibroblasts, both EMG1 and NOP14 were nucleolar in all cells examined (Figure 4.8 B). This suggests cell-type specific interactions between EMG1 and NOP14, affecting sub-cellular localization. Once again, no differences in localization were observed between WT and D86G EMG1 in relation to NOP14.


(A)
HeLa
zoom
EMG1
NOP14
Hoechst
Merge

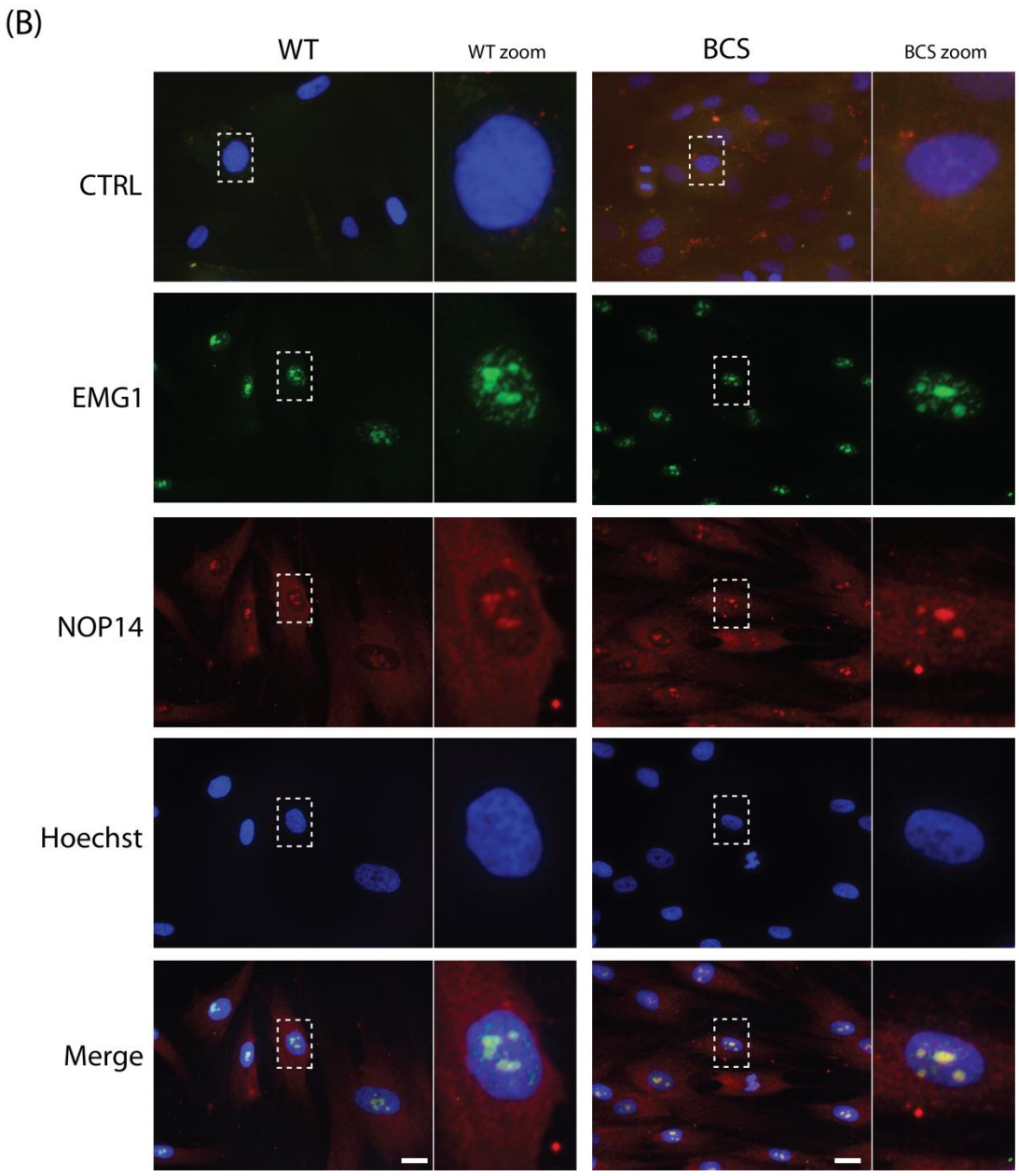


Figure 4.8 EMG1 and NOP14 localization in HeLa cells and fibroblasts.

(A) EMG1 and NOP14 co-localize in only a subset of HeLa cells. HeLa cells were transiently transfected with FLAG-tagged NOP14 and detected using an anti-FLAG antibody (red), while endogenous EMG1 was detected using an anti-EMG1 antibody (green). When EMG1 was localized to the nucleoli, NOP14 was also concentrated in the nucleoli (see arrow in merged photo). When EMG1 was excluded from the nucleoli and was found only in the nucleoplasm, NOP14 was excluded from the nucleus (arrowhead in merged photo). DNA was stained with Hoechst 33342 to label nuclei (blue). Original magnification, 63 x; scale bar = 20 µm. **(B)** EMG and NOP14 co-localize in the nucleoli of fibroblasts. EMG1 was detected using an anti-EMG1 antibody (green), and NOP14 was detected using an anti-NOP14 antibody (red). Cells that were not exposed to the primary antibody were used as a control. Original magnification, 63 x; scale bars = 20 µm.

4.5 Discussion

All characterized members of the SPOUT family of methyltransferases exist as dimers in solution, with the extended surface created upon dimerization being necessary for binding the substrate RNA[118]. The catalytic residue is provided by the opposite subunit[118], and thus the dimeric structure is intimately linked with its enzyme function. In EMG1, a member of the SPOUT family methyltransferases, the substitution of aspartate 86 for a glycine in Bowen-Conradi syndrome is predicted to expose hydrophobic residues, reducing protein stability[147]. This was evidenced by a reduction in D86G EMG1 levels in the nuclear fraction of patient fibroblasts and lymphoblasts as discussed in Chapter 3. To further examine EMG1 stability, we overexpressed WT and D86G EMG1 in BHK cells and found that D86G EMG1 was preferentially associated with an insoluble fraction, resulting in a severe decrease in soluble D86G EMG1. An examination of EMG1 stability using pulse-chase metabolic labelling revealed that WT and D86G EMG1 were initially produced at similar levels in the soluble fraction, but that D86G EMG1 levels were rapidly reduced to a fraction of WT levels. Conversely, D86G EMG1 in the insoluble fraction accumulated quickly in comparison with WT EMG1. It is likely that the overexpressed D86G EMG1 misfolded and formed aggregates when hydrophobic residues were exposed to the surface, leaving less D86G EMG1 protein in a soluble state. This supports the Chapter 3 results which suggested that the defect caused by the mutation in *EMG1* takes place at the protein level. Additionally, using the yeast two-hybrid system, the D86G substitution in human EMG1 has been shown to increase the affinity of the EMG1 monomer for itself, and subsequent work using analytical gel filtration analysis indicated increased dimerization in yeast D90G EMG1 (the equivalent

of D86G in human EMG1)[72,147]. These data support the conclusion that the D86G substitution in EMG1 alters the conformation and binding properties of the protein.

Because ribosome biogenesis takes place primarily in the nucleolus, many studies of the molecular causes of ribosomopathies have focussed on sub-cellular localization of the mutated proteins, with variable results. In a study identifying nucleolar targeting sequences in RPS19, Da Costa *et al*[160] examined two Diamond-Blackfan anemia-causing mutations which had the potential to negatively affect sub-cellular localization of the protein. When GFP fusion proteins bearing these substitutions were expressed in Cos-7 cells, distribution of the protein was altered from the nucleolus to the nucleoplasm or cytoplasm. Notably, several other disease-associated mutations did not alter RPS19 nucleolar localization[159,160,168]. Similar studies of mutations in Treacle, the protein associated with Treacher Collins syndrome, showed mislocalization to the nucleoplasm[158]. However, SBDS proteins carrying different mutations retained their normal nucleolar localization in Shwachman-Diamond syndrome patient cells[169]. To better understand the influence of the D86G substitution on EMG1 localization, we employed immmunofluorescence using an anti-EMG1 antibody. In HeLa cells stably overexpressing WT and D86G EMG1, the sub-cellular localization was dynamic, shuttling between the nucleolus and the nucleoplasm. EMG1 localization may be cell-cycle dependent in HeLa cells, as EMG1 in HeLa cells synchronized by serum starvation showed much more uniform distribution (data not shown). An evaluation of the proportions of HeLa cells in which EMG1 was found in the nucleolus versus in the

nucleoplasm did not reveal any significant differences between the WT- and the D86G EMG1-expressing cells.

Endogenous EMG1 in fibroblasts was found primarily in nucleoli (co-localizing with fibrillarin) and in scattered nuclear puncta which were not observed in the HeLa cells. Quantification of the EMG1 signal in fibroblasts showed that the D86G staining was less intense than the WT staining, in agreement with the immunoblot results shown in Chapter 3, which indicated severely reduced levels of D86G EMG1. However, the immunofluorescence results did not indicate any apparent difference in localization between WT and D86G EMG1 in fibroblasts. These data are contrary to results published by Meyer *et al* in 2011[72], who showed that GFP-tagged D86G EMG1 in HEK293T cells and D90G Emg1 (the yeast equivalent of the D86G mutation) in yeast were mislocalized to the nucleoplasm and the cytoplasm. These proteins were overexpressed and under the control of a tetracycline-inducible promoter. It is possible that the overexpression of EMG1 in these experiments, combined with the influence of the GFP tag, led to its mislocalization. Our experiments detected endogenous D86G EMG1 in control and BCS patient fibroblasts, using an anti-EMG1 antibody, and showed that the D86G substitution alone is insufficient to alter the localization of EMG1. It seems more likely that the defects seen in BCS patients are the result of insufficient EMG1 levels in proliferating cells, rather than of its mislocalization. Moreover, our results suggest that EMG1 localizes differently in different cell types, which may play a role in the tissue-specific manifestation of the disease. Future experiments will be necessary to determine how the sub-cellular localization of EMG1 is regulated in different cell types and tissues.

In an effort to determine the nature of the EMG1-containing nuclear puncta in fibroblasts, we assessed whether it co-localized with markers of Cajal bodies, coilin or SMN. No co-localization was detected, indicating that EMG1 likely does not play an extra-ribosomal role in snRNP biogenesis or in splicing[167,170,171]. It is possible that these nucleoplasmic sites correspond to proteasomes where EMG1 is degraded, since treatment with the proteasome inhibitor MG132 caused an accumulation of EMG1 in these structures. Proteasomes have been shown to be present in the nucleus but not the nucleolus (reviewed in[172,173]), and play a role in regulating nuclear function, including ribosome biogenesis[153,155]. Co-localization with ubiquitin, or with the 26S subunit of the proteasome, would confirm this hypothesis.

We showed that D86G EMG1 levels were reduced in patient fibroblasts and lymphoblasts by immunoblot (Chapter 3) and by immunofluorescence (Figure 4.3). To determine the mechanism by which EMG1 is lost, we explored the proteasome pathway, which accounts for the majority of the protein degradation in a cell[174,175]. The ubiquitin proteasome pathway has been shown to regulate ribosome biogenesis at several levels, including rRNA transcription by RNA polymerase I, processing of the 90S pre-ribosome, and production and export of mature 28S rRNA[155]. Interestingly, ubiquitin in eukaryotic cells is produced either as a polyubiquitin protein, arranged in tandem head-to-tail repeats, or as N-terminal fusions with ribosomal proteins S27a[176,177] or L40[176]. The ubiquitin moiety acts as a chaperone during the assembly of the proteins into the nascent

ribosome, and is then cleaved for use in the ubiquitin proteasome pathway[176,178]. It therefore seems very likely that the ubiquitin proteasome system and ribosome biogenesis are intimately linked. Upon treatment with proteasome inhibitors MG132 or PI1, a putative polyubiquitinated form of EMG1 accumulated in the nuclear fraction of fibroblast cells, indicated by the appearance of a higher molecular weight band. This suggests that EMG1 is normally degraded by the proteasome pathway, although further experiments are necessary to confirm that EMG1 is tagged via lysine 48 polyubiquitin chains and is thus targeted to the proteasome. In a study of Diamond-Blackfan anemia, proteasome inhibition increased the levels of mutated RPS19 protein[168]. However, proteasome inhibition did not restore D86G EMG1 to normal levels, indicating that a proportion of the D86G EMG1 may also aggregate. Degradation by the proteasome is in competition with aggregation for misfolded substrates[148] and consequently the proteasome must degrade misfolded proteins before they have a chance to aggregate. Therefore, it is possible that a large proportion of D86G EMG1 in fibroblasts aggregates following proteasome inhibition, rendering it undetectable in the nuclear fraction and disrupting EMG1 turnover.

To evaluate whether the D86G substitution in EMG1 altered its ability to bind to protein partners, immunoprecipitation experiments were performed with NOP14 in HeLa cells. A weak interaction between NOP14 and wild type EMG1 was observed, in agreement with results seen in yeast[113,161]. However, NOP14 binding with D86G EMG1 was not detected, although it was impossible to tell if the binding was inhibited by the structural change in D86G EMG1 or simply because there was not enough of the mutated protein in the cell to

perform an immunoprecipitation experiment. Immunofluorescence showed co-localization of NOP14 and EMG1 even in the presence of the D86G substitution. Therefore, a difference in sub-cellular localization of EMG1 and NOP14 does not appear to be a contributing factor to the failed D86G EMG1 immunoprecipitation. Thus, either D86G EMG1 could not physically interact with its usual protein partner NOP14 due to structural changes, or was present in insufficient quantities to detect an interaction. Although only NOP14 was examined, it is likely that binding with other small subunit processome components would be affected as well, including Mpp10 and the U3 snoRNA, known partners in yeast[59].

The fact that EMG1 localization in HeLa cells was dynamic adds a temporal aspect to EMG1 activities: it must be in the right place, at the right time, to function correctly. In HeLa cells transiently expressing FLAG-tagged NOP14, the nucleolar localization of EMG1 depended on the concentrated presence of NOP14 in nucleoli. Diffuse NOP14 resulted in exclusion of EMG1 from the nucleoli, indicating that EMG1 transportation into the nucleolus may be dependent upon NOP14 shuttling. EMG1 and NOP14 co-localized only in a subset of HeLa cells, and may indicate multiple roles for EMG1 in these cells, depending on its localization. NOP14 associates with the nuclear pore complex[165,166], and ablation of Nop14 in yeast results in the cytoplasmic accumulation of Emg1, rather than the nucleoplasmic accumulation seen here[113]. However, endogenous EMG1 and NOP14 in fibroblast cells did not show the same dynamic localization, with both EMG1 and NOP14 remaining nucleolar in all cells examined. Despite the static localization, EMG1 still co-localized with only a small proportion of NOP14, with the

majority of NOP14 being found in the cytoplasm, in agreement with the weak association seen in the immunoprecipitation assays. The differences in localization between HeLa cells, fibroblasts, and yeast warrants further study, as it may be an indication of distinct modes of regulation in different cell types. Ribosomopathies often display tissue-specific effects, the cause of which remains unclear. Cell-type specific sub-cellular localization and regulation of ribosome biogenesis proteins could play a part in the sensitivity of certain tissues to defects in ribosome production.

In summary, our results indicate that the D86G EMG1 protein which is responsible for BCS is unstable, causing a rapid reduction of soluble EMG1 levels following translation. Both wild type and D86G EMG1 are likely degraded by the proteasome pathway, but exposure of hydrophobic residues in D86G EMG1 also results in its aggregation. The D86G substitution does not, however, alter the sub-cellular localization of EMG1. The reduction in stable, soluble EMG1 protein is not only detrimental to its own function, it also likely alters the association of EMG1 with its protein partners in the small subunit processome. By extension, this may lead to a reduction in ribosome levels in the cells of BCS patients, negatively affecting cell proliferation during development.

Chapter 5 : Assessment of EMG1 function in ribosome biogenesis in BCS cells

Acknowledgements: Protein synthesis rate experiments were performed by Nehal Patel and ribosome subunit experiments were performed by Dr. Richard Hemming. Cell cycle analysis was performed in a University of Manitoba core facility by Dr. Monroe Chan.

5.1 Introduction

Ribosome synthesis and assembly is an extremely complex and energy-intensive cellular process. In proliferating yeast cells, roughly half of all RNA polymerase II transcription initiation events are on ribosomal protein genes, and approximately 80% of the total RNA in the cell is ribosomal[179]. A mature eukaryotic ribosome is composed of the small subunit, containing the 18S rRNA and 33 ribosomal proteins, and the large subunit, containing the 28S, 5.8S and 5S rRNAs and 49 ribosomal proteins. The 18S, 5.8S and 28S rRNAs are all expressed as a single polycistronic 45S precursor transcript which is processed to yield the mature forms, while the 5S rRNA is transcribed separately. Biogenesis of a fully functioning ribosome involves the coordinated efforts of all three RNA polymerases, roughly 200 assembly proteins, and 75 small nucleolar RNAs (snoRNAs), which aid in the maturation and assembly process[20-22]. These factors modify nascent rRNA concurrent with transcription, including cleavage of the precursor transcript by endo- and exonucleases, 2′-OH ribose methylation, pseudouridylation, base methylation, and assembly with ribosomal proteins. These modifications are thought to optimize rRNA folding, stability and function. Understandably, ribosome biogenesis is tightly regulated, and the TP53 pathway is a powerful checkpoint for ribosome assembly. In a healthy cell, the E3-ubiquitin ligase MDM2 binds and ubiquitylates TP53, targeting it for degradation to maintain low levels of the protein. In the event of aberrant ribosome biogenesis, free ribosomal proteins which are unable to assemble into mature ribosomes bind and segregate MDM2, leading to the accumulation of TP53[79,80,82,180]. Subsequent cell-cycle arrest or apoptosis ensures that only cells with intact ribosome biogenesis machinery can proliferate.

In yeast, the precursor 45S rRNA is modified co-transcriptionally by the small subunit processome, a complex rivalling the small subunit itself in size, which contains the U3 snoRNA and at least 40 proteins[59,60]. Emg1 in yeast is associated with the small subunit processome[59] and catalyses the transfer of a methyl group from a SAM donor to the hyper-modified m1acp3-ψ at position 1191 of yeast 18S rRNA (1248 in human 18S rRNA)[72,120]. Other roles for Emg1 during ribosome biogenesis have been proposed, including recruitment of the ribosomal small subunit protein RPS19 to the maturing ribosome, and removal of snR57, the snoRNA component of the snoRNP responsible for the 2′-OH ribose methylation of G1570 in yeast 18S rRNA[115]. Recently, it was posited that Emg1 may act as a chaperone in 18S rRNA folding[121]. As *Saccharomyces cerevisiae* and human EMG1 are 51% identical at the protein level, and human EMG1 complements yeast Emg1 deficiency, it is probable that they have similar functions[114].

Most small subunit processome proteins are essential for cell survival[59,60,73], and depletion causes specific ribosome biogenesis deficiencies. In yeast, depletion of Emg1 results in defects in production of 18S rRNA, and processing of precursor rRNA to mature 18S rRNA is delayed in cells lacking the hypermodification of pseudouridine m1acp3-ψ[59,64]. In addition, yeast cells incorporating the equivalent BCS-causing mutation in Emg1 show reduced growth rates[72]. Assuming similar functions for yeast and human EMG1, we hypothesized that the reduction in EMG1 protein levels seen in BCS patients would result in altered 18S rRNA processing, and therefore deficient biogenesis of the

small ribosomal subunit. The reduction of healthy ribosomes would lead to reduced protein synthesis rates concomitant with slower rates of cell proliferation. Because the levels of ribosomal proteins are tightly regulated and linked to rRNA transcription to ensure optimal ribosome biogenesis rates, we expected that disruption in ribosome protein levels would also lead to TP53 stabilization. A delay in rRNA processing could consequently be sufficient to disturb proper ribosome biogenesis during periods of increased cell proliferation, such as embryonic development. Similar defects have been found in ribosomopathies in humans, a clinically variable group of disorders which are associated with mutations in ribosome proteins and biogenesis factors (reviewed in[89,90,141]).

To study parameters of ribosome biogenesis in BCS patients, we examined dermal fibroblasts taken from skin biopsies, or lymphoblasts generated from peripheral blood leukocytes. In general, two BCS patient cell lines per cell type were compared with two to three age-matched control cell lines. Where quantification of data is shown, the average result of the two BCS patient cell lines is compared with the average of the control cell lines.

5.2 Cell cycle and cell proliferation analyses

It has been shown previously that errors in biogenesis of the small subunit cause a stall in the cell cycle at G1/S, due to a checkpoint which assesses adequate ribosome numbers[59,79,84]. To determine if BCS patient cells had an altered cycle, cell sorting of

propidium iodide-stained cells was performed. This assay separates cells into G1, S, or G2/M populations based on their DNA content. Surprisingly, lymphoblasts from BCS patients had a significantly increased proportion of cells in G2/M (Figure 5.1 B), with no significant difference in any other phase of the cell cycle. On the other hand, cell cycle was not altered in BCS patient fibroblasts when compared to controls (Figure 5.1 A).

A lengthening of G2/M phase would be expected to manifest itself as a slower cell proliferation rate, as cells would take longer to progress through a complete cell cycle. The MTT assay was employed to determine whether proliferation rates of BCS patient cells were affected when compared to controls. This assay employs a yellow tetrazolium salt substrate which is reduced to a purple insoluble formazan complex by the mitochondria. The amount of colour produced is therefore directly proportional to the number of metabolically active cells. In agreement with the cell cycle analysis, we found no significant difference in the proliferation rates over a 72-hour period between unaffected control fibroblasts and BCS patient fibroblasts (Figure 5.1 C). The mean doubling time for control fibroblasts was 48.81 hours, and 48.10 hours in BCS fibroblasts. However, the BCS patient lymphoblasts showed significantly slower proliferation rates relative to control lymphoblasts (Figure 5.1 D), with a mean doubling time of 52.25 hours, compared to a mean of 38.14 hours for control lymphoblasts. To ensure that the MTT assay results in lymphoblasts were due to reduced cell proliferation and not changes in cell metabolism, cells were counted at twelve hour intervals, using trypan blue to exclude dead or dying cells (Figure 5.1 G). BCS lymphoblasts proliferated significantly slower than control cells, confirming the MTT assay results. These data

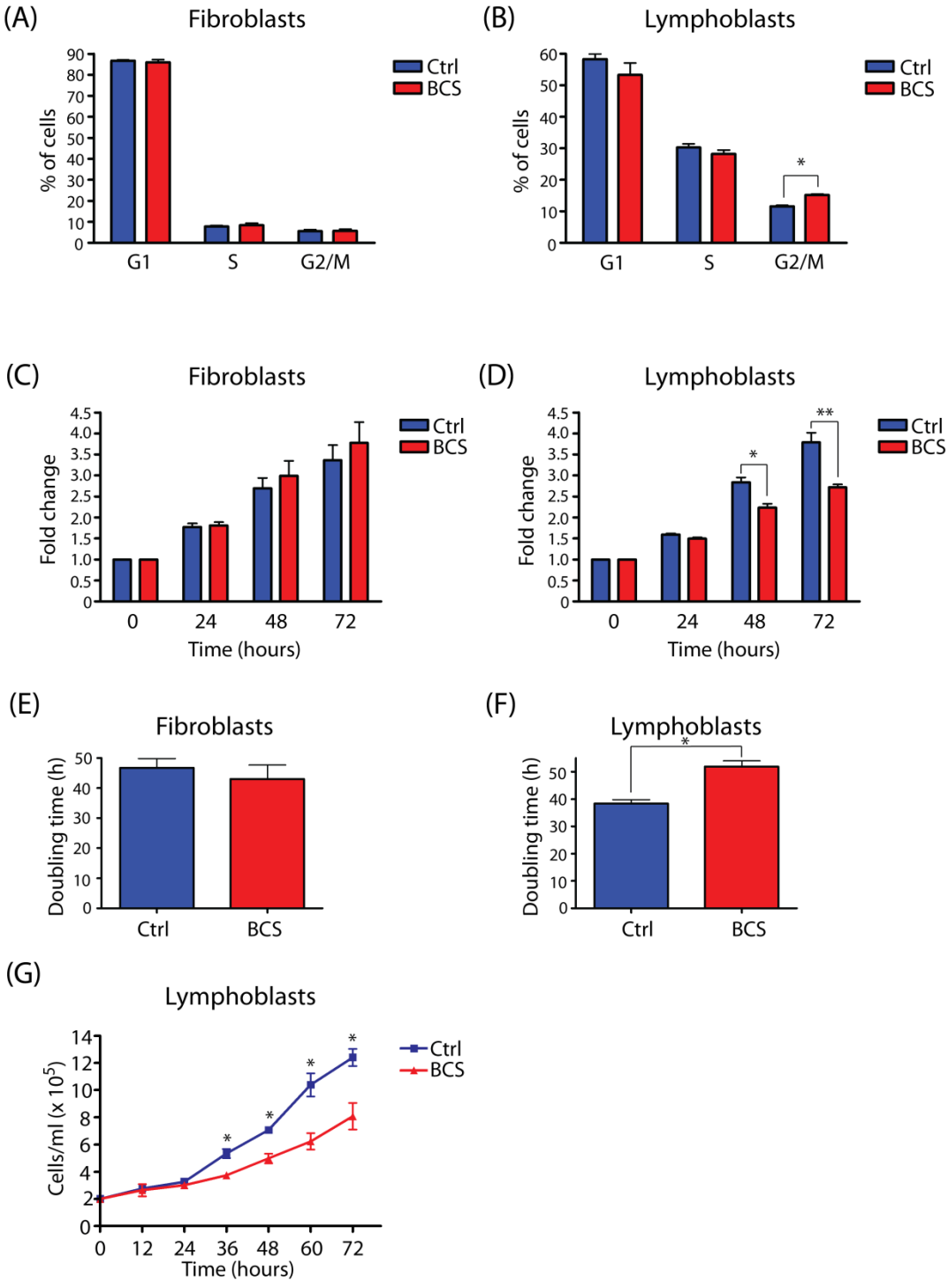

(A)
Fibroblasts
% of cells
Ctrl
BCS
G1
S
G2/M
(B)
Lymphoblasts
% of cells
Ctrl
BCS
*
G1
S
G2/M
(C)
Fibroblasts
Fold change
Ctrl
BCS
0
24
48
72
Time (hours)
(D)
Lymphoblasts
Fold change
Ctrl
BCS
*
**
0
24
48
72
Time (hours)
(E)
Fibroblasts
Doubling time (h)
Ctrl
BCS
(F)
Lymphoblasts
*
Doubling time (h)
Ctrl
BCS
(G)
Lymphoblasts
Cells/ml (x 10^5)
Ctrl
BCS
*
0
12
24
36
48
60
72
Time (hours)

Figure 5.1 Cell cycle analysis and cell proliferation rates of control and BCS patient cells.

(A, B) Cell cycle analysis of fibroblasts (A) and lymphoblasts (B). Cells were treated with propidium iodide and sorted based on DNA content. In the lymphoblasts, the percentage of cells in G2 phase was significantly higher in BCS cells than in control cells (*p = 0.00185). The mean of three experiments and SEM are shown. **(C, D)** Cell proliferation rates, evaluated using the MTT assay, in fibroblasts (C) and lymphoblasts (D). At 24 hour intervals, cells seeded in a 96-well plate were incubated with the MTT reagent for four hours, lysed, and the absorbance was read at 570 nm. The fold change relative to time 0 was plotted for each time point. The mean of three experiments, performed in triplicate, is shown for the fibroblasts and four experiments for the lymphoblasts, expressed as fold change from time 0. BCS lymphoblasts were significantly lower than controls at 48 hours (*p < 0.01), and at 72 hours (**p < 0.001). Error bars show SEM. **(E, F)** Doubling time of fibroblasts (E) and lymphoblasts (F). Doubling time was calculated from the results of the MTT assay performed in (C) and (D). In the lymphoblasts, doubling time was significantly longer (*p<0.0001). **(G)** Lymphoblast cell concentrations in cells per millilitre, counted at 12 hour intervals. BCS lymphoblasts numbers were similar to control numbers until the 36 hour time point, and were significantly lower than controls from 36 to 72 hours (*p<0.01). The mean of two experiments, performed in duplicate, and SEM are shown.

suggest that a stall at the G2/M checkpoint results in a longer cell cycle in BCS patient lymphoblasts, but that patient fibroblasts are not appreciably affected.

5.3 Determination of 60S/40S subunit 28S/18S rRNA ratios

5.3.1 Steady-state levels of ribosomal subunits and rRNA

In order to determine if the deficiency of EMG1 protein seen in BCS patient cells was causing an alteration in the synthesis of the small ribosomal subunit, the relative amounts of the intact small and large ribosomal subunits were evaluated. Ribosomes were isolated from patient and unaffected control lymphoblasts and fibroblasts, and separated in a sucrose gradient (Figure 5.2 A, B). The ratio of the large subunit to the small subunit was compared between BCS patients and controls, and no significant difference was found in either the lymphoblast or fibroblast lines (Figure 5.2 C, D). Similar results were found in fibroblasts in the presence of higher concentrations of Mg^{2+}, which allows the isolation of 80S ribosomes and polysomes (not shown).

Previous studies have shown that defects in precursor rRNA processing result in a reduction in ribosomal protein levels, and that a reduction in proteins from one subunit can affect the other[74]. Therefore, the levels of proteins RPL26 and RPL7 from the large subunit, and RPS19 and RPS6 from the small subunit, were examined by immunoblot in BCS patient and unaffected control fibroblasts (Figure 5.2 E). There were no consistent differences in protein levels from either subunit. Taken together with the sucrose gradient

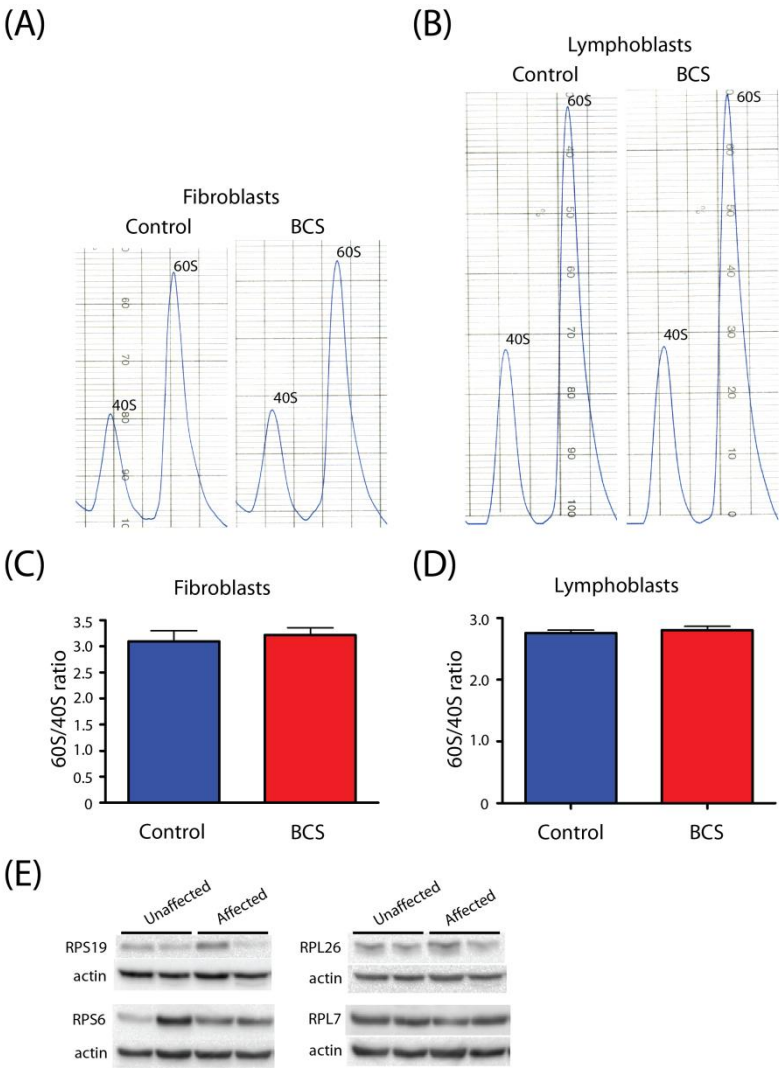


Figure 5.2 Ribosome levels at steady state in control and BCS patient cells.

(A, B) Small (40S) and large (60S) ribosome subunits from BCS and unaffected control cell lysates were separated by centrifugation on a sucrose gradient and detected by absorbance at 254 nm. Examples of the resultant tracings are shown for fibroblasts (A) and lymphoblasts (B). **(C, D)** The area under the curve was determined for individual peaks, and the ratio of large to small subunit was calculated. No significant difference between unaffected control and BCS- affected cells was found in fibroblasts (C) or lymphoblasts (D). The mean of six individual experiments and SEM are shown. **(E)** Immunoblots detecting ribosomal proteins from the small (RPS) and large (RPL) subunits of the ribosome showed no consistent differences in the proteins between two unaffected control and two BCS-affected fibroblast lines. Actin was used as a loading control in each case.

ribosome subunit experiments, these data indicate that at steady state, the small subunit of the ribosome is at normal levels in BCS patient cells relative to the large subunit.

The relative levels of steady-state rRNA found in cells from BCS patients were also assessed. Total RNA was isolated from either fibroblasts or lymphoblasts from BCS patients or normal controls, and was analyzed using the RNA 6000 Nano LabChip Kit in the Agilent 2100 Bioanalyzer. Examples of the resultant electropherograms are shown in Figure 5.3 A and B. The area under the curve for each peak was calculated, and total amounts of 18S rRNA were evaluated relative to 28S rRNA. In each case, there were no significant differences in the ratio of 28S rRNA to 18S rRNA between any of the samples (Figure 5.3 C, D).

5.3.2 Processing of 18S rRNA

Although the steady-state levels or rRNA remain unaffected, it is possible that the processing rate of precursor 45S to mature 28S and 18S rRNA is delayed in BCS patient cells. This was previously shown in the ribosomopathy Shwachman-Diamond syndrome, where sucrose gradient analysis showed similar steady-state levels of ribosomal subunits, yet the formation of mature rRNA was significantly slower in SDS patient fibroblasts[37]. To evaluate rRNA processing rates, pulse-chase analysis was employed using either $^{32}P_i$ or ^{3}H-methyl methionine to label nascent rRNA. Using this technique, it is possible to follow the transformation of precursor 45S rRNA to its mature 28S and 18S rRNA forms over time. Following a pulse of radioactive label, the cells were chased in cold media,

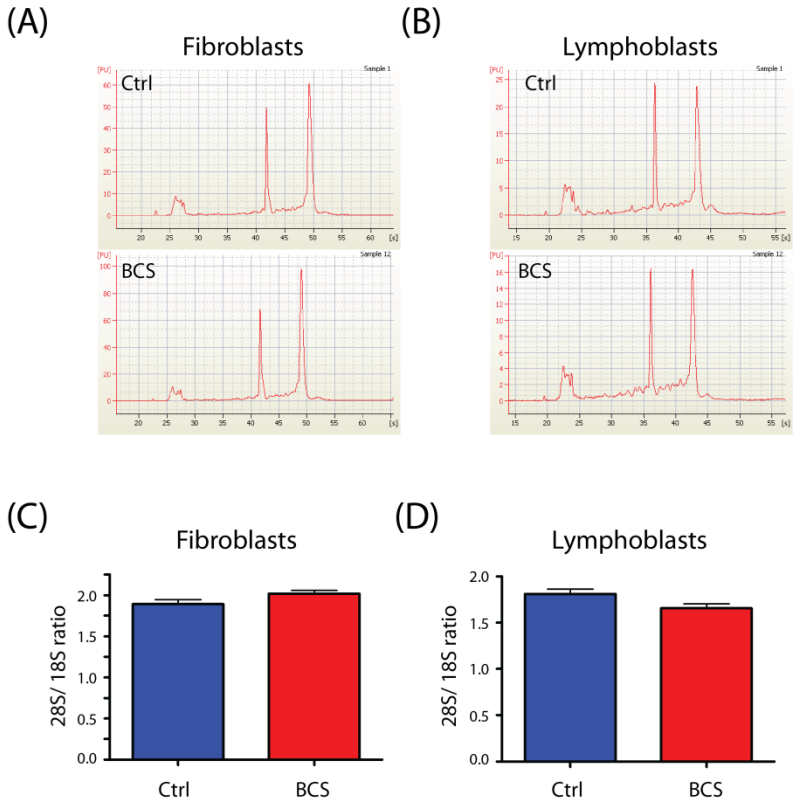


Figure 5.3 Ribosomal RNA levels at steady state.

(A, B) Ribosomal RNA electropherograms. Total rRNA was isolated from unaffected control and affected fibroblasts (A) and lymphoblasts (B), and separated by capillary chromatography using the 6000 RNA Nano kit in an Agilent Bioanalyzer. The resulting traces show the 18S and 28S peaks, as well as a smaller peak which encompasses the 5S, 5.8S, and tRNA. **(C, D)** 28S/18S ratios for fibroblasts (C) and lymphoblasts (D). The area of each peak was calculated by the Agilent software, and the ratios of 28S to 18S rRNA were calculated. The mean of three individual experiments, performed in triplicate, and SEM are shown. No significant difference was found between the unaffected and affected cells, indicating that 18S rRNA levels are normal at steady state in patient cells.

and samples were collected at regular intervals. Ribosomal RNA was isolated and separated using gel electrophoresis, then visualized using autoradiography. Using Image Lab software to quantify band intensity, there was a consistent reduction in labelling of 18S rRNA in BCS patient cells, while 28S rRNA was unaffected (Figure 5.4). Although this difference was only statistically significant at a specific point in processing (see time 20 minutes of 18S rRNA in 5.4 B), it was consistent through several independent experiments, employing both the $^{32}P_i$ and the ^{3}H methyl-methionine labels, and in both lymphoblasts and fibroblasts. This, together with the fact that steady state rRNA levels were normal, indicates that overall rRNA production occurs at a normal rate in BCS cells, and that processing of 18S rRNA is instead delayed. This delay does not detectably affect steady-state levels of 18S or 28S rRNA, nor does it alter steady-state levels of the ribosomal subunits.

5.4 Evaluation of protein synthesis rates

A delay in the 18S rRNA processing could result in sub-optimal ribosome function, and perturb overall protein synthesis rates. We therefore examined the rate of protein synthesis in each fibroblast cell line by incorporation of ^{35}S Cys/Met over a four-hour period, calculated as the amount of incorporated radioactivity in CPM per μg of protein. Similar results were found for BCS patient and unaffected control fibroblasts (Figure 5.5), indicating that the delay in 18S rRNA processing does not alter the rate of translation by the ribosome. This is consistent with the conclusions that the delay in 18S rRNA processing is transient, and does not affect steady-state levels of the ribosomal

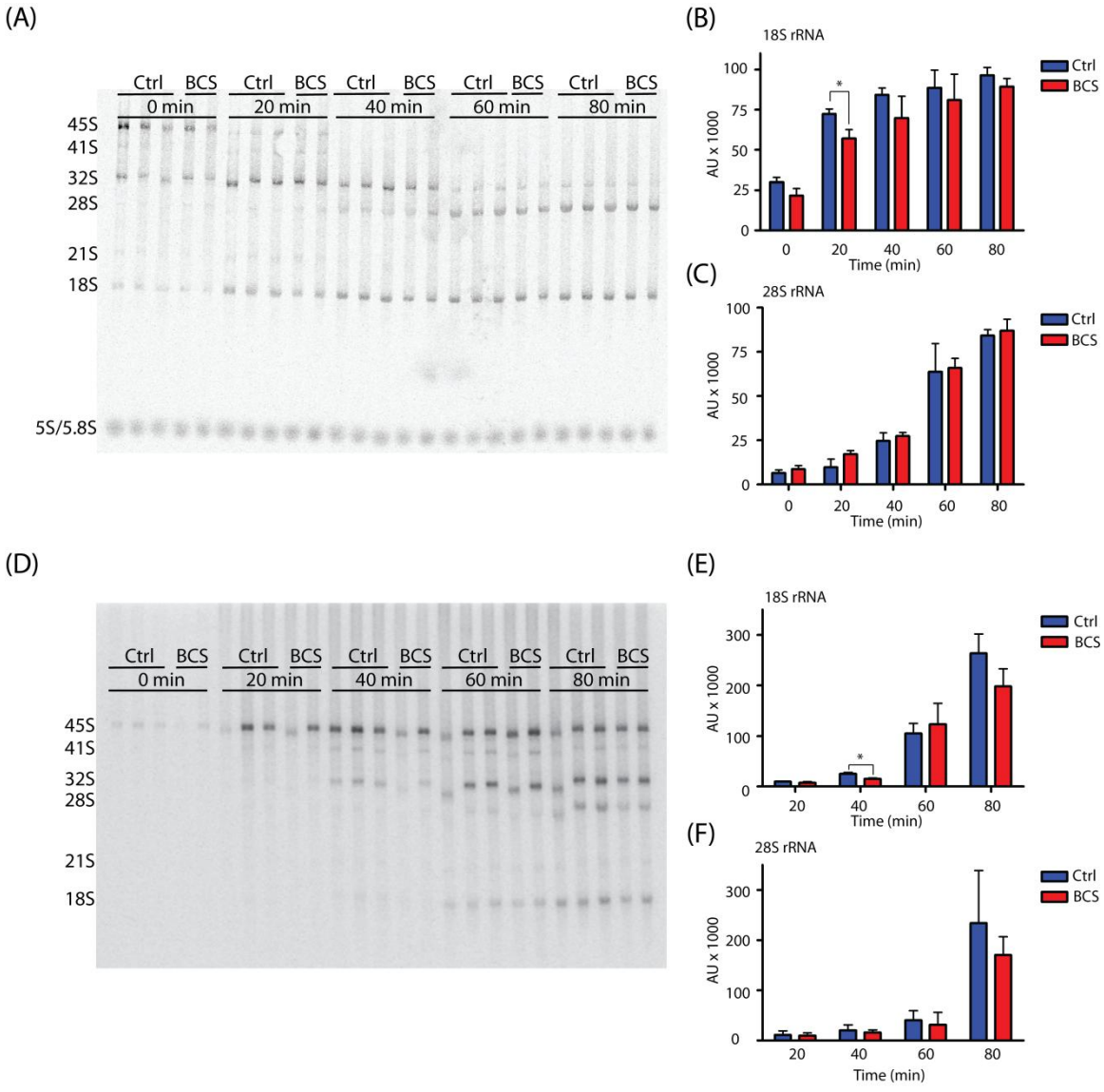

(A)
Ctrl BCS Ctrl BCS Ctrl BCS Ctrl BCS Ctrl BCS
0 min 20 min 40 min 60 min 80 min
45S
41S
32S
28S
21S
18S
5S/5.8S
(B)
18S rRNA
AU x 1000
Time (min)
Ctrl
BCS
(C)
28S rRNA
AU x 1000
Time (min)
Ctrl
BCS
(D)
Ctrl BCS Ctrl BCS Ctrl BCS Ctrl BCS Ctrl BCS
0 min 20 min 40 min 60 min 80 min
45S
41S
32S
28S
21S
18S
(E)
18S rRNA
AU x 1000
Time (min)
Ctrl
BCS
(F)
28S rRNA
AU x 1000
Time (min)
Ctrl
BCS

Figure 5.4 Ribosomal RNA processing.

(A) An example of a ribosomal RNA processing pulse-chase experiment in fibroblasts. Nascent rRNA was metabolically labelled using (methyl-^{3}H) methionine for thirty minutes, and samples were isolated at twenty minute intervals to follow the processing rate of the large 45S rRNA precursor to the mature 18S and 28S species. Equal counts were separated on a 0.7% agarose gel, and transferred to a positively-charged nylon membrane. The membrane was exposed to a phosphor storage screen, which was scanned using a phosphorimager. The precursor 45S rRNA, the 28S and 18S mature species, and the intermediates are indicated on the left side of the diagram. **(B** and **C)** The intensity of each band in (A) was quantified using Image Lab software. The graphs indicate the mean density in arbitrary units of the 18S **(B)** and 28S rRNA **(C)** bands at each time point, and error bars represent standard deviation. For this experiment, the average of three control cell lines and two BCS lines are shown. The 18S rRNA was significantly reduced in BCS cells at the 20 minute time point (*p = 0.0287), indicating a delay in the 18S rRNA processing in BCS cells which does not affect the 28S rRNA. **(D)** An example of an rRNA processing pulse-chase experiment in lymphoblasts. The experiment was performed essentially as in (A), except the rRNA was labelled with $^{32}P_i$ for a twenty minute pulse. Following gel electrophoresis to separate the RNA species, the gel was dried and directly exposed to a phosphor storage screen. **(E)** and **(F)** show the intensity of each band. The 18S rRNA was significantly reduced in BCS cells at the 40 minute time point (*p = 0.0292).

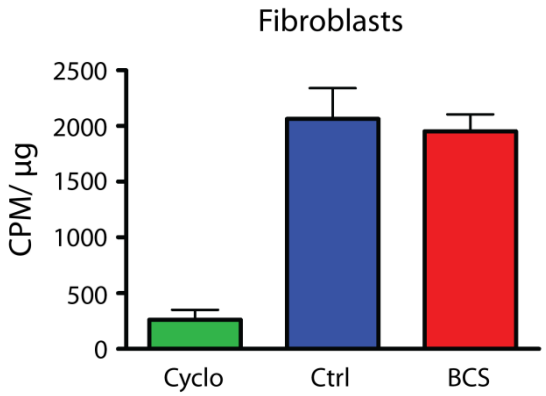


Figure 5.5 Protein synthesis rates.

Fibroblasts were incubated with ^{35}S-cysteine/methionine to label newly synthesized protein, while cycloheximide was included with control cells to inhibit protein synthesis as a negative control. Proteins were TCA-precipitated from cell lysates, filtered, and washed. Labelled protein was detected using a liquid scintillation counter, protein concentration was determined, and counts per microgram of protein was calculated. No statistical difference was found. The mean of three experiments and SEM are shown.

subunits. Given that the ribosomal RNA and subunit levels were similarly unaffected in BCS patient lymphoblasts, the protein synthesis rate experiments were performed on fibroblasts only.

5.5 Assessment of TP53 stabilization

Because dysregulation of ribosome biogenesis has been shown to lead to TP53 stabilization and cell cycle arrest[79-82], levels of TP53 protein in BCS patient cells and unaffected controls were compared using immunoblot analysis. TP53 protein levels varied between cell lines, with no clear difference between BCS patient cells and unaffected controls (Figure 5.6). It is possible that the variable levels of TP53 reflected differences in cell cycle of the different cell populations that were used.

5.6 Discussion

In yeast, a reduction in Emg1 levels causes a decrease in 18S rRNA levels with a corresponding accumulation of its immediate precursor in the rRNA processing pathway[113]. This lack of 18S rRNA leads to an imbalance between the large and small ribosomal subunits which in turn causes a reduction in protein synthesis and cell proliferation rates[114]. Thus, we expected that since EMG1 levels are severely reduced in BCS patient cells, we would see a corresponding decrease in cell proliferation and protein synthesis rates due to a lack of small ribosomal subunits. Instead, we found an increase in the proportion of BCS lymphoblasts in G2/M, accompanied by slower proliferation rates, while neither was affected in BCS fibroblasts. A consistent, transient delay in 18S rRNA

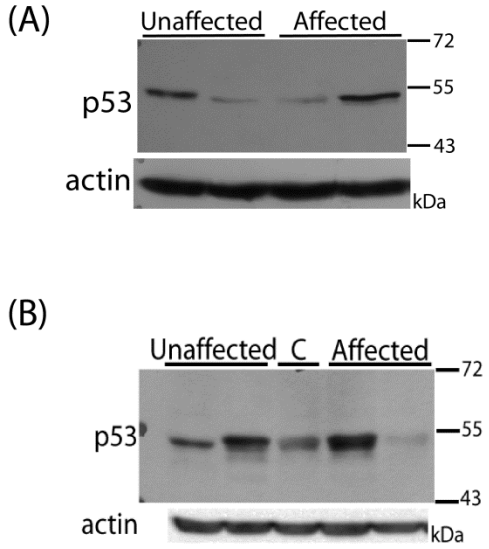


Figure 5.6 TP53 levels in control and patient cells.

Levels of TP53 protein were evaluated in **(A)** fibroblasts and **(B)** lymphoblasts in unaffected controls and BCS-affected patients by immunoblot. Equal amounts of protein were separated in a 10% gel by SDS-PAGE, transferred to a nitrocellulose membrane, and probed using an anti-p53 antibody. Levels of TP53 were found to be variable, with no clear differences between controls and BCS cells. Each immunoblot is representative of three independent experiments. C, carrier of *EMG1* A>G mutation.

processing in both fibroblasts and lymphoblasts did not result in an alteration of steady-state levels of 18S rRNA or small ribosomal subunits. Protein synthesis rates were also unaffected in BCS patient fibroblasts.

The reduced proliferation rate, observed only in the BCS lymphoblasts, can be explained by a lengthening of the cell cycle in G2/M, which was also found only in lymphoblasts. While rRNA processing defects typically cause G1 arrest[59,79,84], an increase in the percentage of cells in G2/M phase has been reported in cells depleted of dyskerin, the protein responsible for the X-linked form of the ribosomopathy dyskeratosis congenita[181,182]. In addition, changes in RPL13a expression can induce cell cycle arrest at G2/M[183], mutations in RNAse MRP causing cartilage hair hypoplasia lead to cell cycle delay in anaphase[184], and a mutation in *pescadillo* in zebrafish, a gene implicated in ribosome biogenesis, disrupts oligodendrocyte formation due to failure to progress from S phase to mitosis[86]. The stall at G1/S seen in many cases of ribosome biogenesis disruption has been shown to be the result of a checkpoint which verifies that adequate numbers of ribosomes are present before proceeding with the cell cycle. We found that steady-state levels of the ribosome subunits are normal in BCS cells, which may explain why they are able to progress through the G1/S checkpoint. It is difficult to say whether this is the case in the examples of cell cycle delay at G2/M as described above, since levels of ribosomal subunits were not evaluated in those studies.

Intriguingly, in X-linked dyskeratosis congenita, ribosomes lacking pseudouridine are deficient in IRES-dependent translational initiation due to failure of the ribosome to bind to the IRES sequence[98,185]. In general, cap-dependent translation is inhibited at the beginning of mitosis in favour of the translation of mRNAs containing IRES elements, including one isoform of cyclin-dependent kinase 11, which is necessary for progression through M phase[186,187]. In X-linked dyskeratosis congenita, it is thought that pseudouridine-deficient ribosomes are unable to recognize IRES elements, causing a stall at M phase due to decreased translation of cell-cycle related proteins. In the same vein, it is conceivable that changes in rRNA methylation or folding occasioned by the D86G substitution in EMG1 alter the specificity of the ribosome in such a way that it can no longer recognize or bind IRES elements efficiently. Further support for this hypothesis comes from the fact that IRES-dependent mRNA translation is down-regulated in mice expressing mutated RPS19, the protein most commonly associated with Diamond-Blackfan anemia[188]. It has been postulated that EMG1 promotes RPS19 binding to the nascent ribosome[115,121]. Alterations in IRES-mediated translation could lead to the delay or failure of progression through G2/M seen in BCS patient lymphoblasts. A standard assay employing a dicistronic *Renilla* luciferase/ firefly luciferase reporter, linked by a known IRES element, could be used to determine if ribosomes from BCS patient cells are deficient in IRES-dependent translation[68,189].

To determine if a defect in ribosome biogenesis was causing the delay at G2/M and the concomitant reduction in cell proliferation rates, processing of precursor 45S rRNA to its mature forms was examined. We detected a slight but reproducible decrease in the

processing rate of 18S rRNA in both fibroblasts and lymphoblasts, evidence that a defect in ribosome biogenesis underlies Bowen-Conradi syndrome. As EMG1 has more than one proposed role in 18S rRNA processing, it is difficult to determine the exact nature of the defect. Given that EMG1 is a methyltransferase, the use of ^{3}H-methyl methionine as a label makes it impossible to determine if the reduction in detectable radioactivity is the result of reduced rRNA production, or if rRNA is produced at normal levels but is under-methylated. Several lines of evidence point to the former scenario. First, yeast Emg1 bearing the equivalent mutation of the BCS protein conserves its rRNA methyltransferase activity in an *in vitro* methylation assay[72] despite causing a reduction in growth rates. Second, Emg1 mutants unable to bind the methyl donor SAM maintain normal ribosome biogenesis[117]. Finally, similar results were observed using inorganic phosphate as a radiolabel, which incorporates into the RNA backbone and does not affect rRNA methylation. It is therefore likely that the reduction in EMG1 causes a delay in 45S to 18S rRNA processing via its role in RPS19 recruitment or rRNA folding rather than its catalytic activity. However, until methylation of 18S rRNA at ψ1248 is measured *in vivo*, it should still be considered a possibility that 18S rRNA is under-methylated in patient cells. Despite the fact that a delay in processing was observed in both fibroblasts and lymphoblasts, steady-state ratios of 60S to 40S ribosomal subunits were unaffected in both cell types, and consequently protein synthesis rates were also unchanged. This is similar to what was found in the ribosomopathy Shwachman-Diamond syndrome, where a reduction in rRNA production in patient cells was not associated with an imbalance in 60S, 40S, or 80S ribosomal subunit levels at steady state[37]. Given the transient nature of the 18S rRNA defect observed in BCS patient cells, it is unlikely that a 60S/40S subunit

ratio imbalance would be detected using sucrose gradient analysis of ribosome subunits at steady state, as we did here.

Underlining the importance of the TP53 checkpoint in proper ribosome biogenesis, previous studies have shown that removal of *Trp53* can rescue the defect in certain mouse models of ribosomopathies[105,124]. However, crossing *Trp53*$^{-/-}$ mice with *Emg1*$^{+/-}$ mice failed to rescue the pre-implantation growth arrest phenotype of EMG1 knockouts[122]. Taken together with our TP53 immunoblot studies which did not show stabilization of TP53 in BCS patient cells, we concluded that the EMG1 defect likely does not activate the TP53 pathway. Although the link between the TP53 pathway and ribosome biogenesis has been exploited in animal models of ribosome biogenesis, the contribution of TP53 to the regulation of ribosome biogenesis is far from clear, as recent studies have shown variable responses to TP53 inhibition or knockout. In *Rpl38*, *Rps19*, and *Rpl11* mutant mice, the phenotype is refractory to TRP53 inhibition[97,188]. In a zebrafish model of Shwachman-Diamond syndrome, ablation of the TP53 checkpoint in *sbds* knockdown fish ameliorates the skeletal phenotype and improves the overall health of the embryos[109]. However, the spleen hypoplasia is not rescued, suggesting that the phenotype is only partially mediated by the TP53 pathway. Similarly, TP53 knockdown in a Diamond-Blackfan anemia model rescued the morphological abnormalities, but did not alter the erythopoietic defect[110,111]. The selective contribution of the TP53 checkpoint to individual ribosomopathies, and even different tissues within a single ribosomopathy, is just beginning to be appreciated. Therefore, stabilization of TP53 is still possible in other BCS patient tissues or cell types.

The fact that lymphoblasts were affected, while fibroblasts were not, supports the hypothesis that proper levels of EMG1 are critical in rapidly proliferating cells. Slower-growing cells, such as fibroblasts, have less of a demand for ribosomes and therefore are not as strongly affected by defects in ribosome biogenesis; thus the delay in rRNA processing seen in fibroblasts does not detectably affect cell cycle or proliferation rates. By extension, the rapidly proliferating cells of the embryo could be severely affected by a delay in 18S rRNA processing, resulting in the developmental delay seen in BCS patients. The different results in lymphoblasts and fibroblasts may also be a reflection of the tissue specificity which is often seen in ribosome biogenesis disorders. In BCS, the phenotype primarily affects the central nervous system, making it likely that EMG1 deficiency has a greater impact on the rapidly developing brain than it does on other tissues. Expression profiling via microarray could be used to evaluate differences in gene expression between the fibroblasts and lymphoblasts, and perhaps indicate genes which compensate for the delay in 18S rRNA processing in fibroblasts. Taken together, these results confirm conservation of EMG1 function in yeast and human cells, and link BCS with the growing number of ribosomopathies.

.

Chapter 6 : Establishment and characterization of a mouse model of BCS

Acknowledgements: Nehal Patel constructed the EMG1 targeting vector. Dr. Hao Ding and Xiaoli Wu performed the ES cell manipulations, created the chimeras, and performed BrdU injections.

6.1 Introduction

The mutation causing BCS produces profound effects on fetal development long before birth and until death at an early age. While it has been established that mutations in ribosomal proteins and ribosomal biogenesis factors can differentially affect multiple tissues[89,95,97], autopsies on infants who have died from BCS are rare, and the majority of investigations have been performed on patient cells obtained from peripheral blood or skin biopsies. In order to study the effect of the EMG1 mutation on multiple tissues in organism development and survival, we proposed to create a mouse model of BCS. Human and mouse EMG1 are 89% identical at the protein level, and both human and mouse EMG1 complement a yeast deficiency[113,114] making it very likely that the homologues share similar functions. In addition, the D86 substituted in BCS and the regions surrounding it are conserved in the mouse EMG1 protein[147]. Mouse models of disease have been crucial in deciphering the molecular causes of several ribosomopathies, and have aided in the elucidation of the normal functions of the mutated proteins involved[94,190-192]. For example, a murine model of the craniofacial disorder Treacher Collins syndrome (TCS) showed that *Tcof1*, which is mutated in TCS, is not only a regulator of ribosome biogenesis, but is also required for the proliferation of neural crest cells during development[94,104]. Moreover, the same mouse model was later used to show that the defect in TCS mice could be rescued by disruption of the TRP53 checkpoint[105].

A previous *Emg1* null (*Emg1*$^{-/-}$) mouse produced in collaboration with Dr. Hao Ding replaced exons 2-6 with a SA-IRES-βgeo cassette[122]. The *Emg1* knockout is lethal very early in development while the heterozygous mice show no detectable phenotype,

suggesting that EMG1 is essential for development prior to implantation. In the *Emg1*$^{-/-}$ mice, growth is arrested between embryonic day (E) 2.5 and E3.5, accompanied by increased cell death and reduced cell proliferation. The very early stage of lethality made it impossible to examine tissues or generate cell lines deficient in EMG1, precluding further study. Unlike in the mouse models of Treacher Collins syndrome and other ribosome disorders[95,105,124], crossing *Emg1*$^{+/-}$ mice with TRP53-deficient mice failed to rescue the phenotype, suggesting that the defect in these mice is not mediated by the TRP53 pathway.

To generate a mouse model more closely resembling human BCS, we designed a targeting vector to insert the equivalent A>G mutation into the *Emg1* gene (*Emg1*$^{A/G}$). The inclusion of *loxP* sites flanking exons 2-6 of the gene meant that the allele could be deleted by Cre recombinase, allowing conditional knockout of the gene at different ages or in different tissues depending on the promoter driving Cre expression[193]. Alternatively, Cre could be expressed in cultured cell lines containing the targeted *Emg1* allele, creating an *Emg1*$^{-/G}$ line in which we could study the effects of severely reduced EMG1 levels. We hypothesized that mice homozygous for the mutation (*Emg1*$^{G/G}$) would be small for their gestational age when compared to wild type littermates, possibly with craniofacial abnormalities, and that they would be unable to feed properly, resulting in death shortly after birth. These mice would allow detailed analyses of the defect in all tissues throughout development, in particular the central nervous system which appears to be a major target of BCS in humans. Given the growth retardation in BCS patients, *Emg1*$^{G/G}$

mice permitted us to examine the roles of cell proliferation and cell death in the etiology of BCS.

6.2 EMG1 expression throughout normal mouse development

To better understand the role of EMG1 during development, particularly in the central nervous system, we examined expression of EMG1 in mice from embryonic day 7.5 through to adulthood. We employed a previously generated *Emg1* knockout strain which incorporated an SA-IRES-βgeo cassette in place of exons 2-6[122]. Heterozygous mice therefore expressed the *LacZ* reporter gene under the control of the endogenous *Emg1* promoter. Embryos from timed pregnant heterozygous intercross matings were taken at embryonic days 7.5, 8.5, 9.5, 10.5, and 12.5 (n = 2 pregnant females / time point). Tissues were also collected from pups at postnatal days 7 and 14, and from adult heterozygotes at six weeks of age (n = 2 / time point). In addition to the *LacZ* reporter, EMG1 was also examined in these tissues using a polyclonal anti-EMG1 antibody. To validate the EMG1 antibody for use in immunohistochemistry in mouse tissues, the antibody was first incubated with full-length EMG1 tagged with maltose binding protein. Experiments performed using this competed anti-EMG1 antibody eliminated nuclear staining (Figure 6.1 A), although some background cytoplasmic staining remained. These results, combined with data generated using the *LacZ* reporter, indicated that the EMG1 antibody was appropriate for use in immunohistochemistry. To detect the β-galactosidase enzyme encoded by the *LacZ* reporter gene, tissues were incubated with the substrate X-gal, producing a blue product in the presence of the enzyme, which is visualized as a pink product using dark field

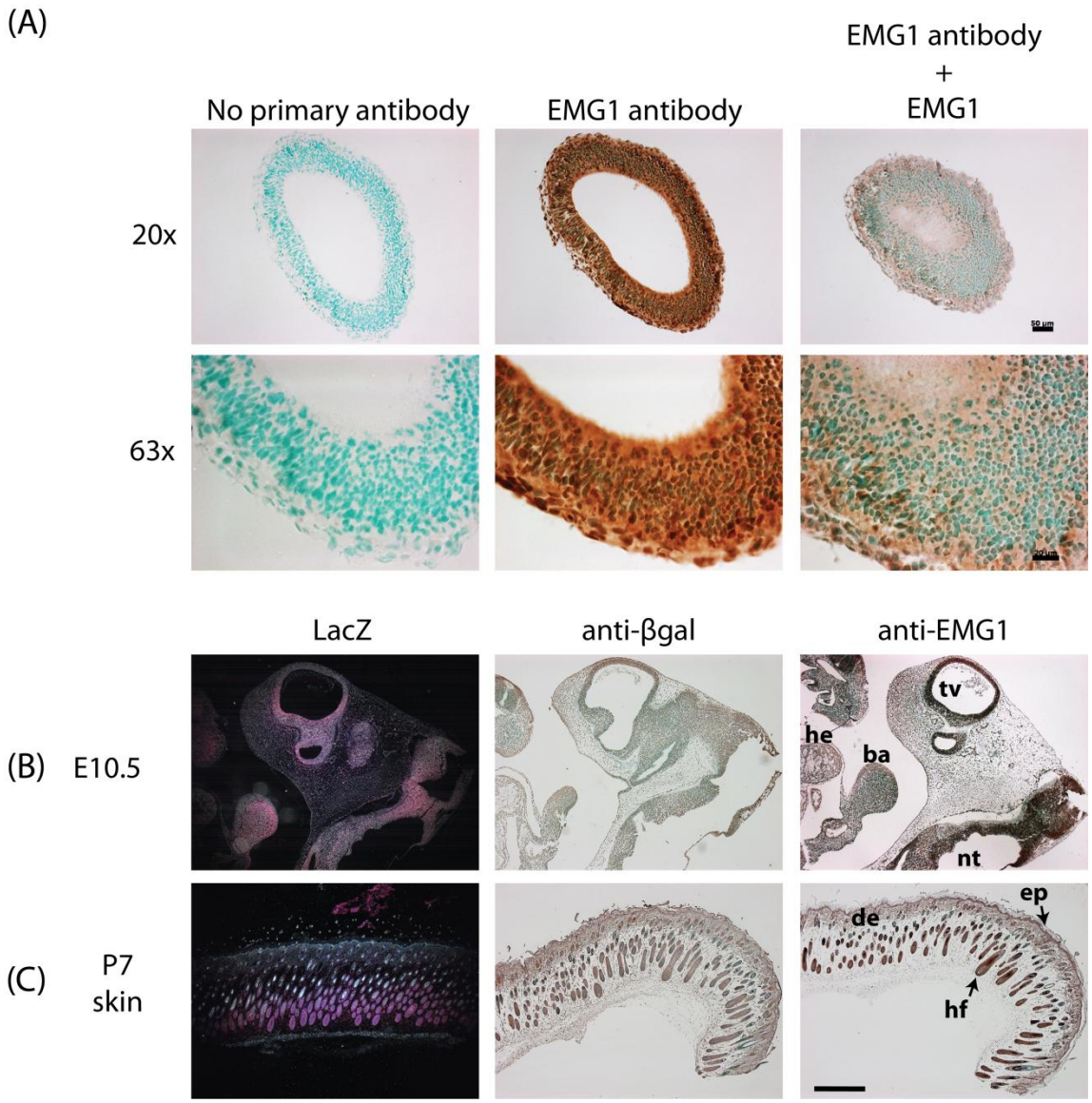


Figure 6.1 EMG1 detection in mouse tissues. (**A**) EMG1 antibody competition assay. EMG1 polyclonal antibody (ProteinTech Group) was incubated at a dilution of 1:500 with purified MBP-tagged EMG1. Uncompeted and competed anti-EMG1 antibodies were then incubated on the same slide with formaldehyde-fixed, paraffin-embedded sections of E10.5 wild type mouse tissue (transverse sections through the top of the head), and detected using the ABC/DAB system. A section not exposed to the primary antibody was used as a negative control, and methyl green was used as counterstain. A comparison of the competed and uncompeted sections revealed specific EMG1 staining in cell nuclei, with some residual background cytoplasmic staining. Scale bar at 20x = 50 µm, and at 63x = 20 µm. (**B-C**) Comparison of β-galactosidase activity and anti-EMG1 antibody. The LacZ reporter gene was detected by X-gal staining (pink) in mid-sagittal sections of E10.5 embryos (**B**) and skin of postnatal day seven pups (**C**). This showed similar areas of expression as detection of the β-galactosidase enzyme by an anti-βgal antibody (1:500 dilution), or detection of EMG1 using an anti-EMG1 antibody (1:500 dilution). Expression was ubiquitous in the E10.5 embryos, with strong staining in the neuroepithelium, and in the hair follicles of the skin. Original magnification, 5x; scale bar = 500 µm. **ba**, branchial arch; **de**, dermis; **ep**, epidermis; **he**, heart; **hf**, hair follicle; **nt**, neural tube; **tv**, telencephalic vesicle.

microscopy. The results were confirmed with an anti-β-galactosidase antibody (Figure 6.1 B, C).

During development, EMG1 expression was strong and widespread in all tissues, though it was strongest in the earliest embryos examined and became more restricted over time (Figure 6.2). It was also found in all tissues at postnatal day 7 (Figure 6.3) and in the adult (Figure 6.4). Areas of strongest expression included eye, skin, intestine, ovary, and testes, in agreement with a previous assessment of EMG1 expression in mouse using *in situ* hybridization[122]. Although expression was ubiquitous as expected for a cell-essential protein, it became more restricted during development. For example in the eye, in later embryos and postnatally, expression became restricted to the ganglion cell layer and the inner nuclear layer in the retina, whereas expression was low or absent in the outer nuclear layer (Figure 6.5 A-D). In the intestine, EMG1 staining appeared as a gradient with highest expression in the crypts, gradually decreasing to the ends of villae (Figure 6.5 E). Similarly, the strongest expression in testes was found in the outer layer of spermatogonia and primary spermatocytes, likely corresponding to cells in mitosis and meiosis, with much lower expression in mature spermatozoa (Figure 6.5 F). In most cells, EMG1 was localized to the nucleus, although strong cytoplasmic staining was also found in the liver of postnatal day 7 mice.

In the brain, EMG1 expression was similarly ubiquitous, and found in most areas at all ages examined. EMG1 expression was strongest and most widespread early in

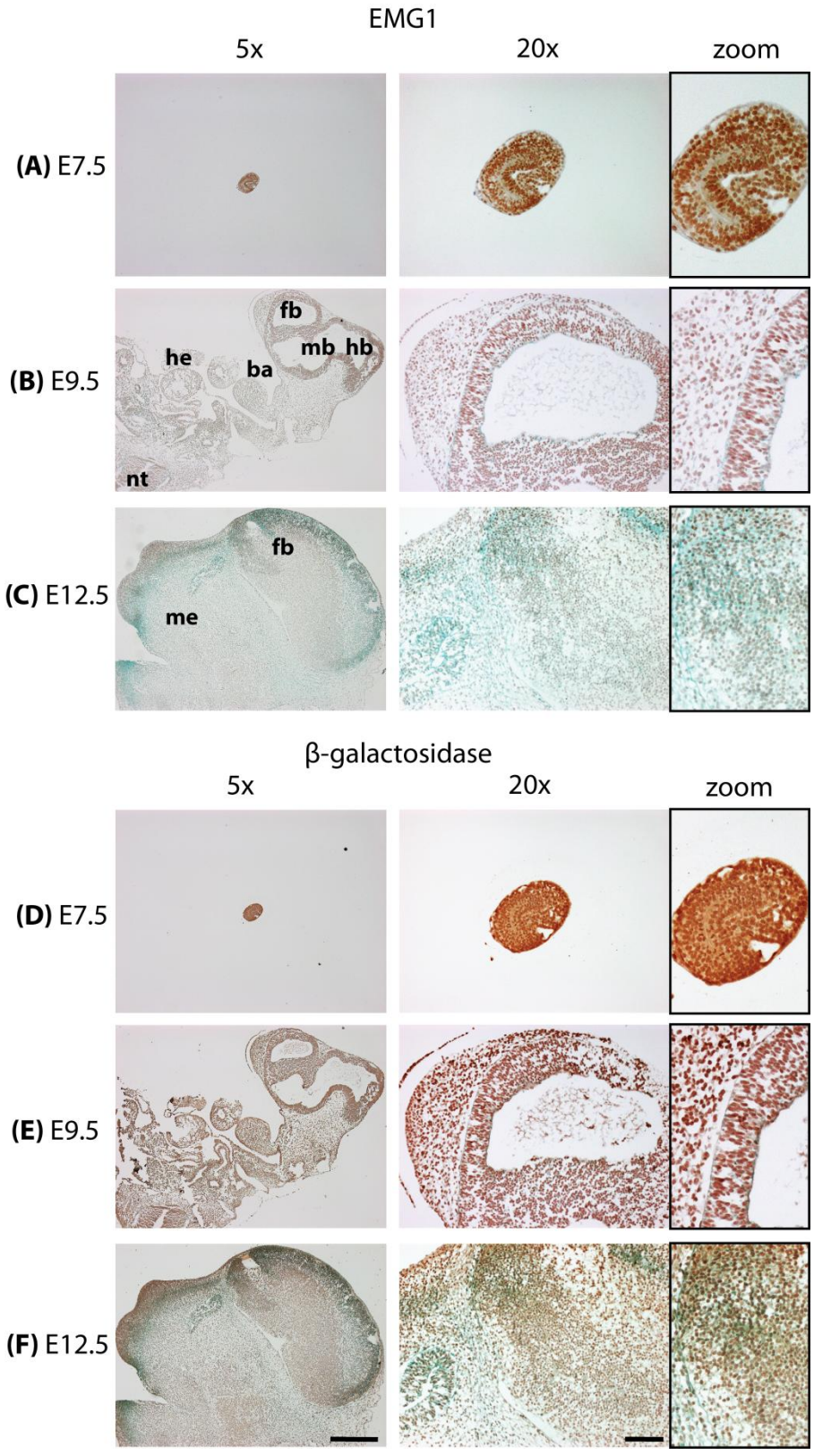
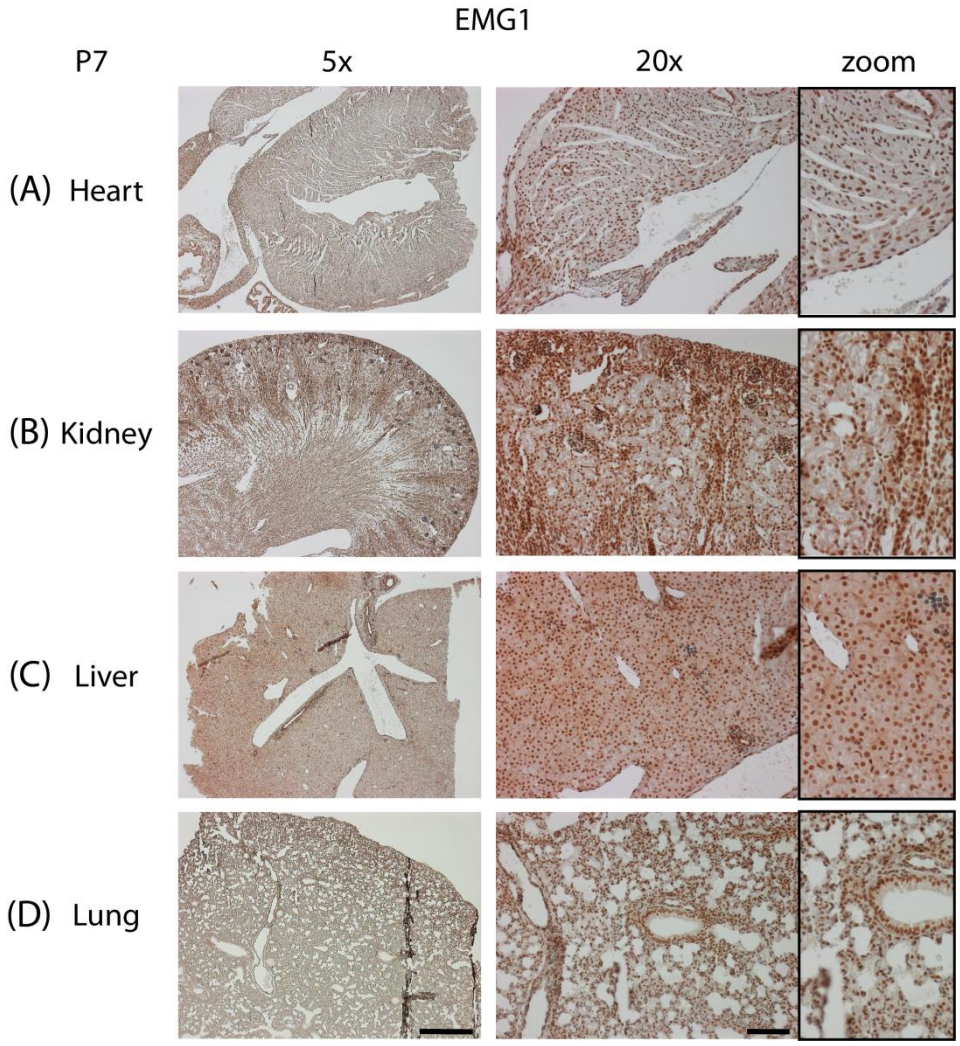

EMG1
5x
20x
zoom
(A) E7.5
(B) E9.5
fb
mb hb
he
ba
nt
(C) E12.5
fb
me
β-galactosidase
5x
20x
zoom
(D) E7.5
(E) E9.5
(F) E12.5

Figure 6.2 EMG1/ β-galactosidase during development.

EMG1 expression was strongest early in development (**A**, transverse section of E7.5 embryo), and diminished in later embryos (**B**, sagittal section of E9.5 embryo, **C**, sagittal section of E12.5 embryo). Although EMG1 was ubiquitous in later embryos, the level of expression was reduced compared to the E7.5 embryo. EMG1 was detected in formaldehyde-fixed, paraffin embedded embryos either with an anti-EMG1 antibody to detect the endogenous protein, or with an anti-β-galactosidase antibody to detect the product of the reporter gene under the control of the *Emg1* promoter (both 1:1000 dilution). Tissues were counterstained with methyl green. Scale bar at 5x = 500 µm and at 20x = 100 µm. **ba**, branchial arch; **fb**, forebrain; **hb**, hindbrain; **he**, heart; **mb**, midbrain; **me**, mesenchyme; **nt**, neural tube.


EMG1
P7
5x
20x
zoom
(A) Heart
(B) Kidney
(C) Liver
(D) Lung

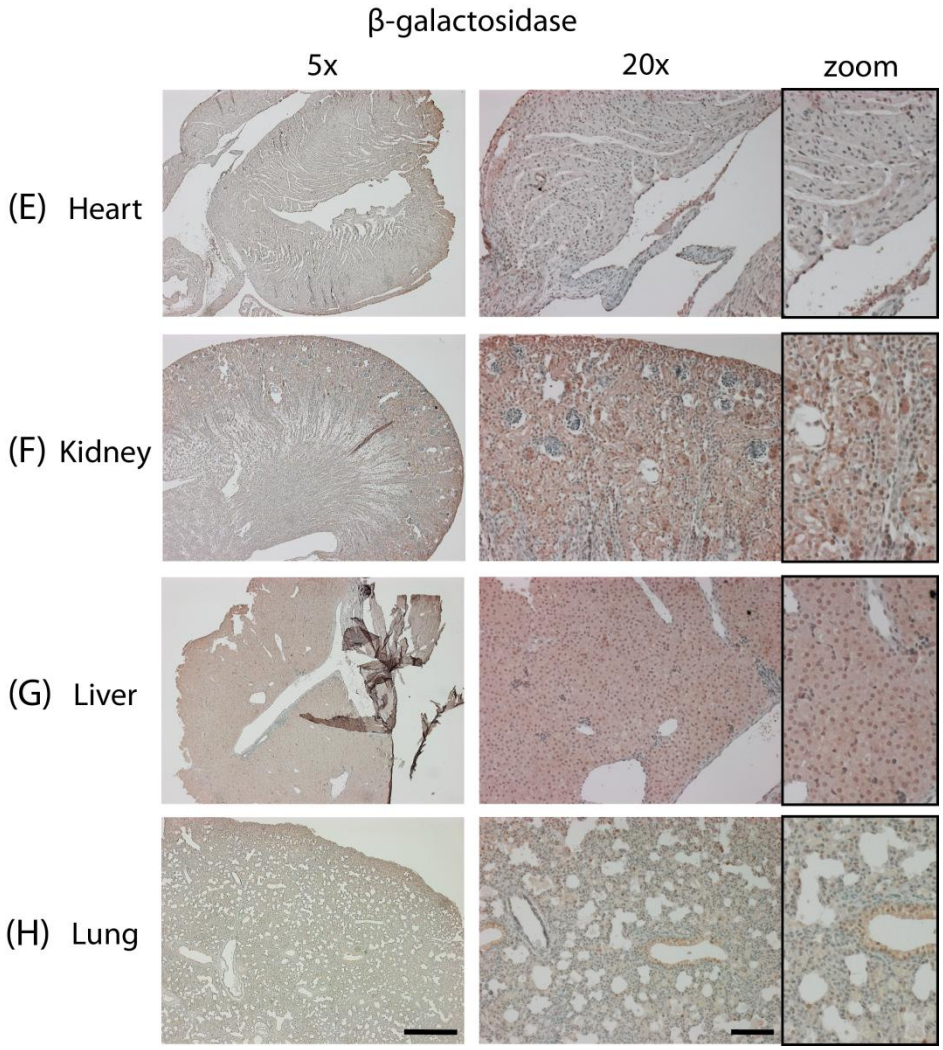
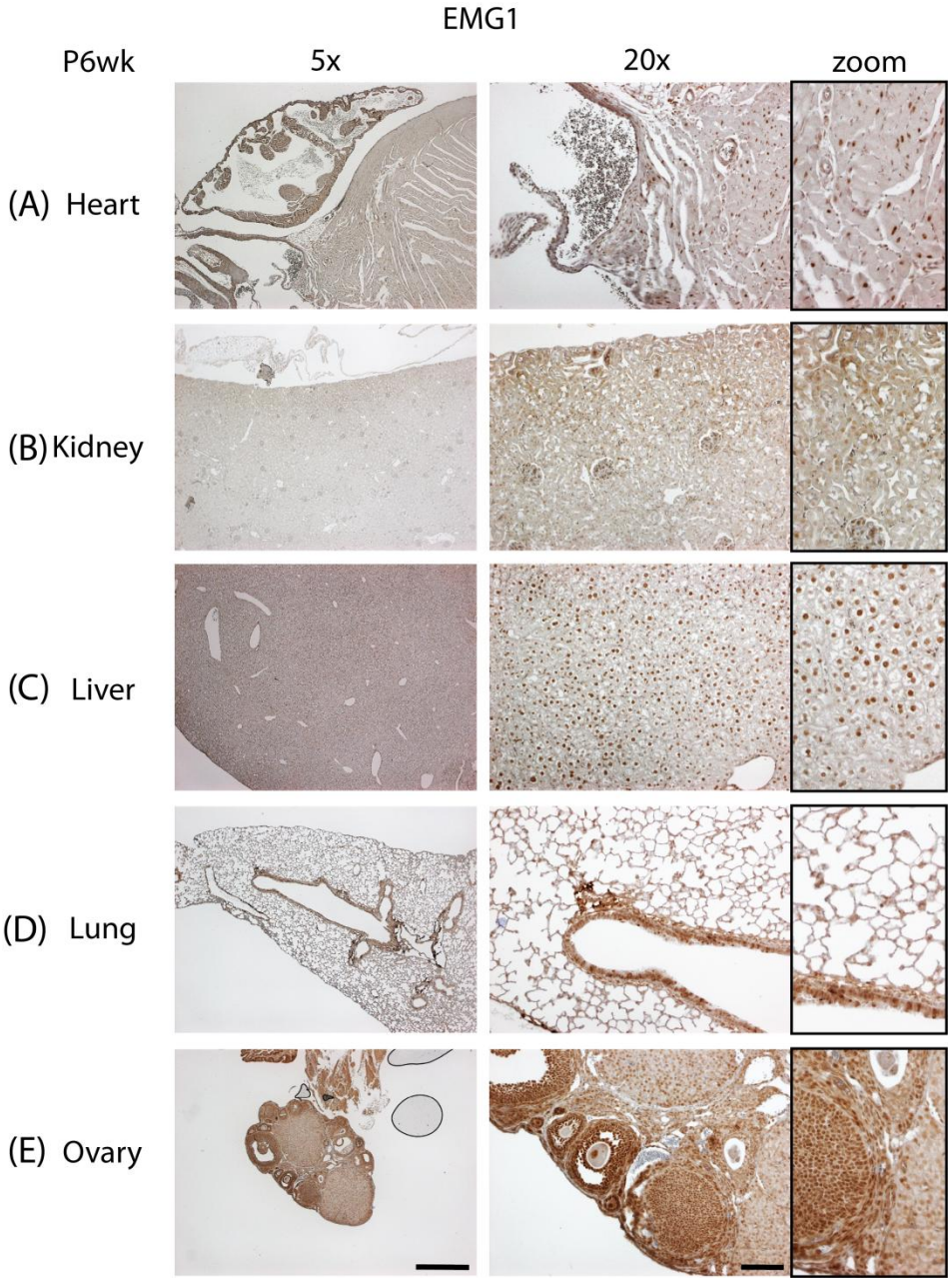


Figure 6.3 EMG1/ β-galactosidase at postnatal day 7.

EMG1 was found in all tissues examined, including the nuclei of myocardial cells in the heart **(A)**, in the cortex of the kidney with higher expression in the tubules than the glomeruli **(B)**, hepatocytes of the liver with both nuclear and cytoplasmic localization **(C)**, and in the lung, with highest expression in the bronchiole epithelium **(D)**. EMG1 was detected in formaldehyde-fixed, paraffin embedded tissues either with an anti-EMG1 antibody to detect the endogenous protein, or with an anti-β-galactosidase antibody to detect the product of the reporter gene under the control of the *Emg1* promoter (1:500 dilution). Tissues were counterstained with methyl green. Scale bar at 5x = 500 µm and at 20x = 100 µm.


EMG1
P6wk
5x
20x
zoom
(A) Heart
(B) Kidney
(C) Liver
(D) Lung
(E) Ovary

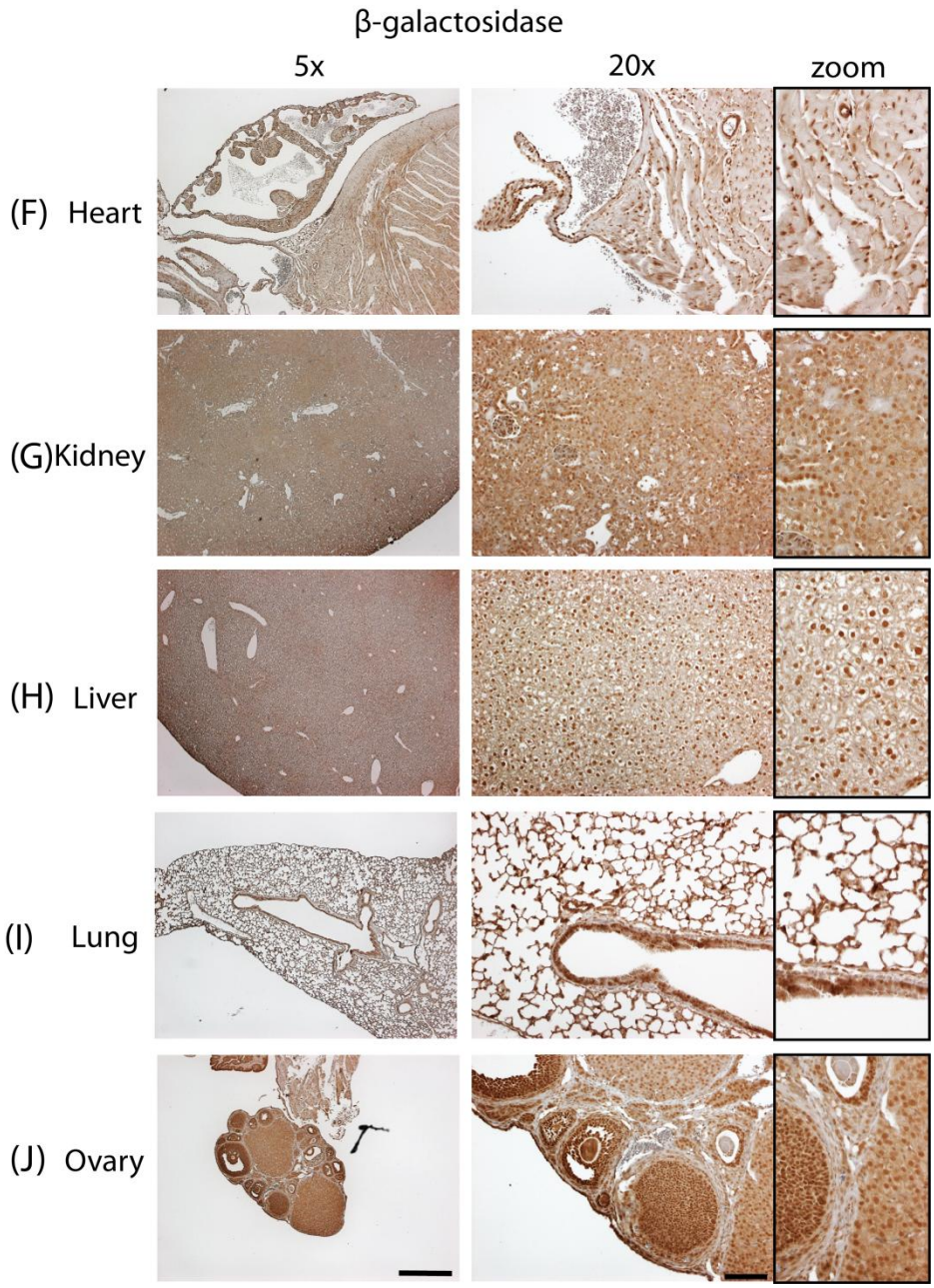


Figure 6.4 EMG1/ β-galactosidase at six weeks.

EMG1 expression was very similar to that at 7 days postnatal (Figure 6.3), although expression in the liver was more confined to the nuclei, and was less cytoplasmic. In the ovary **(E)**, EMG1 was strongest in the nucleus of the oocycte and the surrounding follicular cells, with lower levels found in the corpus luteum. EMG1 was detected in formaldehyde-fixed, paraffin embedded tissues either with an anti-EMG1 antibody to detect the endogenous protein, or with an anti-β-galactosidase antibody to detect the product of the reporter gene under the control of the *Emg1* promoter. Tissues were counterstained with methyl green. Scale bar at 5x = 500 µm and at 20x = 100 µm.

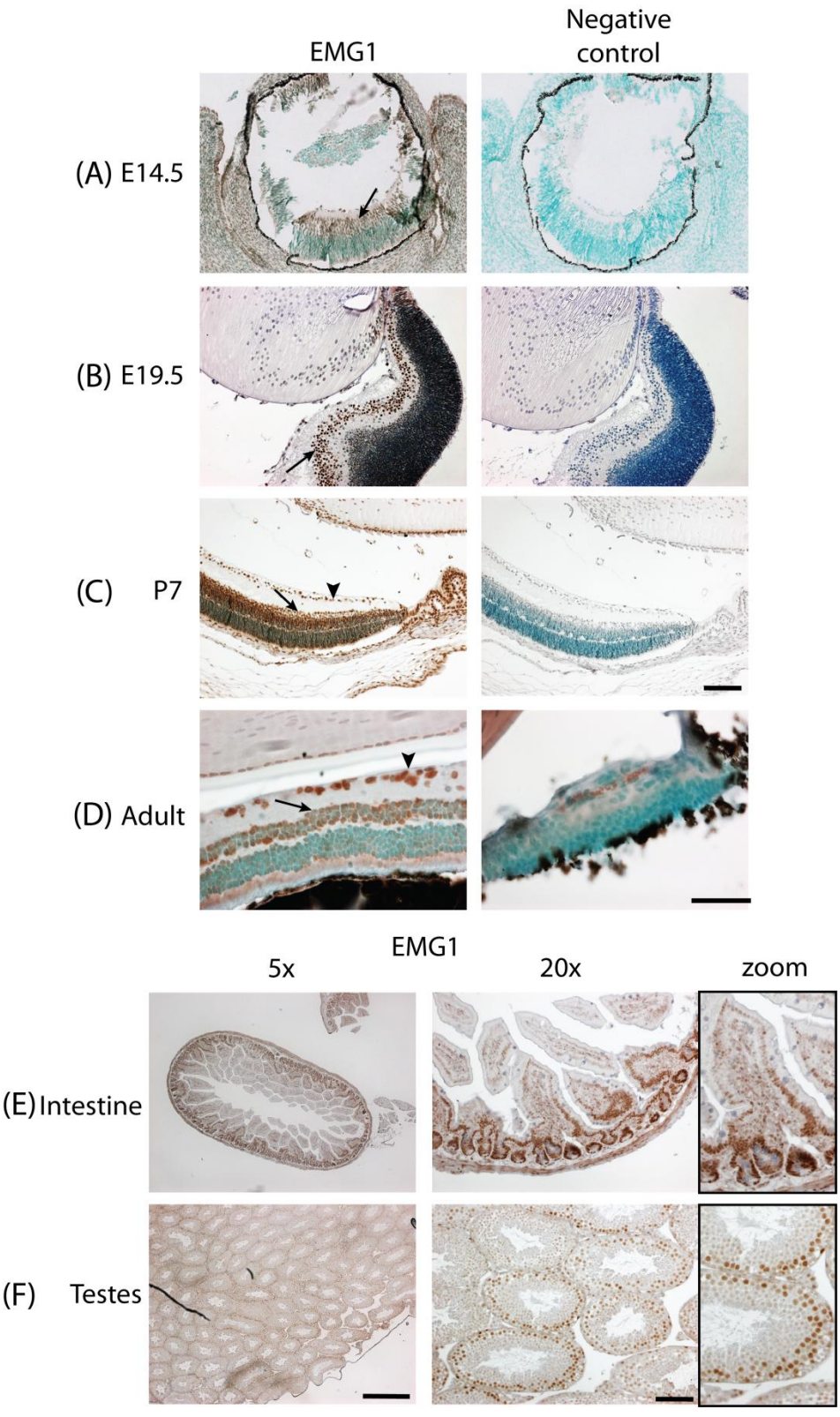

EMG1
Negative control
(A) E14.5
(B) E19.5
(C) P7
(D) Adult
EMG1
5x
20x
zoom
(E) Intestine
(F) Testes

Figure 6.5 EMG1 expression in eye, intestine and testes.

(A-D) EMG1 expression in mouse eye becomes restricted to the inner nuclear layer (arrows) and the ganglion cell layer (arrowheads) during development. In adult eye, EMG1 is absent in the outer nuclear layer (D). EMG1 was detected in formaldehyde-fixed, paraffin-embedded tissue sections using an anti-EMG1 antibody, and counterstained using methyl green. A brown precipitate indicates the presence of EMG1. The photos are oriented so that the inner nuclear layer is at the bottom and the ganglion cell layer is at the top. A-C, original magnification 20x; scale bar = 100 µm. D, original magnification 63x; scale bar = 50 µm. **(E)** EMG1 expression in the adult intestine is very strong in crypts, where cell proliferation is strongest, and diminishes towards the ends of the villae. **(F)** Similarly, EMG1 in adult testes is very strong in the outer layers of cells, with diminished expression in mature spermatocytes. Tissues were taken from 6 week old mice. Scale bars = 500 µm at 5x and 100 µm at 20 x.

development, as seen in the neuroepithelium of E7.5 embryos (Figure 6.6 A), and became more restricted as the brain developed (Figure 6.6 C). At E16.5, the area of the most concentrated expression was in the forebrain, in the superficial stratum of the cortex (arrow in Figure 6.6 C). Postnatally, EMG1 was found at varying levels in all areas of the brain, though the hippocampus and the cerebellum showed the strongest and most concentrated expression (Figure 6.6 D-F). In the hippocampus, expression in the adult mouse was restricted to the dentate gyrus. In the cerebellum, EMG1 was found in the granule cell layer, as well as strong expression in Purkinje neurons (arrowhead in Figure 6.6 F).

Taken together, these experiments suggest that EMG1 is required in all mouse tissues, at all stages of development and in adulthood. However, within certain tissues such as the retina, seminiferous tubules, or intestinal villae, EMG1 expression may be reduced or restricted to certain cell types over time. While these data indicate that EMG1 is required in highly proliferative cells such as in the intestinal crypts and testes, it was also found in post-mitotic cells, such as neurons of the adult brain.

6.3 Construct strategy

With a better understanding of EMG1 expression throughout development, we moved on to explore the role of the EMG1 D86G substitution in a mouse model of BCS. The etiology of BCS throughout development and in multiple tissues is best studied in the context of the whole organism. We therefore developed a strategy to produce a knock-in

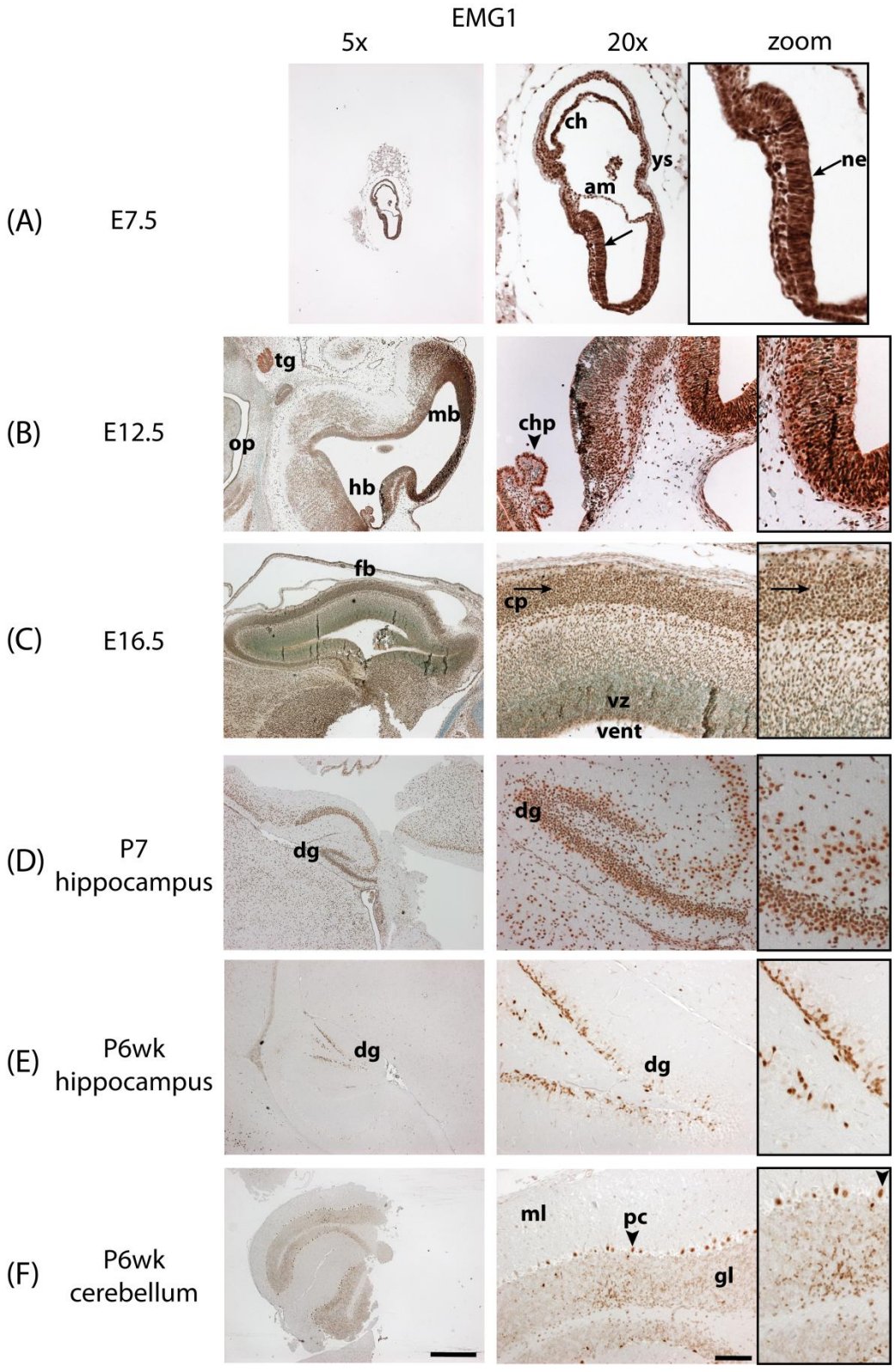

EMG1
5x
20x
zoom
(A)
E7.5
ch
ys
am
ne
(B)
E12.5
tg
mb
op
hb
chp
(C)
E16.5
fb
cp
vz
vent
(D)
P7
hippocampus
dg
dg
(E)
P6wk
hippocampus
dg
dg
(F)
P6wk
cerebellum
ml
pc
gl

Figure 6.6 EMG1 expression in brain throughout development and in adulthood.

Strong, even EMG1 expression was observed in the neuroepithelium at E7.5 (arrow) **(A)**, in the mid- and hindbrain and the choroid plexus (arrowhead) at E12.5 **(B)**, and began to become more restricted by E16.5 **(C)**, when strongest expression in the cortex was found in the superficial layer of the cortical plate (arrow). Postnatally, the most concentrated areas of expression were the dentate gyrus of the hippocampus at postnatal days seven and six weeks **(D)** and **(E)**, and the Purkinje neurons (arrowhead) and granular layer neurons of the cerebellum at six weeks **(F)**. EMG1 was detected in formaldehyde-fixed, paraffin-embedded sagittal tissue sections using an anti-EMG1 antibody, and counterstained using methyl green. Scale bars = 500 µm at 5x and 100 µm at 20 x. **am**, amnion; **ch**, chorion; **chp**, choroid plexus; **cp**, cortical plate; **dg**, dentate gyrus; **fb**, forebrain; **gl**, granular layer; **hb**, hindbrain; **mb**; midbrain; **ml**, molecular layer; **ne**, neuroepithelium; **op**, oropharynx; **pc**, Purkinje cell; **tg**, trigeminal ganglion; **vent**, ventricle; **vz**, ventricular zone.

mouse harbouring the equivalent BCS-causing A to G mutation in the mouse *Emg1* gene via homologous recombination. A targeting construct was generated containing exons 2-6 of *Emg1* with a single A to G mutation in exon 2 that would cause the equivalent aspartate 86 to glycine substitution in the mouse EMG1 protein (Figure 6.7). The construct carried a SA-IRES-βgeo cassette in intron 1 flanked by *FRT* sites for removal by Flp recombinase, a diphtheria toxin A cassette at the 3' end for negative selection, and 3′ and 5′ arms for homologous recombination at the wild type *Emg1* locus. The SA-IRES-βgeo cassette contained a splice acceptor, an internal ribosomal entry site, and a fusion gene of β-galactosidase and neomycin resistance for positive selection. Exons 2-6 were flanked by *loxP* sites for future conditional deletion of the *Emg1* gene using Cre recombinase.

6.4 Screening of ES cell clones

The targeting construct was electroporated into R1 *C129* mouse ES cells, and cells were selected in medium containing G418 for those harbouring the neomycin resistance gene. Surviving ES cell colonies were picked and expanded, and 192 clones were screened for single, targeted integration of the construct using both PCR assays and Southern blotting following restriction enzyme digestion (Figure 6.8 and Table 6.1). The 5′ and 3′ arms of homology were detected by Southern blot using specifically designed probes. Using the 5′ probe (Figure 6.8 A), restriction digestion with ScaI was expected to produce a band of 14978 bp in wild type cells and an additional band of 20198 bp in heterozygotes, while KpnI digestion was expected to produce a single 16717 bp band in wild type cells and an additional 9910 bp band in heterozygotes. Using the 3′ probe (Figure 6.8 B), ScaI

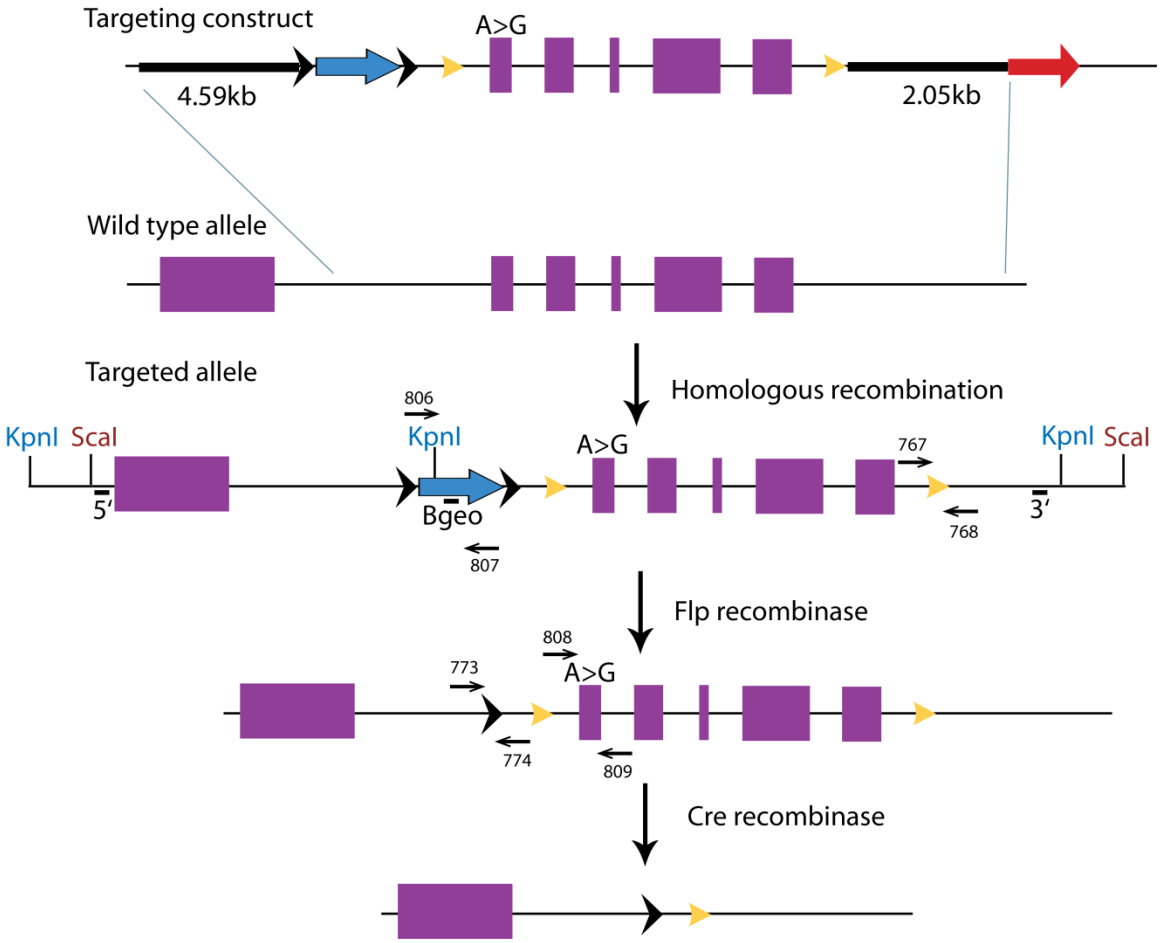


Figure 6.7 Diagram of the targeting construct and homologous recombination at the *Emg1* locus.

The targeting construct contained 5' and 3' arms of homology (black bars), an SA-IRES-βgeo cassette for positive selection (blue arrow), exons 2-6 of *EMG1* (purple boxes) with the BCS-causing A>G mutation in exon 2, and a Diphtheria Toxin A cassette for negative selection (red arrow). The βgeo cassette, which contains a splice acceptor, an internal ribosomal entry site, and a fusion of the β-galactosidase and neomycin resistance genes, is flanked by *FRT* sites (black arrowheads) for excision of the cassette by Flp recombinase. Similarly, exons 2-6 are flanked by *loxP* sites (yellow arrowheads) for future removal by Cre recombinase. Following homologous recombination with the wild type allele, the targeted allele contains three KpnI and two ScaI restriction digest sites for screening of correctly-targeted clones. The positions of the 5', 3' and βgeo probes for screening by Southern blot are shown as small black bars. Removal of the βgeo cassette in a correctly-targeted cell is mediated by Flp recombinase. The approximate positions of PCR primers to detect the βgeo cassette (WPG 806 and 807), the 3' *loxP* site (WPG 767 and 768), the single *FRT* site (WPG 773 and 774), and the A>G mutation (WPG 808 and 809), are indicated by small black arrows. Upon expression of Cre recombinase, recombination will take place between the two *loxP* sites, excising exons 2-6. Diagram is not to scale.

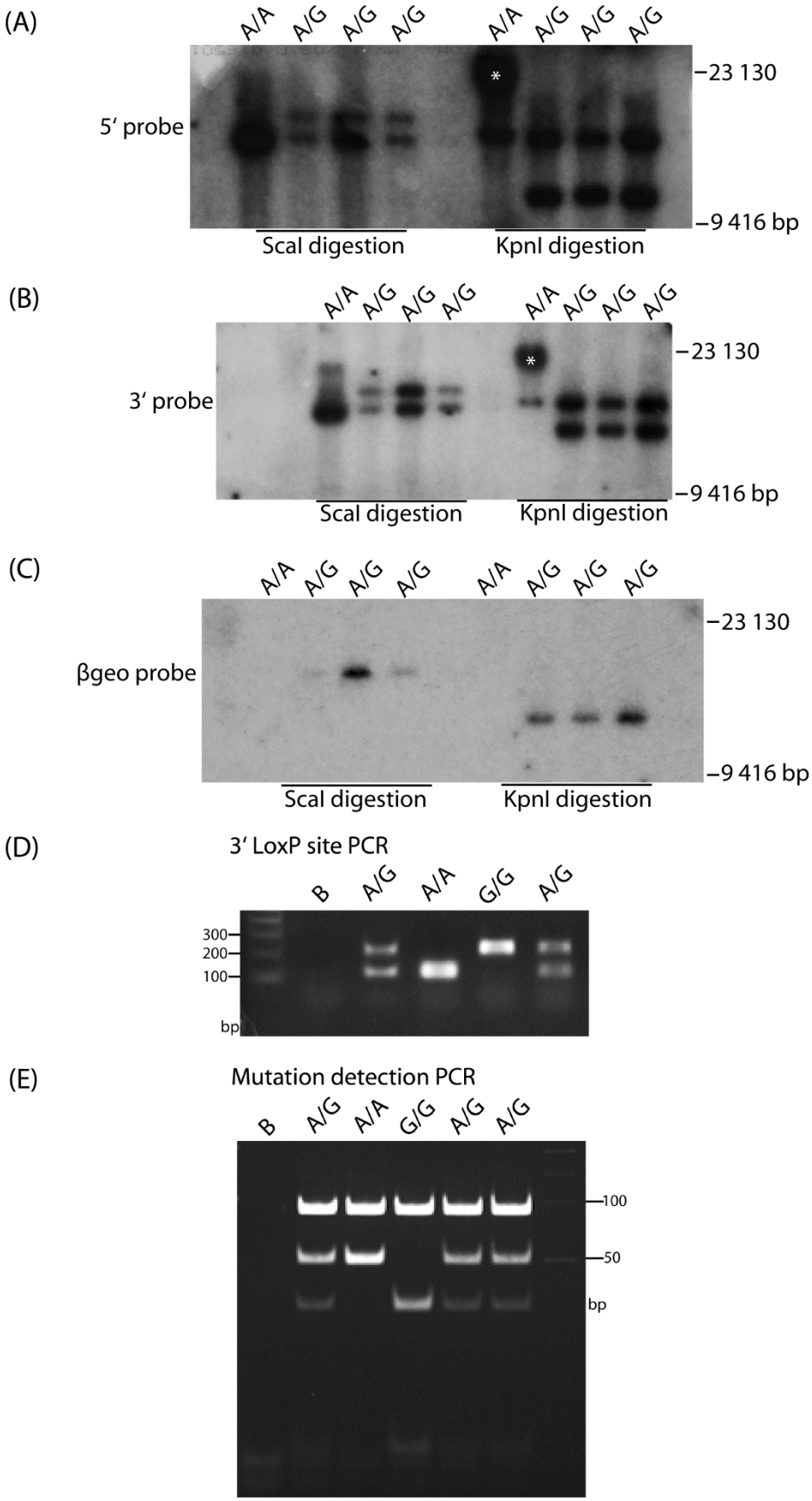

(A)
A/A
A/G
A/G
A/G
A/A
A/G
A/G
A/G
5′ probe
*
23 130
9 416 bp
ScaI digestion
KpnI digestion
(B)
A/A
A/G
A/G
A/G
A/A
A/G
A/G
A/G
3′ probe
*
23 130
9 416 bp
ScaI digestion
KpnI digestion
(C)
A/A
A/G
A/G
A/G
A/A
A/G
A/G
A/G
βgeo probe
23 130
9 416 bp
ScaI digestion
KpnI digestion
(D)
3′ LoxP site PCR
B
A/G
A/A
G/G
A/G
300
200
100
bp
(E)
Mutation detection PCR
B
A/G
A/A
G/G
A/G
A/G
100
50
bp

Figure 6.8 Embryonic stem cell screening.

(A-C) Southern blot screening of targeted embryonic stem cells with predicted band sizes. Briefly, DNA was isolated from ES cells, digested with either ScaI or KpnI, and separated by gel electrophoresis. DNA was transferred to a positively-charged nylon membrane and hybridized with specific radio-labeled probes generated by PCR. The 5' probe in **A** was amplified using WPG 759 and WPG 760 , the 3' probe in **B** was amplified using WPG 761 and WPG 762, and the βgeo probe in **C** was generated using WPG 763 and WPG 764. Film was exposed to the labeled membrane at -80°C, and developed. Wild type (A/A) DNA was run with each set of samples as a control (note that KpnI digestion was incomplete for the control in the 5' and 3' blots, causing an aberrant upper band, indicated by an asterisk). **(D-E)** PCR screening of targeted ES cells. PCR using primers specific for the 3' *loxP* site (**D**) (WPG 767 and WPG 768) and the A>G mutation (**E**) (WPG 808 and WPG 809) was performed, and the product from the mutation detection reaction was digested using ScrFI. DNA was separated by gel electrophoresis in a 1.5% agarose (**D**) or a 12% acrylamide gel (**E**), and visualized by staining with ethidium bromide. B, blank.

digestion was expected to produce a single 14978 bp band in wild type cells and an additional 20198 bp band in heterozygotes, while digestion with KpnI would yield a single 16717 bp band in wild type cells and an additional 12027 bp band in heterozygotes. A single integration of the targeting construct was verified by Southern blot with a βgeo-specific probe (Figure 6.8 C), which was expected to yield a 20198 bp band after ScaI digestion and a 12027 bp band following digestion with KpnI. The 3′ *loxP* site was detected by PCR. This yielded a single 207 bp band in the presence of the *loxP* site, and a band of 114 bp in its absence (Figure 6.8 D).The A to G mutation was also detected using PCR, followed by restriction digestion with ScrFI, which produced two bands at 88 and 50 bp in wild type cells, and three bands at 88, 36 and 15 bp in the presence of the mutation (Figure 6.8 E). Of the 192 clones screened, three independent correctly targeted colonies were identified, for a successful targeting rate of 1.6%. Undesired recombination between the mutation site and the 3′ *loxP* site, due to the distance of 1646 base pairs between them, was a major contributor to the low successful targeting rate. A much higher number of ES cell colonies (14.1%) showed correct targeting for the remaining four markers (5′, 3′, and βgeo probe Southern blots, and mutation detection PCR). Additionally, insufficient DNA isolation from the ES cell colonies resulted in an unsuccessful Southern blot in several experiments, and reduced the percentage of successfully targeted colonies.

	Restriction enzyme	WT	BCS
5'	ScaI	14978	20 224
	KpnI	16717	9906
3'	ScaI	14978	20 224
	KpnI	16717	12057

βgeo	**ScaI**	N/A	20198
	KpnI	N/A	12027
3' *loxP*	**N/A**	114	207
Mutation	**ScrFI**	88, 50	88, 36, 15

Table 6.1 **Expected band sizes in bp for Southern blot and PCR screening of ES cells**

6.5 Identification of chimeras

Two independent correctly targeted ES cell clones, denoted 2B10 and 2C5, were aggregated with wild type *CD1* morulae and cultured overnight to form blastocysts. Embryos were implanted into pseudopregnant females and the level of chimerism in the progeny was assessed based on coat colour (Figure 6.9). Males with a high level of chimerism from both the 2B10 and the 2C5 ES cell lines were crossed with wild type *CD1* females to verify germline transmission of the mutant allele, and offspring heterozygous for the mutation ($Emg1^{A/BgeoG}$) were produced. Thereafter, 2B10 and 2C5 mice were maintained as separate lines. The phenotypes of each line were indistinguishable; therefore the following descriptions apply to both.

6.6 Heterozygous mutant mice

6.6.1 Heterozygous mice with the βgeo cassette

$Emg1^{A/BgeoG}$ mice were intercrossed, but genotyping revealed that homozygous mutant $Emg1^{BgeoG/BgeoG}$ mice were not viable; out of 161 births, 50 were wild type (31.1%) and 111 (68.9%) were $Emg1^{A/BgeoG}$. This result was not unexpected since the targeted gene still contained the ~5 kb SA-IRES-βgeo cassette in the first intron. The strong splice acceptor in the cassette has been shown to effectively create a null allele by preventing

Figure 6.9 *Emg1* A>G chimeric mice.

The white coat colour is from the *CD1* morulae, while the brown coat colour originates from the targeted *C129* ES cells. Chimeras from two independent ES cell lines were produced, and maintained separately thereafter.

the incorporation of the downstream endogenous exons[194]. Since *Emg1* expression was likely compromised and it had already been established that loss of *Emg1* is lethal prior to implantation[122], studies to determine the embryonic stage of lethality were not pursued and we instead proceeded to excise the SA-IRES-βgeo cassette.

6.6.2 Removal of the βgeo cassette in heterozygous mice by Flp recombinase

The SA-IRES-βgeo cassette was flanked by *FRT* sites, and was therefore removable by expression of Flp recombinase (*Flpe*). Heterozygous *Emg1*$^{A/BgeoG}$ males derived from each independently targeted ES cell line were crossed with *C57BL/6 Flpe*-expressing females (B6.129S4-Gt(ROSA)26Sor$^{tm1(FLP1)Dym}$/RainJ) purchased from Jackson Labs (jaxmice.jax.org). In these deleter mice, a thermostable version of Flp recombinase is expressed from the *ROSA26* locus, a common site for the insertion of transgenes[194]. Using these mice, target gene recombination has been shown to take place in all tissue types, including the developing germ line[135]. The presence of the *Flpe* transgene in progeny was detected by PCR (Figure 6.10 A). However, Flp recombinase efficiency is known to be low compared to Cre recombinase[195], creating mosaicism, wherein site-specific recombination takes place in only a subset of cells. Therefore, removal of the floxed SA-IRES-βgeo cassette was verified using two different parameters: one PCR to detect the presence or absence of the SA-IRES-βgeo cassette, and another for the presence of a single remaining *FRT* site following recombination (Figure 6.10 B and C). Following removal of the βgeo cassette, a single *FRT* site remained in intron 1 (see black arrowhead in Figure 6.7), which was detected as a 229 bp band. Because the primers were positioned on either side of the βgeo cassette, the intervening 5 kb sequence was too

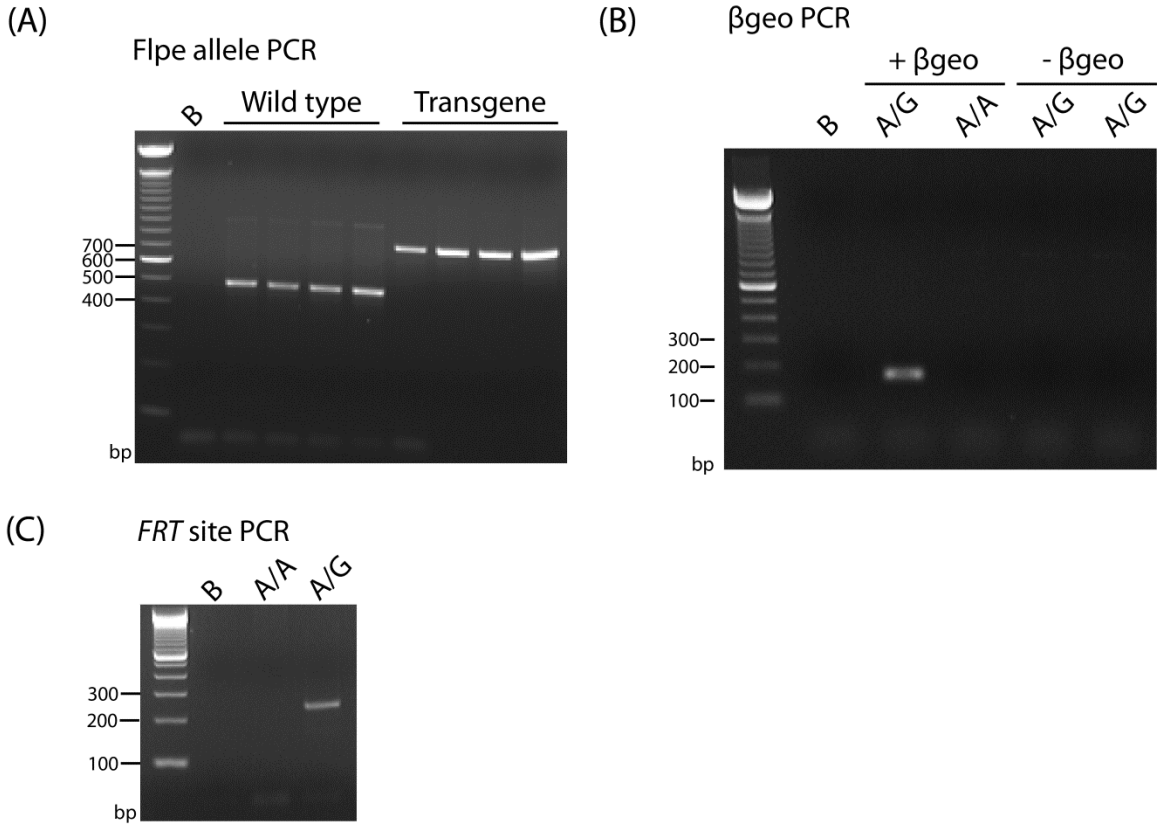


Figure 6.10 Strategy for detection of βgeo cassette removal.

(A) PCR to detect the wild type ROSA26 locus and *Flpe* transgene. *Flpe*-expressing mice harboured the *Flpe* transgene at the ROSA26 locus, creating a 725 bp band using the *Flpe* transgene primers (WPG 769 and WPG 770). The wild type ROSA26 locus was detected using a separate primer pair, giving a 500 bp band (WPG 771 and WPG 772). **(B)** PCR to detect the SA-IRES-βgeo cassette (primers WPG806 and WPG807). In the presence of the βgeo cassette, a band was detected at 172 bp in heterozygous mice. Following Flp recombinase expression and subsequent excision of the βgeo cassette, the band was undetectable in heterozygous mice. **(C)** PCR to detect the presence of a single *FRT* site after removal of the SA-IRES-βgeo cassette (Primers WPG 773 and WPG 774). Primers were located on either side of the βgeo cassette, and in heterozygous mice, a single *FRT* site is detectable at 229 bp following excision of the βgeo cassette. The *FRT* site amplification product is not evident for the allele with the βgeo cassette because the intervening sequence is too long, at 5 kb, to be amplified under the conditions utilized.

long to be amplified under the conditions utilized, and did not produce a band if the βgeo cassette remained (Figure 6.10 C). However, this strategy would still amplify a band representing a single *FRT* site in the case of mosaicism, confounding the results. A PCR designed using primers internal to the βgeo cassette (Figure 6.10 B) was therefore used to confirm complete removal of the cassette. An animal was considered to be βgeo-free only in presence of a single *FRT* site, and in the absence of the βgeo cassette.

At the same time, mice heterozygous for the *Emg1* mutation and for *Flpe* were crossed in an effort to produce mice homozygous for the *Flpe* transgene, to increase the efficiency of Flp recombinase-mediated recombination. Unfortunately, the majority of mice which were positive for the *Flpe* transgene also expressed wild type *Emg1*, since the *ROSA26* locus and *Emg1* in mice are both on chromosome 6, within approximately 11.6 Mb of each other. This translates to a recombination rate between the two loci of approximately 5.3%[196], meaning extensive breeding was necessary to obtain mice expressing both *Flpe* and mutated *Emg1*. Hence linkage between the *ROSA26* locus and *Emg1*, combined with embryonic lethality of the homozygous *Emg1* mutation, meant that only 10 out of 353 mice (2.8%) heterozygous for the mutation and homozygous for the *Flpe* transgene were obtained. Once mice expressed *Flpe* and the mutated form of *Emg1* together, the SA-IRES-βgeo cassette was excised.

Once heterozygous mice lacking the SA-IRES-βgeo cassette (*Emg1*$^{A/G}$*)* were established, they were intercrossed to produce homozygous *Emg1*$^{G/G}$ or "BCS" mice. Genotyping of

progeny revealed 33.8% wild type and 66.2% heterozygous mice, and no homozygous mutants (χ^2 = 23.68, p< 0.0001,Table 6.2, "weanlings" row). Thus, the mutation in the homozygous state was lethal before genotyping was performed at 21 days, even in the absence of the SA-IRES-βgeo cassette.

6.6.3 Determination of stage of lethality of Emg1$^{G/G}$

Newborn litters were carefully monitored, and since no evidence of cannibalism by the mother was observed, it seemed likely that *Emg1*$^{G/G}$ mice died before birth. To establish the embryonic stage of lethality, timed *Emg1*$^{A/G}$ intercrosses were performed, and the morning of the discovery of a vaginal plug was considered embryonic day (E) 0.5. Embryos collected at E8.5 to E13.5 were genotyped using DNA extracted from the yolk sac or from formaldehyde-fixed, paraffin embedded tissue sections. The results are summarized in Table 6.2. Embryos were also taken at E6.5 (n = 24 embryos) and E7.5 (n = 28 embryos), but were not genotyped due to lack of material and were identified only by morphology and expected Mendelian ratios. Resorption sites in the uterus, which represent sites in which the embryo died after implantation, were noted, and tissue was collected for genotyping whenever possible.

		% observed			
Age	**N**	**A/A**	**A/G**	**G/G**	**resorptions**
Weanlings total	**71**	**33.8**	**66.2**	**0**	**N/A**
Embryonic	**319**	**22.3**	**51.1**	**10.6**	**16.0**

total					
E 8.5	44	20.5	61.4	15.9	2.3
E 9.5	67	20.9	52.2	16.4	10.5
E 10.5	42	38.1	40.5	11.9	9.5
E 11.5	92	14.1	54.4	9.8	21.7
E 12.5	65	24.6	46.2	3.1	26.2
E 13.5	9	33.3	44.4	0	22.2

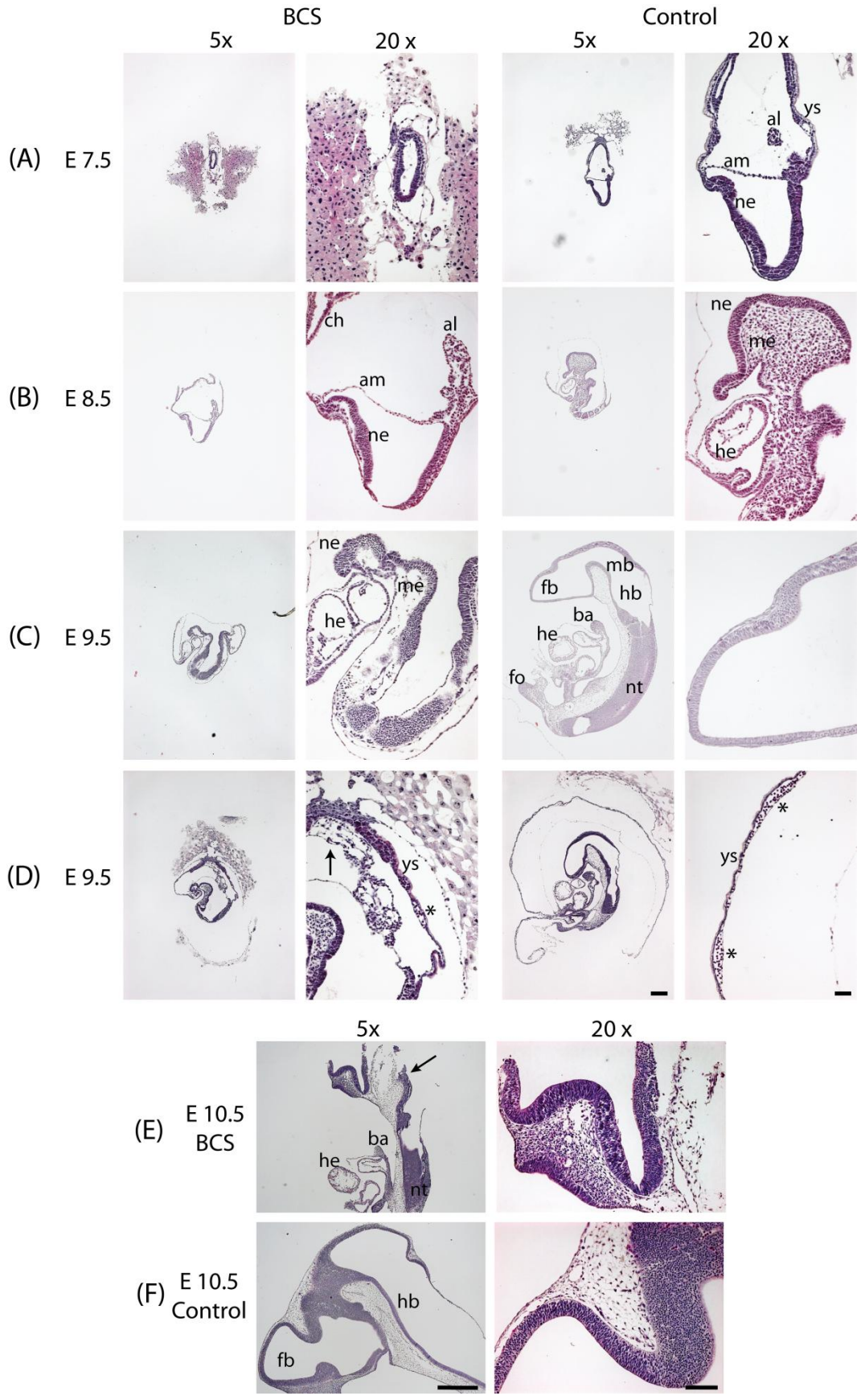
Table 6.2 Genotype distributions from heterozygous intercrosses

Contrary to the pre-implantation death found in *Emg1* knockout mice[122], *Emg1*$^{G/G}$ embryos were identified between E8.5 and E12.5. However, only 10.6% of embryos were homozygous mutants, which is significantly below the expected Mendelian proportion of 25% ($\chi^2 = 22.8$, $p < 0.0001$ with 2 degrees of freedom) (Table 6.2). The missing homozygous mutants may be explained by the 16.0 % of resorption sites found in the uterus. Although genotyping of resorptions was not possible in most instances, nearly all of those genotyped were homozygous mutants, suggesting that the BCS embryos died post-implantation. Notably, the number of homozygous mutants decreased with embryonic age, accompanied by an increase in resorption sites, indicating that more *Emg1*$^{G/G}$ embryos are viable earlier in development. The presence of resorption sites at every stage examined suggests that the homozygous mutation is lethal after implantation at variable ages. After E12.5, no *Emg1*$^{G/G}$ embryos were found. Taken together, these results indicate that the stage of lethality is variable, occurring between E8.5 and E12.5.

6.6.4 Morphology of homozygous mutant embryos

Upon dissection, all *Emg1*$^{G/G}$ embryos were easily distinguishable from littermates due to their severely reduced size, with a consistent developmental delay of approximately one day. At E6.5 to E12.5, embryos were collected, fixed in 4% formaldehyde, embedded in paraffin, serially sectioned, and stained with hematoxylin and eosin for morphological examination. Despite the growth retardation, BCS embryos were generally normal morphologically (Figure 6.11, compared with *Emg1*$^{A/G}$ littermates except where indicated otherwise). A beating heart was observed in most BCS embryos from E9.5 to E12.5, and blood islands were visible in the yolk sac in sagittal sections (Figure 6.11), suggesting that the homozygous mutant embryos were viable, despite their developmental delay.

Aside from the growth retardation, the only major morphological abnormality was an open neural tube in many of the BCS embryos from E10.5 to E12.5. Proper cranial neural tube closure begins at approximately E8.5 in the hindbrain, and includes two other discrete points of closure, at the forebrain-midbrain boundary and the rostral extremity of the forebrain. The tube zips shut between these points and the process terminates by E9.5[197]. Approximately 50% of BCS embryos that were otherwise beyond the stage of neural tube closure, had an open cranial neural tube (Figure 6.12, A and B). Eversion of the neural folds led to exencephaly by E12.5 (Figure 6.11 H and 6.12 E). No BCS embryos showed defects of spinal neural tube closure. However, transverse sections at the level of the second pharyngeal arch at E10.5, E11.5 (Figure 6.12, C, D) and E12.5


BCS
Control
5x
20 x
5x
20 x
(A) E 7.5
al
ys
am
ne
(B) E 8.5
ch
al
am
ne
ne
me
he
(C) E 9.5
ne
me
he
mb
fb
hb
ba
he
fo
nt
(D) E 9.5
ys
*
ys
*
*
5x
20 x
(E) E 10.5 BCS
ba
he
nt
(F) E 10.5 Control
hb
fb

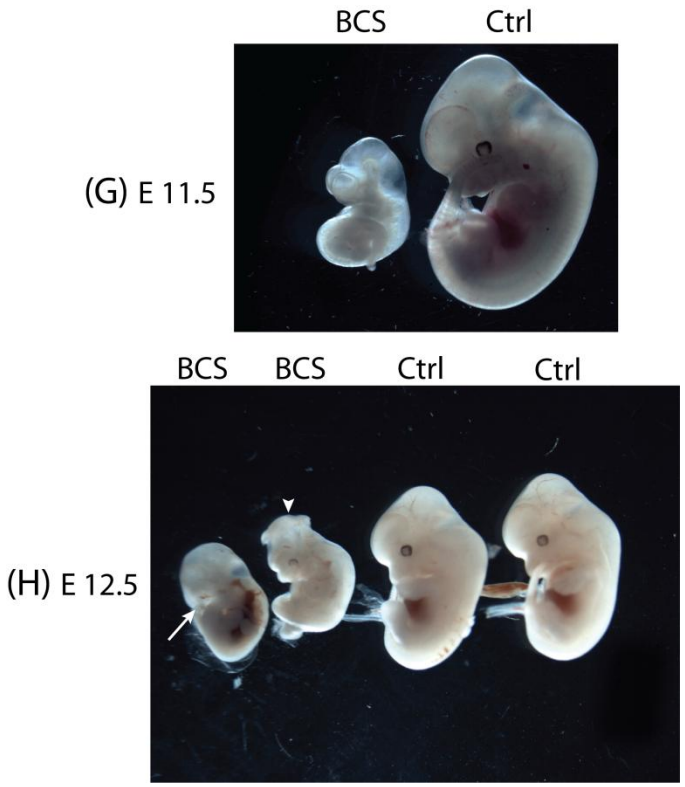


Figure 6.11 BCS embryo morphology.

(A-E) Mid-sagittal sections of formaldehyde-fixed, paraffin-embedded embryos at different stages of development, stained with hematoxylin and eosin to show morphology. BCS and control embryos at each time point are taken from the same litter, and a heterozygous control is shown wherever possible. In each photomicrograph, ventral is to the left, dorsal is to the right, rostral is at the top and caudal is at the bottom. Original magnification 5x and 20x. **(A)** Putative BCS embryo and control from the same litter at E7.5. The embryos were not genotyped due to a lack of available material, but were identified based on morphology and Mendelian ratios. **(B)** BCS and heterozygous control embryos from the same litter at E8.5. Note the similarities between the BCS embryo at E8.5 and the control embryo at E7.5. **(C)** BCS and heterozygous control embryos from the same litter at E9.5. **(D)** Chorioallantoic fusion and blood island formation at E9.5. Note the attachment of the allantois to the chorion (arrow), and the presence of blood islands in the yolk sac (asterisks). Scale bar at 5x = 200 µm, at 20x = 50 µm. **(E** and **F)** BCS and heterozygous control embryos from the same litter at E10.5. Note the open neural tube in the cranial region of the BCS embryo (arrow). Scale bar at 5x = 500 µm, at 20x = 100 µm. **(G)** Photograph taken using a dissecting microscope of a BCS and a control embryo from the same litter at E11.5. The neural tube in the BCS embryo is closed in this instance. **(H)** Photograph of BCS and control embryos from the same litter at E12.5. The first BCS embryo has a facial cleft (arrow), and the second has exencephaly in the midbrain region (arrowhead). **am**, amnion; **al**, allantois; **ba**, branchial arch; **ch**, chorion; **fb**, forebrain (telencephalon); **fo**, forelimb bud; **he**, heart; **hb**, hindbrain (metencephalon); **mb**, midbrain (mesencephalon); **me**, mesenchyme; **ne**, neuroepithelium; **nt**, neural tube; **ys**, yolk sac.

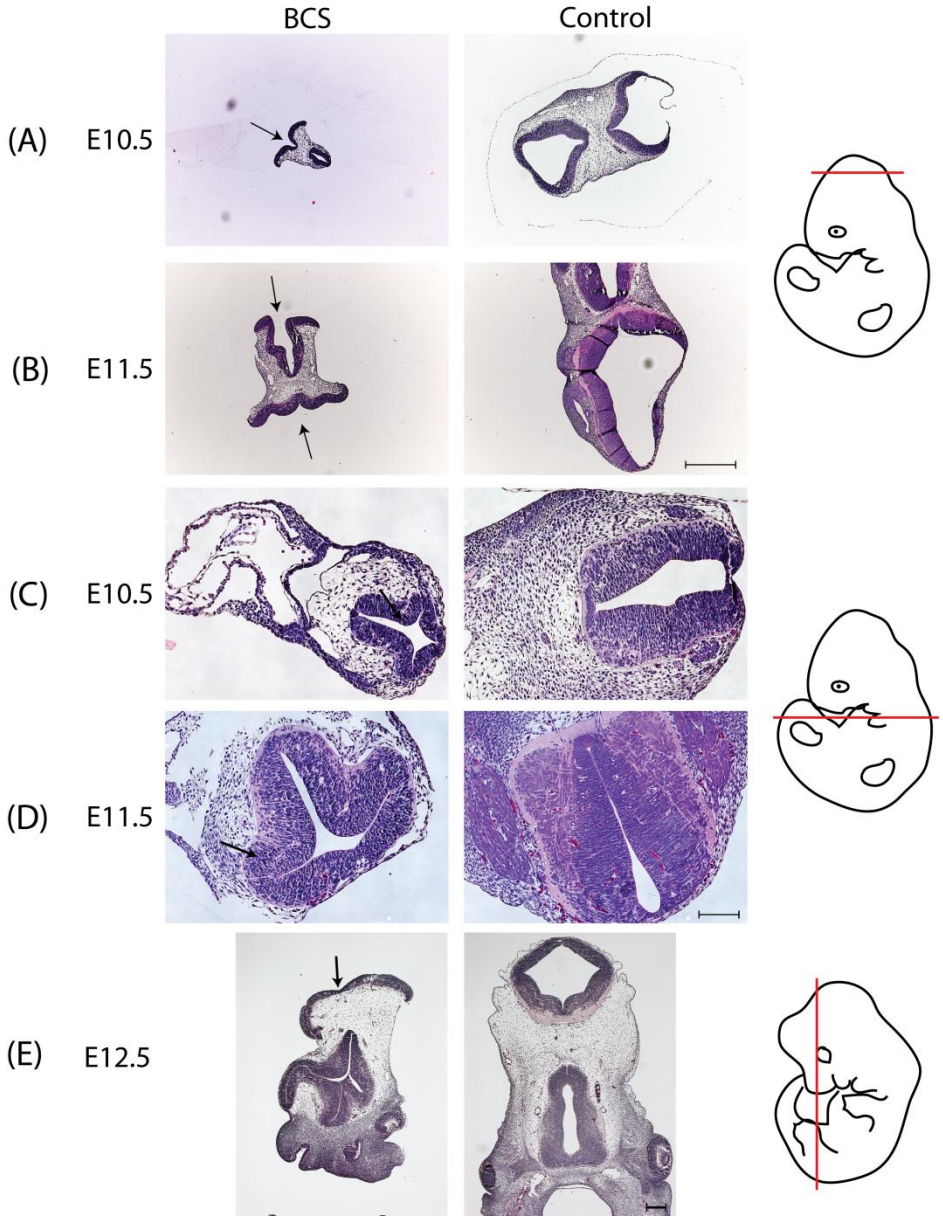


Figure 6.12 Neural tube defects in BCS embryos.

Sections were stained with hematoxylin and eosin. In A and C, dorsal is to the left and ventral is to the right, while in B and D, dorsal is to the bottom and ventral is to the top. Diagrams show the approximate plane of the section. **(A** and **B)** Transverse sections through the top of the head showed an open neural tube in the BCS embryos (arrows), while the neural tube had already closed in control mice. Original magnification 5x, scale bar = 500 µm. **(C** and **D)** More caudal sections in the same embryos, at the level of the second pharyngeal arch, showed an abnormal cross-shaped neural tube in the BCS embryos (arrows), compared to the straight neural tube seen in controls. Original magnification 20x, scale bar = 100 µm. Note that the control in (A) and (C) is *Emg1*$^{A/A}$. **(E)** Coronal sections of the head region of the second and fourth embryos pictured in Figure 6.11 H, showing an exencephalic brain in the BCS embryo (arrow). Original magnification 5x, scale bar = 200 µm.

(not shown) in embryos with an open neural tube revealed an abnormal cross-shaped neural tube, while controls from the same litter had a straight neural tube in cross-section. The defect was only present at this level in the embryos, as sections caudal to this point showed a normal, straight neural tube (not shown). These results indicate that the A to G mutation in *Emg1* has profound effects on both embryonic growth and development of the central nervous system, and may preferentially target the cranial and cranial-spinal boundaries of the central nervous system.

A minority (n = 4/319) of wild type and heterozygous embryos displayed other morphological defects, including disorganization, relatively small head, and dysgenesis of the posterior section of the embryo. However, these embryos were usually easily distinguishable from the $Emg1^{G/G}$ embryos, which were considerably smaller and never displayed overt morphological defects other than the open neural tube. These few unusual embryos are likely naturally occurring malformations on the *CD1/C129/C57* background.

6.6.5 EMG1 levels

In BCS patient cells, the level of EMG1 protein is drastically reduced compared to control cells (see Chapter 3). We therefore expected a similar reduction in EMG1 protein in the BCS mouse model. EMG1 was detected in control and BCS mouse sections using an anti-EMG1 antibody, showing reduced staining intensity in BCS mice (Figure 6.13). This suggests that the A to G mutation in *Emg1* has a similar destabilizing effect on the

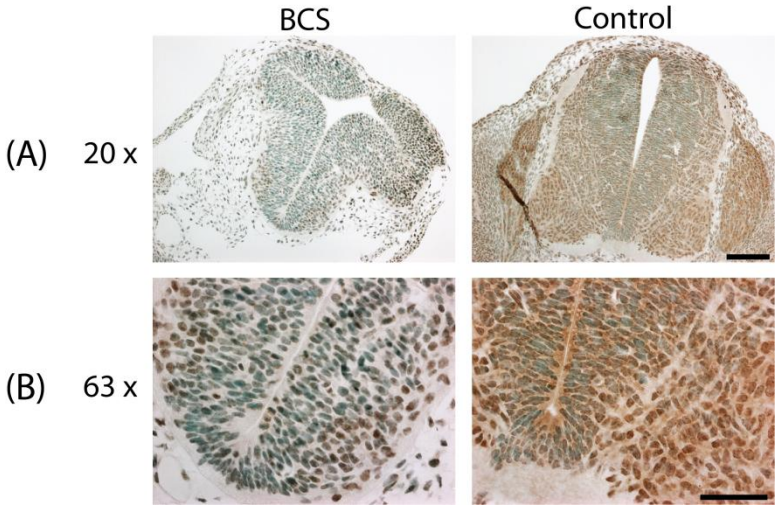


Figure 6.13 Reduction of EMG1 staining in BCS embryos.

Formaldehyde-fixed, paraffin-embedded transverse sections of the neural tube from BCS and control embryos at E11.5 were probed for the presence of EMG1 using an anti-EMG1 antibody and the ABC/DAB system. Note the reduced detection of brown EMG1 staining in BCS embryos. Dorsal is at the top of the image, and ventral is at the bottom. In B, the ventral portion of the neural tube is magnified. Nuclei were counterstained with methyl green. Scale bar at 20x = 100 µm, at 63x = 50 µm.

EMG1 protein in BCS patients and in *Emg1*$^{G/G}$ mice. Immunoblot analysis using lysates from BCS mouse tissue is necessary to confirm these results.

6.6.6 Cell proliferation in homozygous mutant embryos

The growth retardation in BCS embryos could be the result of reduced cell proliferation occasioned by a ribosome biogenesis defect, similar to what was seen in BCS patient lymphoblasts (Chapter 5). In order to assess cell proliferation rates in mouse embryos, pregnant females were injected with bromodeoxyuridine (BrdU), a thymidine analogue which incorporates into the DNA of cycling cells during S phase, and sacrificed two hours later (n = 3 pregnant females per time point). BrdU was detected in formaldehyde-fixed, paraffin embedded embryo sections using an anti-BrdU antibody (Figure 6.14). Time points from E7.5 to E9.5 were chosen because they reliably yielded viable *Emg1*$^{G/G}$ embryos, whereas a BCS embryo was less frequent at later time points. Reflecting the high rate of cell proliferation in embryonic development, control embryos consistently showed strong staining in the majority of cells. By contrast, staining in BCS embryos was much more variable. Some embryos displayed strong staining similar to that of controls (n = 3/8 BCS embryos), but others had reduced or nearly absent BrdU levels (n = 5/8). In fact, BCS embryos even within the same litter were observed with differing levels of staining (Figure 6.14 C and D). The variable staining reflects the varying neural tube defects and ages of death of BCS embryos, and suggests that some embryos continue growing while others stop proliferating and eventually die. The almost complete

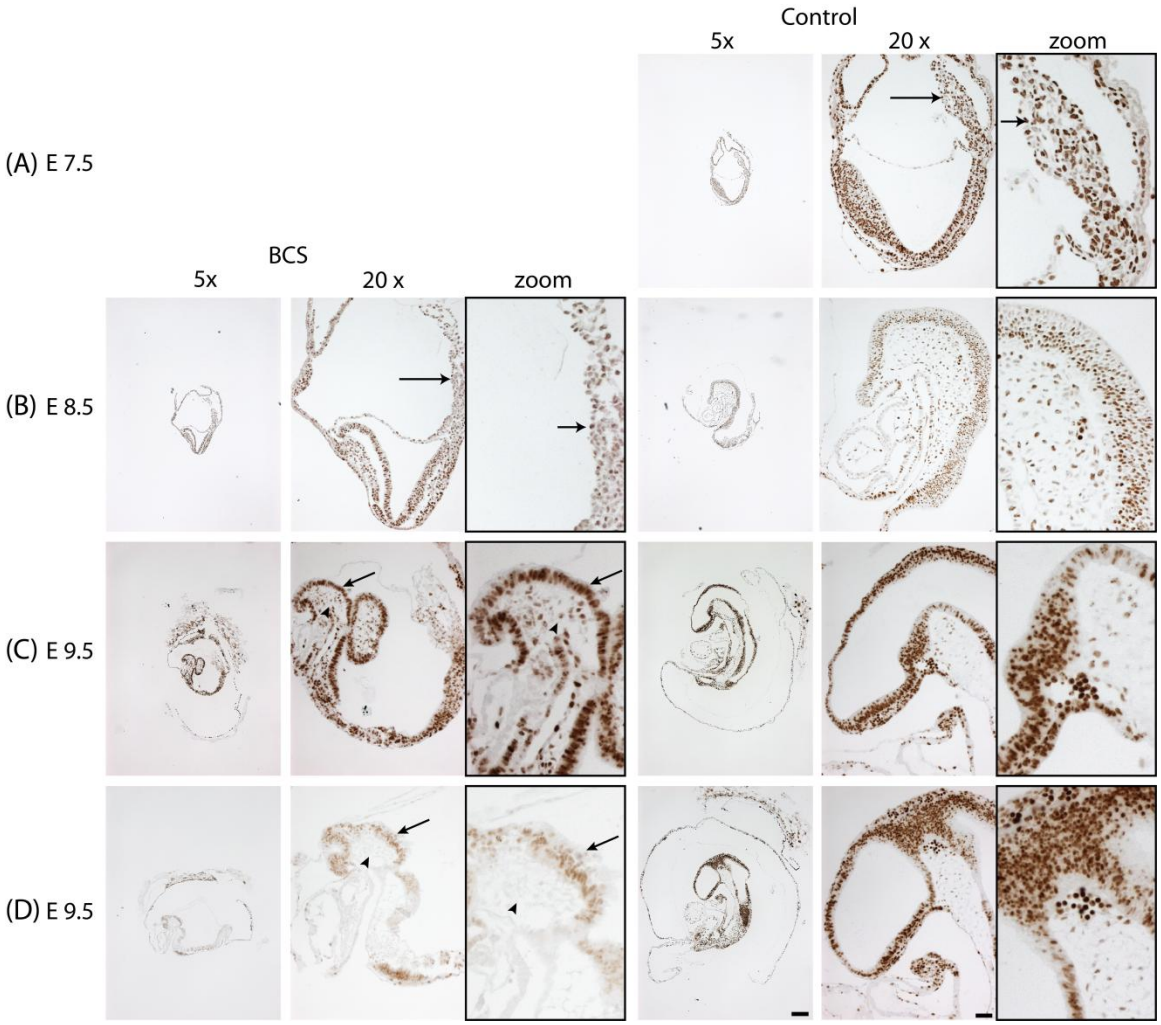


Figure 6.14 Cell proliferation in BCS embryos.

(A-D) Proliferating cells in BCS and control embryos were detected by injecting pregnant females with bromodeoxyuridine to label DNA in S phase cells. Embryos were collected two hours post-injection, and mid-sagittal sections were stained using an anti-BrdU antibody. Compare the strong staining in the allantois of the control embryo in **(A)** with the weaker staining in the allantois of the BCS embryo in **(B)** (arrows). **(C)** and **(D)** show differing levels of BrdU staining in BCS embryos from the same litter, while control embryos have similar levels of staining. Compare the strong staining in the neuroepithelium of the BCS embryo in **(C)** with weak staining in **(D)** (arrows), and the absence of BrdU in the underlying cranial mesenchyme in **(D)** (arrowheads). Scale bar at 5x = 200 µm, at 20x = 50 µm.

lack of cell proliferation observed in some embryos could indicate that they were already dead, however no obvious signs of necrosis were observed upon dissection.

Many *Emg1*$^{G/G}$ embryos at E8.5 and E9.5 showed strong BrdU staining in the embryo itself, but reduced staining in the allantois compared to controls (Figure 6.14 B, C), suggesting that a possible failure of umbilical cord or placental formation may contribute to the growth retardation in BCS mouse embryos (n = 3/5 BCS embryos in which the allantois was visible). However, no overt placental defects were found in the BCS embryos that survived beyond E10.5, and the allantois still appeared to fuse with the chorion in BCS embryos (see Figure 6.11 D). Thus an allantoic defect may not play a role in growth restriction in all cases, but a delay in chorioallantoic fusion in BCS embryos due to slower proliferation of the allantois could further exacerbate developmental delay.

These data indicate that the D86G substitution in EMG1 causes a reduction of cell proliferation during development, in some cases affecting growth of the allantois, leading to severe growth retardation. As the cell proliferation defect varies in severity, even within a litter, the developmental stage to which an embryo can survive varies as well.

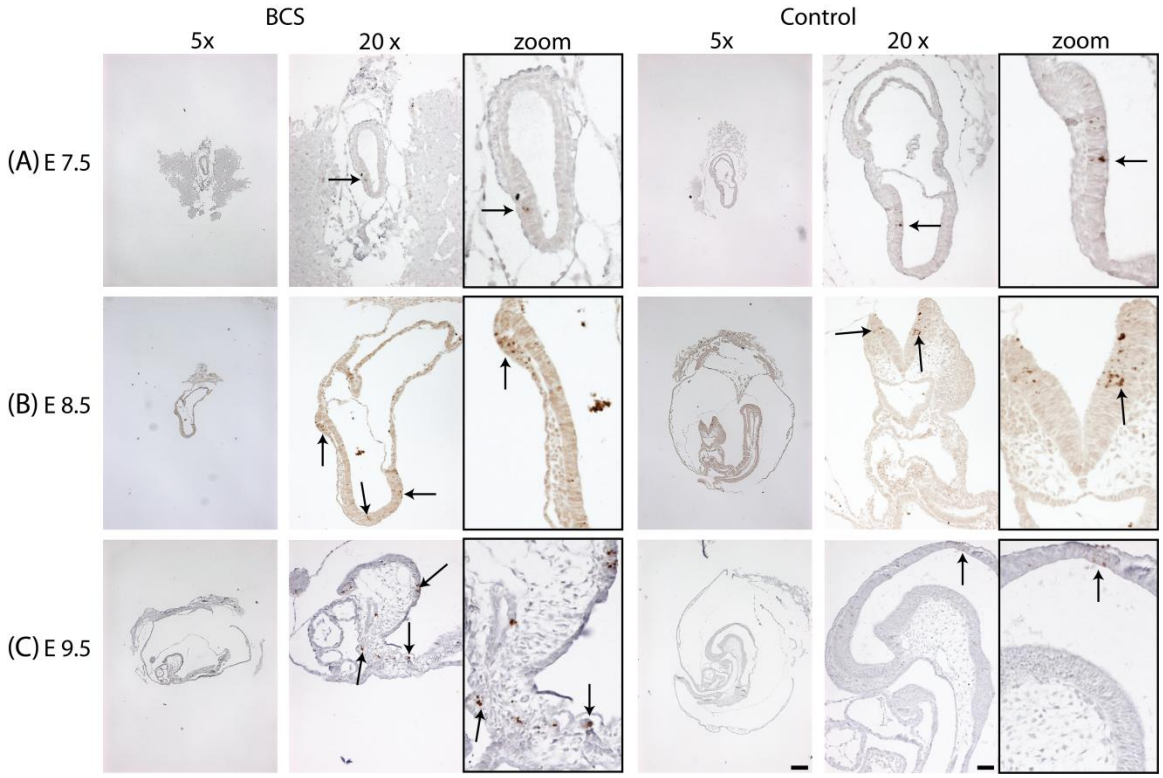
6.6.7 Cell death in homozygous mutant embryos

The small size of the BCS embryos may be explained in part by a reduction in cell proliferation rates, but may also be accompanied by increased cell death. To investigate

this possibility, we assessed apoptosis using an antibody against cleaved caspase 3 at E7.5- E11.5. During apoptosis, procaspase 3 is cleaved as part of a signaling cascade, activating it and propagating the cell death signal. Apoptosis is a normal and essential process in development, for example, in the midline of the recently-closed neural tube at E9.5, and to delineate and form toes during paw development[198,199]. In control embryos, cleaved caspase 3 labelling was concentrated in areas of expected apoptosis (for example, in the folding cranial neural tube at E8.5, see control in Figure 6.15 B). However, caspase staining in BCS embryos was frequently more widespread and not confined to areas where apoptosis is known to occur in healthy embryos, indicating that reduced cell proliferation is accompanied by increased apoptosis. The increased number of apoptotic cells could eventually lead to the death of the embryo, or result in hypoplasia of certain cell populations or tissues.

6.7 Discussion

Since a complete knockout of *Emg1* was preimplantation lethal, precluding further study of the embryos[122], we sought to establish a mouse model which more closely reflected human BCS. Mice homozygous for the mutation were expected to allow the study of the effects of BCS throughout embryonic development and in tissues that were previously unavailable in BCS patients, including the brain. When heterozygous *Emg1*$^{A/G}$ mice were intercrossed, live homozygous mutant "BCS" mice were not seen upon genotyping, and it became apparent that mice homozygous for the mutation in *Emg1* died prenatally. By examining litters of timed pregnant females, the age of embryonic death was established to be between E8.5 and E12.5, accompanied by severe intrauterine growth retardation


BCS
Control
5x
20 x
zoom
5x
20 x
zoom
(A) E 7.5
(B) E 8.5
(C) E 9.5

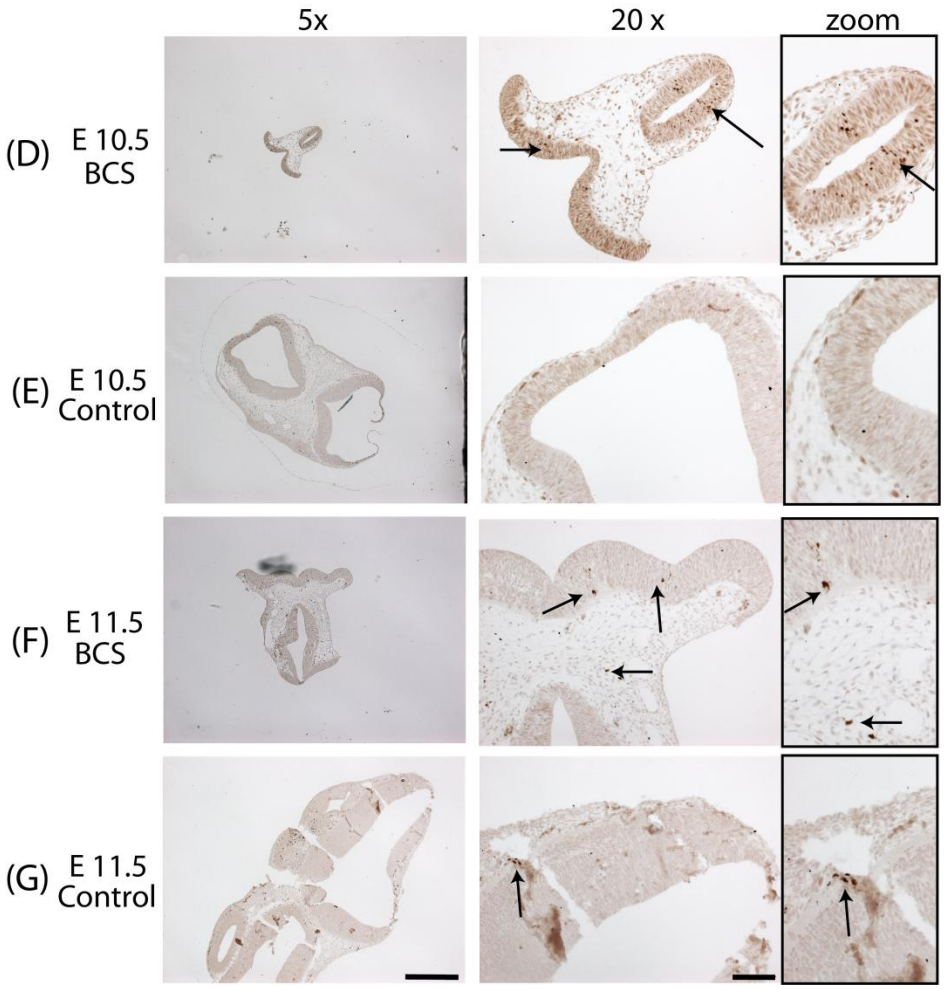


Figure 6.15 Cleaved caspase 3 in BCS embryos.

(A-C) Apoptotic cells were labeled in mid-sagittal sections using an anti-activated caspase 3 antibody in formaldehyde-fixed, paraffin-embedded embryos from E7.5 to E9.5. Scale bars at 5x = 200 µm, 20x = 50 µm. Arrows indicate areas of strong staining, such as in the hinge of the closing neural tube of E8.5 control embryos (B). **(D-G)** Cleaved caspase 3 was detected in transverse sections through the top of the head of embryos at E10.5 and E11.5, to show apoptosis in the neural tube (similar to the section plane pictured in 6.12 A and B). Note more generalized staining in BCS embryos. Scale bars at 5x = 500 µm, 20x = 100 µm.

and failure of neural tube closure. Death was apparently caused by a reduction in cell proliferation concomitant with increased apoptosis. Therefore, the BCS mouse model described in this chapter recapitulates the severe developmental delay seen in human BCS patients, and confirms that the mutation in *Emg1*causes a defect in the development of the central nervous system.

To better define the role of EMG1 in BCS, EMG1 expression throughout mouse development and into adulthood was examined. We found that, as expected for an essential protein with a function linked to ribosomes, it was ubiquitously expressed. During early development, EMG1 expression was uniform throughout the embryo, though the level of expression diminished and became more restricted over time. In the adult, tissues with the highest expression were, in general, also highly proliferative tissues, such as cells of the intestinal crypts (Figure 6.5 E). Logically, cells which are rapidly proliferating would have an increased need for ribosomes, and therefore an increased need for EMG1. However, EMG1 was also found in post-mitotic neurons of the retina and the brain (Figure 6.5 D and Figure 6.6). Neural cells such as those found in the ganglion cell layer of the retina, and Purkinje neurons of the cerebellum, are nevertheless highly metabolically active due to their need to mediate signal transduction between large numbers of neurons, requiring numerous ribosomes[151]. Although EMG1 was primarily localized to the nucleus in most tissues, some tissues such as liver appeared to also have a cytoplasmic localization. As discussed in Chapter 4, changes in EMG1 sub-cellular localization could reflect a difference in function or in binding partners in

different tissues. Biochemical fractionation of EMG1 in different tissues could be used to determine if EMG1 is preferentially associated with different sub-cellular compartments.

EMG1 expression in the brain was of particular interest due to the CNS defects in BCS patients. Expression of EMG1 in mouse brain was broadly similar to that in the human brain (see Chapter 3). In the embryonic central nervous system, EMG1 expression was strong and uniform in the neuroepithelium of early embryos, and became more specifically localized during development (Figure 6.6). Proliferative cells in the developing brain are found primarily in ventricular and subventricular zones; however EMG1 staining did not correlate exclusively in these areas. For example, in the brain of E16.5 embryos, EMG1 staining was strongest in the superficial layer of the cortex in the forebrain, which would be expected to contain post-mitotic neural cells which migrated there after being born in the ventricular or subventricular zone. Although these cells are no longer proliferative, they are still metabolically active, differentiating and extending processes. In the adult mouse brain, strong levels of EMG1 were found in the Purkinje neurons and the granule cell layer of the cerebellum. The cerebellum is largely responsible for motor control[200], and a defect in EMG1 in these cells during cerebellar development could result in a hypoplastic cerebellum, and contribute to the lack of psychomotor control displayed by BCS patients. In addition, the dentate gyrus of the hippocampus showed strong EMG1 staining in adult mouse brains. Neurons of the granule cell layer of the hippocampal dentate gyrus have been shown to be proliferative in adulthood, and contribute to the production of new neurons in the adult brain even in humans[201-204]. The strong levels of EMG1 detected in this area may reflect a small

population of proliferative neural progenitor cells. Since EMG1 expression in the developing brain was broad, it is plausible that a reduction of EMG1 levels, due to the instability caused by the D86G substitution, could have severe deleterious effects on brain development and function. A mouse model of Diamond-Blackfan anemia, caused by mutations in the small ribosomal subunit protein RPS7, resulted in significant reduction of the size of the cerebral cortex and the hippocampus[87]. The brain defects were shown to be the result of increased TRP53-mediated apoptosis during development of the central nervous system, and the hippocampal defect was underlined by behavioural tests showing deficits in working memory. These results are suggestive of an important role for ribosomes in the development of the brain, including the cerebral cortex and the hippocampus. Due to the embryonic death of BCS embryos, it is difficult to assess the contribution of EMG1 during brain development. A brain-specific *Emg1* knockout could be used to examine the role of EMG1 during neuronal development without the confounding effects of EMG1 deficiency in the entire organism.

Given the ubiquitous expression of EMG1, it is not surprising that the growth retardation in $Emg1^{G/G}$ embryos was uniform throughout the embryo from the earliest time point examined (Figure 6.11). It is likely that growth delay is present from the two-cell stage, when zygotic transcripts begin to take over from the maternal supply. Until the neural tube defect became apparent, there were no overt morphological defects in any particular tissue. The appearance of BCS embryos was consistent between litters: embryos homozygous for the mutation in *Emg1* were smaller in size than even the smallest control littermates, and were morphologically very similar to normal controls approximately one

day younger. BCS embryos which survived beyond E9.5 turned normally, and organogenesis in the oldest embryos examined appeared normal. Death occurred between E8.5 and E12.5, but the equivalent mutation in *EMG1* in humans is not known to be lethal until after birth, with some children surviving several years beyond the average age of death of 13 months[3]. Some Hutterite women with BCS-affected children are reported to have had spontaneous abortions[1,2], however the literature is not comprehensive on the subject and it seems unlikely that D86G EMG1 causes many cases of prenatal death in humans. The differences seen between BCS in humans and mice may reflect the high rates of proliferation necessary in mouse embryonic development[205]. Lethal mutations in mouse embryos between E8.5 and E12.5 often affect the initiation or maintenance of blood circulation in the yolk sac, or disrupt placental formation[206]. Indeed, ribosome biogenesis defects are frequently associated with hematopoietic defects[90,140,141]. However, blood islands were visible in the yolk sac of BCS embryos, and pericardial swelling, which is associated with placental defects, was not observed. In addition, chorioallantoic fusion was seen in E9.5 BCS embryos, indicating that placental formation was not disturbed at this stage (Figure 6.11 D). It therefore seems unlikely that embryonic death was caused by a disruption of yolk sac blood circulation or placental formation, although the placentas of the BCS embryos surviving to later stages were not examined microscopically. Moreover, a placental defect cannot entirely account for the growth retardation, since developmental delay was observed from E7.5, long before placental formation takes place (Figure 6.11 A). However, hematopoiesis in BCS embryos should be examined further to determine progenitor cell function, for example using a hematopoietic colony formation assay[101].

Notably, approximately half of all *Emg1*$^{G/G}$ embryos which were otherwise past the stage at which the neural tube should close, displayed an open cranial neural tube (that is, BCS embryos at E10.5 to E12.5 which had the appearance of normal control embryos at E9.5 to E11.5). In the mouse, the neuroepithelium begins to differentiate during gastrulation, forming the neural plate. The edges of this neural plate begin to fold and elevate, eventually meeting at the midline to close and form the neural tube. At the site of adhesion, epithelial remodeling involving apoptotic cell death completes the process. Formation of the neural tube involves three initial closure points (reviewed in[197,207-209]). At approximately E8.0 to E8.5, closure 1 takes place at the hindbrain-cervical boundary, followed by closure 2 in the future midbrain, though the position of this closure varies between mouse strains[210]. The final cranial neural tube closure, closure 3, takes place in the future forebrain, at the rostral extremity of the neural tube. The neural tube zips shut bidirectionally between these sites. Spinal neural tube formation continues after the cranial neural tube is shut, and is complete by approximately E10.0. Failure of neural tube closure at site 1 results in an open neural tube along the entire midline from the midbrain to the lower spine, called craniorachischisis. On the other hand, failure of closure at sites 2 or 3 results in an open neural tube in the cranial region only, called exencephaly, where the developing brain is open to the external environment. In the BCS embryos, the neural tube defect affected the cranial region only, therefore closure 1 likely took place as usual. The *Emg1* mutation either affected sites 2 or 3, or prevented subsequent closure between them. Disruption of neural tube closure can have many causes, including failure of neural fold elevation or fusion. The neuroepithelium is very

highly proliferative, and neuroepithelial cells exit the cell cycle and begin to differentiate only after the neural tube has closed[197]. In addition, increased proliferation in the mesenchyme subjacent to the neuroepithelium is necessary for elevation of the neural folds[208]. Given the reduced BrdU labelling in BCS embryos, the neural tube defect may simply be a reflection of inadequate numbers of proliferating cells in the neuroepithelium and cranial mesenchyme required to extend and close the neural tube folds. A second hypothesis involves the disruption of actin microfilaments during neural fold bending, which can cause exencephaly[163,211]. If a reduction in EMG1 levels occasions a delay in ribosome biogenesis, it could therefore reduce protein synthesis rates, disrupting the rapid formation of actin microfilaments required in the developing neural tube. Interestingly, defects in the actin cytoskeleton appear to play a more critical role in the cranial neural tube than in the spinal neural tube[211]. It should be noted, however, that failure of neural tube closure cannot be considered the cause of death in BCS embryos, since embryos with exencephaly and craniorachischisis can survive until birth[197,207-209].

A closer examination of the neural tube in transverse sections revealed an unusual cross shape in BCS embryos at E10.5 and E11.5, but only at approximately the level of the second pharyngeal arch (Figure 6.12 C and D). This may correspond to the position of the initial closure of the neural tube, suggesting aberrant neural tube closure at all three cranial closure sites. Closure at site 1 might occur, albeit abnormally, causing a misshapen tube, while closure at sites 2 and 3 is delayed or absent. Neural tube patterning is controlled at several levels by different signalling pathways. Patterning along the rostral-caudal axis is controlled by the differential expression of the *Hox* genes[200], and

disruption can result in homeotic transformations, which were not observed in BCS embryos. Dorsal-ventral patterning of the neural tube is controlled by counteracting gradients of sonic hedgehog (SHH) ventrally and bone morphogenetic proteins (BMPs) dorsally[200,212]. Disruption of these gradients can result in dorsalization or ventralization of the neural tube, with disorganization of the associated neurons (D1-D6 interneurons for BMP, and V0-V3 interneurons and motor neurons for SHH). Further experiments are necessary to determine if the BMP or SHH gradients are disturbed, and if this is reflected along the entire neural tube, or just at the sites of abnormal morphology. If dorsal-ventral patterning of the neural tube is misregulated in BCS embryos, formation or organization of motor neurons could be disrupted, contributing to the lack of motor control seen in BCS patients. These results indicate that the mutation in *Emg1* has profound deleterious effects on the development and the patterning of the central nervous system in the mouse.

Analysis of cell proliferation and cell death in BCS embryos also showed a somewhat variable phenotype, even between BCS embryos from the same litter (see Figure 6.14 C and D). At E8.5, BrdU labelling was largely similar in all embryos examined, even though BCS embryos were already smaller than control littermates. Proliferation rates at this stage of development are extremely high, and it may be possible that, in order to detect a reduction in BrdU labelling, a BrdU injection time shorter than the two hours used here would be necessary. Notably, BCS embryos frequently showed markedly reduced BrdU labelling in the allantois, a structure which contributes to the formation of the umbilical cord and the placenta (compare control allantois at E7.5 with BCS at E8.5 and E9.5, A-C). By E9.5, $Emg1^{G/G}$ embryos were found which showed very little BrdU

labelling at all (D), while others still had strong labelling (C). This pattern mirrors the variable age at death, making it likely that the reduced cell proliferation contributes to embryonic death. Conversely, an examination of the apoptotic marker cleaved caspase 3 revealed very little labelling in putative BCS embryos at E7.5, but more generalized labelling at E8.5 and E9.5 (Figure 6.15). Although cleaved caspase 3 labelling was more widespread in BCS embryos, the number of stained cells was still relatively low, making it unlikely that apoptosis alone can account for embryonic death. Taken together, it is possible that the combined effects of reduced proliferation and increased apoptosis reduce the available cell population until the embryo is no longer viable. Further analysis is necessary to establish if specific cell populations are affected. In mouse models of Treacher Collins syndrome, mutations in the ribosome biogenesis gene *Tcof1*result in reduced proliferation specifically in migrating cranial neural crest cells, leading to craniofacial hypoplasia[94]. This was determined by intercrossing the *Tcof1* mutants with mice expressing GFP in neural crest cells, enabling isolation and quantification of neural crest cell numbers. In the BCS embryos, evidence of apoptosis was observed in both the neuroepithelium and the cranial mesenchyme, though it remains to be determined if specific cell types are affected. Additionally, TRP53 plays an important role in apoptosis and has been shown to link the cell cycle with ribosome biogenesis[59,79,84]. Indeed, the phenotypes of many animal models of ribosome biogenesis disorders can be rescued by eliminating the TRP53 checkpoint[95,105,123,124,213]. Levels of TRP53 in BCS embryos have not yet been examined, and future experiments should explore its contribution to the phenotype. If TRP53 or its downstream targets are found to be increased in *Emg1*$^{G/G}$ embryos, it may be possible to extend survival by ablating TRP53, either genetically or

through the use of a drug such as pifithrin-α, which was used to rescue the phenotype of Treacher Collins syndrome mice[105].

Taken together, these data clearly demonstrate that the A>G mutation in *Emg1* has severe deleterious effects on organism development. Interestingly, the mutation in mice recapitulates the severe developmental delay and marked defects in the development of the central nervous system seen in BCS in humans. Although the phenotype is more severe in mice than in humans, this model is currently the best tool available for studying the effects of the A>G mutation in *Emg1* during development. The variable phenotype in BCS embryos, including the age of death, is common in ribosomopathies in humans. The phenotype of Treacher Collins syndrome is highly variable, even between siblings with the disease, despite high penetrance of the mutations responsible[93,214]. Patients with cartilage hair hypoplasia, caused by mutations in *RMRP*, have highly variable clinical manifestations, some displaying no obvious phenotype at all[215]. In addition, a variable phenotype is not unexpected on a mixed *CD1/C129/C57BL/6* background, as the genetic background of mice can have a profound influence on the phenotype of a mutation[101,216,217]. Efforts to produce a congenic *C57BL/6* line are underway, by backcrossing male mice heterozygous for the mutation with female *C57BL/6* mice. This may reduce the variable penetrance of the phenotype and perhaps increase survival. Embryos resulting from intercrosses at five and six backcrossings were examined and showed a very similar phenotype when compared with embryos on the mixed *CD1/C129/C57BL/6* background (not shown). When the congenic strain is established, further analysis will be necessary to determine if the phenotypic variability is improved.

A more in-depth biochemical analysis of the effects of the EMG1 defect, including its influence on ribosome biogenesis similar to what was described in Chapter 5, will be possible once *Emg1*$^{G/G}$ cell lines are established, either from mouse embryonic fibroblasts or from ES cells. These cell lines can initially be used to confirm *Emg1* mRNA levels and EMG1 protein levels. Analysis of cell proliferation rates, protein synthesis rates, ribosome subunit ratios, and rRNA processing should be undertaken to definitively establish that a ribosome biogenesis defect underlies the developmental delay. In the absence of tissue-specific cell lines derived from BCS patients, the mouse ES cell line can be differentiated into any cell type of interest given the proper conditions, allowing comparative studies of ribosome biogenesis in multiple tissue types. Moreover, the mouse ES cell line has the added advantage of *loxP* sites surrounding exons 2-6 of *Emg1*. It is therefore possible to further reduce the levels of EMG1 by expressing Cre recombinase and excising exons 2-6 (*Emg1*$^{G/-}$), allowing the study of ribosome biogenesis defects in cultured cells expressing varying levels of EMG1, and the exploration of a possible dose-dependent effect.

The establishment of a mouse model of BCS opens the door for testing treatments for the disease. Although the phenotype is embryonic lethal, it is still possible to monitor litters for improved survival of homozygous knockin embryos following treatment. One possible treatment is S-adenosyl methionine, the methyl donor in the methyltransferase reaction performed by EMG1. Excess exogenous SAM or SAM synthase rescues an EMG1 mutant in yeast, likely by stabilization of the protein upon SAM binding[114]. Since *EMG1* is transcribed at normal levels in BCS patient cells[147], if SAM stabilizes the D86G

EMG1 protein and prevents its aggregation or degradation, normal levels of EMG1 could be restored. SAM is already approved for use as an anti-depressant in humans[218,219], meaning that it would be readily available to Hutterite families if it showed promise in mouse studies.

Chapter 7 : Conclusions and future directions

Bowen-Conradi syndrome is a lethal disorder for which there is currently no cure or even treatment. The carrier frequency is high amongst Hutterites, at an estimated one in ten[3,125], making it a significant concern for this population. The discovery that a c.257A>G, p.D86G mutation in *EMG1* was responsible for the disease, described in Chapter 3, allowed accurate diagnosis of BCS and carrier testing for prospective parents. The biochemical characterization of EMG1 in Chapters 4 and 5 established that the mutation may have different effects in different cell types, and confirmed that the D86G substitution adversely affects ribosomal RNA processing. Therefore, BCS belongs to the ribosomopathies, which are individually rare but as a group affect many. Research into the molecular causes of any one of these diseases sheds light on all of them, and recent insights into the variable tissue specificity in ribosomopathies may be valuable for BCS. The mouse model described in Chapter 6 further supports the conclusion that the D86G substitution in EMG1underlies BCS, and provides a new tool for studying both the etiology of the disease and treatments which may one day be useful in humans.

Taken together, the results presented in this thesis depict EMG1 as an essential protein, with potentially multiple roles not only throughout development and in different tissues, but even throughout the cell cycle. The conclusions about EMG1 function discussed in Chapters 4 and 5 may be applied to the mouse model of BCS and extrapolated to the human disease. The D86G substitution makes the protein unstable, and it is either degraded by the proteasome or aggregates due to the exposure of hydrophobic residues, which may affect its association with interacting partners. Contrary to previous reports[72], we concluded that EMG1 sub-cellular localization is unaltered by the A>G mutation in

BCS patient fibroblasts and lymphoblasts. Intriguingly, sub-cellular localization was found to be dynamic in HeLa cells, possibly regulated by NOP14. In fibroblasts, on the other hand, EMG1 was localized exclusively to the nucleolus and to nucleoplasmic foci. Although HeLa cells are a cancer cell line and "normal" functions should be extrapolated from them with caution, this was the first indication that EMG1 may play diverse roles in different cell types.

Evidence for different effects of the D86G EMG1 substitution was also found between BCS patient fibroblasts and lymphoblasts. Even though a delay in 18S rRNA processing was evident in both cell types, this resulted in a reduction of proliferation rates due to cell cycle delay at G2/M only in the lymphoblasts. Given these results, it seems likely that an examination of other cell types would reveal further differences. Since it is difficult to obtain tissue samples from BCS patients, it would be possible to derive induced pluripotent stem cells from the already-available fibroblasts. These could then be differentiated into numerous cell types of interest, including neuronal cells to study the effect of the D86G substitution on EMG1 function in the brain. In a complementary approach, ES cell lines should be isolated from BCS and control mice, and used to generate cell types of interest, which would lead to a better understanding of EMG1 functions in different cell types. Furthermore, the contribution of TP53 to the phenotype would be clarified by determining if TP53 preferentially accumulates in specific cell types harbouring D86G EMG1.

Extending the hypothesis that EMG1 has diverse functions in different cell types, multiple tissues may be affected differently by the D86G EMG1 substitution in developing mice and humans. One of the most intriguing aspects of ribosomopathies is the tissue specificity shown in these disorders. One possible explanation by Kondrashov *et al*[97] showed that large ribosomal protein RPL38 is expressed at different levels in different mouse tissues, and that mutations in RPL38 preferentially affect the tissues in which it is most highly expressed. Following this line of thought, BCS is primarily a neurological disorder, and thus we would expect that EMG1 is most highly expressed in the central nervous system. This does not appear to be the case in the mouse, with strong expression in the developing neuroepithelium but equally high levels throughout the entire embryo. However, the qualitative immunohistochemistry results presented here could be improved by using quantitative microarray analysis of *Emg1* expression in different tissues throughout development and in the adult mouse. This would give much more precise data, and could provide target tissues for future analysis in humans. In the adult mouse, strong expression was found in highly proliferating tissues such as testes and intestinal crypts. The fact that EMG1 is strongly expressed in intestine raises the interesting possibility that the growth delay in BCS patients may be caused, at least in part, by an inability to absorb nutrients due to an underdeveloped intestine. BCS patients do not readily gain weight after birth, even upon tube feeding. This question awaits histological examination of BCS patient intestine following autopsy. However, it should be noted that this hypothesis cannot completely explain the tissue specificity of ribosomopathies, since high levels of expression are sometimes found in unaffected tissues. For example, mutations in RPL38 did not affect development of the kidney,

where it is strongly expressed[97]. It may be possible that RPL38 expression in mouse kidney outstrips demand, such that a reduction in RPL38 levels is still sufficient to meet demand during development. Further analysis should determine which tissues are most sensitive to changes in *EMG1* expression, as opposed to simply determining which tissues have the highest expression. Indeed, a tissue with low levels of EMG1 might be more susceptible to an EMG1 defect than one in which EMG1 is produced in excess of demand.

Expression of SBDS, the protein responsible for the ribosomopathy Shwachman-Diamond syndrome, is very similar in mice to that of EMG1, with ubiquitous expression during development, and high levels in proliferative tissues in the adult[102]. A recent study performed in zebrafish on the effects of mutations in *sbds* recapitulated the pancreas-specific effects seen in human patients[109]. Shwachman-Diamond syndrome patients suffer from exocrine pancreatic insufficiency, and it has never been clear whether the symptoms result from reduced protein synthesis rates in the pancreas, leading to insufficient levels of pancreatic enzymes, or whether the pancreas simply does not develop properly. The zebrafish study suggested a proliferation defect in pancreatic precursor cells during development, a defect which precedes the exocrine deficiency. Therefore, mutations in SBDS do not solely affect the production of digestive enzymes as a direct result of reduced protein synthesis rates, but instead occasion a reduction of mature pancreatic exocrine cells. This suggests a much more complex role for ribosome biogenesis during development, beyond a simple need for increased protein synthesis rates. It also underlines the exquisite specificity of ribosome biogenesis defects, in this case affecting

exocrine but not endocrine pancreatic function. How the mutations in SBDS cause pancreas insufficiency, while mutations in EMG1 cause severe CNS deficiencies, remains an open question.

Cell cycle analysis revealed an unexpected delay in G2/M in BCS patient lymphoblasts, leading to a proliferation defect. It is plausible that this scenario holds true in the developing embryo as well. The reduction in BrdU staining in BCS mouse embryos could be the result of a lengthened cell cycle caused by a G2/M block. The cell cycle delay was not, however, found in fibroblasts, and cell types in the embryo may also be differently affected. If, for example, the cell cycle of neural progenitor cells were preferentially perturbed, a reduction in proliferation could lead to hypoplasia in the brain. BCS mouse embryos probably do not survive long enough to observe hypoplasia in a particular area of the brain or in a particular cell type as cortical neurogenesis lasts from E11 to E19[220,221], but an inducible, brain-specific knockout mouse model could be generated and used to evaluate the effects of EMG1 perturbation at different stages of neurogenesis. For example, an inducible Cre recombinase under the control of the Nestin promoter (e.g. Jackson labs strain C57BL/6-Tg(Nes-cre/Esr1*)1Kuan/J)[220], could be employed to ablate *loxP*-flanked *Emg1* coding sequence in a transgenic mouse. In addition, BCS patient brain tissue taken upon autopsy is available and could be used to determine if hypoplasia is present in any particular cell type or area of the brain. Based on patient data and EMG1 staining, the cerebellum is a good candidate, since EMG1 was found in both mouse and human cerebellum, and cerebellar involvement has been reported in cases of BCS[2,3]. It is of relevance that ribosome biogenesis defects often alter proliferation rates of

progenitors, without affecting their capacity to differentiate[94,109-111]. An examination of BCS patient or mouse brain would require careful analysis to determine whether reduced proliferation, as opposed to impaired differentiation, underlies the neurological phenotype in BCS.

The cause of cell cycle delay in G2/M is yet to be determined. Accumulation of TP53 is well known to cause cell cycle delay at the G1 checkpoint, and can induce arrest at the G2/M checkpoint as well[222-224]. Accumulation of TP53 in BCS patient cells and of TRP53 in the BCS mouse has not yet been found. However, the cell cycle is also controlled by TP53-independent pathways (reviewed in[225]), as evidenced in *S. cerevisiae*, which does not have a TP53 homologue. As discussed in Chapter 5, the G2/M growth arrest could be mediated by the defective translation of an IRES-containing isoform of cyclin-dependent kinase 11 required for progression through M phase[186,187]. If this is true, then the translation of other specific mRNAs may be affected in BCS as well. For an unbiased approach to identify changes in translation, polysome analysis, similar to what was done in dyskeratosis congenita patient and mouse cells[98], should be performed. Briefly, ribosomes are separated by sucrose gradient, mRNA is isolated from the fraction of intact ribosomes associated with mRNAs, microarray analysis is performed, and up- or down-regulation of candidate proteins are verified by immunoblot. This would identify mRNAs which are differentially translated in BCS cells compared with controls, offering a more accurate picture than microarray analysis alone. If any IRES-containing mRNAs are identified, a dicistronic *Renilla* luciferase/ firefly luciferase reporter assay could be used to determine if global IRES-regulated translation is deficient in BCS patient or

mouse cells. Following along these lines, the generalized apoptosis found in BCS mouse embryos could be downstream of impaired translation of anti-apoptotic mRNAs which contain IRES elements, including BCL2-like 1 and X-linked inhibitor of apoptosis. Both of these are down-regulated in dyskeratosis congenita patient lymphocytes[98]. Defects in rRNA modification have previously been shown to have subtle effects on ribosome function. For example, specific inhibition of pseudouridylation of m1acp3ψ in yeast results in a reduction in ribosome fidelity[68]. Lack of methylation at m2G966 of *E. coli* rRNA, the bacterial equivalent to m1acp3ψ, is phenotypically silent but results in reduced cell fitness in a competition assay with wild type *E. coli*[226]. If dyskeratosis congenita and BCS both share defects in the translation of IRES element-containing mRNAs, it would indicate that rRNA modification plays a major role in conferring ribosome specificity.

In the long term, the mouse model of BCS described in Chapter 6 can be used to evaluate possible treatments for BCS. As a first step to determine if SAM would be useful to treat BCS, it should be determined if SAM can stabilize D86G EMG1. BCS patient cells can be incubated with SAM, and immunoblot analysis can be used to determine if EMG1 levels are increased in the presence of the drug. Increased levels of EMG1 should improve the cell cycle and cell proliferation defect in BCS patient lymphoblasts. If SAM can stabilize D86G EMG1 *in vitro*, experiments can then be performed to evaluate the effect of injecting SAM into pregnant mice to stabilize the mutant protein in developing embryos. At a minimum, increased age of survival *in utero* would be very valuable for evaluation of the phenotype in developing tissues beyond E12.5, and the birth of viable pups would be very promising for SAM treatment in humans. Notably, heterozygous

mutants are indistinguishable from wild type mice, meaning that a partial rescue of EMG1 function may be sufficient to increase survival and reduce the growth defect. Furthermore, if it is determined that TRP53 is significantly increased in BCS mice, even in certain tissues, BCS mice may be crossed with TRP53 knockout mice or treated with the drug pifithrin-α in an effort to bypass the TRP53 checkpoint and ameliorate survival rates.

In summary, the results presented in this thesis are the starting point for future examination of the intricacies of BCS. It will be important to determine if the D86G EMG1 substitution causes changes in the specificity of ribosome translation, as this will provide new insights about ribosome function. With that information in hand, it may become possible to develop treatments for diseases affecting ribosome biosynthesis and function.

Beyond ribosomopathies

Several diseases affecting the biosynthesis and assembly of mitochondrial ribosomes[227-229] fall under the category of mitochondrial disease since the outcome is a defect in the mitochondrial respiratory chain. Mitochondrial ribosomes differ significantly from cytoplasmic ribosomes due to their separate evolution[229]. The small subunit consists of the 12S rRNA and 29 proteins, while the large subunit consists of the 16S rRNA and 48 proteins. One reported case of mitochondrial disease discussed a patient with low birth weight, dysmorphic facial features, and agenesis of the corpus callosum[228]. This was

found to be caused by a mutation in the mitochondrial small ribosomal subunit protein gene *MRPS16*, leading to reduced production of the 12S rRNA and mitochondrial-specific protein synthesis impairment. The authors noted that the brain phenotype was not likely a result of reduced energy throughout the organism, but rather a tissue-specific effect. This may indicate that the disease was caused not only by a respiratory chain defect, but additional ribosome-related deficiencies. Importantly, mitochondrial ribosomal proteins have been implicated in apoptosis and cell cycle progression[230]. The fact that many of these diseases appear to display tissue-specific defects[231,232], and that mitochondrial ribosomes can be heterogeneous depending on which isoform of mitochondrial ribosome proteins are used[229], suggests that mitochondrial disease and ribosomopathy research have information to offer each other.

Disorders of protein synthesis, caused by mutations in initiation, elongation, and termination factors[233], may also have some light to shed on ribosomopathies. Protein synthesis disorders paradoxically display tissue specificity despite affecting a ubiquitous process, much like ribosomopathies. Indeed, Vanishing White Matter disease (VWM, OMIM #603896), caused by mutations in subunits of translation initiation factor eIF2B, affects only the central nervous system, reminiscent of BCS[234]. Other diseases which are not usually grouped with the ribosomopathies may also shed light on ribosome function and regulation. For instance, the autosomal recessive neurological disorder cystic leukoencephalopathy (OMIM #612951) is caused by a deficiency in RNASET2, a lysosomal ribonuclease responsible for degrading ribosomal RNA[235]. Defects in RNASET2 lead to progressive neurological defects including psychomotor impairment,

seizures, sensorineural hearing impairment, and microcephaly, due to the accumulation of RNA in neurons[235,236]. Therefore, defects in ribosome turnover, in addition to ribosome biogenesis, can lead to human disease. Defects in ribosome function have also been implicated in early stages of Alzheimer disease[237] (OMIM #104300). More broadly, many ribosome biogenesis factors are found to be altered in various cancers[31,238]. RBM28, which is associated with alopecia, neurological defects and endocrinopathy syndrome, is upregulated in oral squamous cell carcinoma[239], levels of the small subunit processome protein (and EMG1 binding partner) NOP14 are increased in pancreatic cancer[240], and mutations in eleven different ribosomal protein genes cause malignant nerve sheath tumors in the zebrafish peripheral nervous system[241]. Clearly, the ribosome and its proper function underlie diverse processes, many of which are incompletely understood, and research to tease apart the complex regulation of ribosomes will reveal details about many human disorders.

References

1. Bowen, P. & Conradi, G.J. Syndrome of skeletal and genitourinary anomalies with unusual facies and failure to thrive in Hutterite sibs. *Birth Defects Orig.Artic.Ser.* **12**, 101-108 (1976).
2. Hunter, A.G. *et al.* The Bowen-Conradi syndrome -- a highly lethal autosomal recessive syndrome of microcephaly, micrognathia, low birth weight, and joint deformities. *Am.J.Med.Genet.* **3**, 269-279 (1979).
3. Lowry, R.B. *et al.* Bowen-Conradi syndrome: a clinical and genetic study. *Am.J.Med.Genet.* **120A**, 423-428 (2003).
4. Pena, S.D. & Shokeir, M.H. Autosomal recessive cerebro-oculo-facio-skeletal (COFS) syndrome. *Clin Genet* **5**, 285-93 (1974).
5. Pena, S.D. & Shokeir, M.H. Syndrome of camptodactyly, multiple ankyloses, facial anomalies, and pulmonary hypoplasia: a lethal condition. *J Pediatr* **85**, 373-5 (1974).
6. Pena, S.D. & Shokeir, M.H. Syndrome of camptodactyly, multiple ankyloses, facial anomalies and pulmonary hypoplasia--further delineation and evidence for autosomal recessive inheritance. *Birth Defects Orig Artic Ser* **12**, 201-8 (1976).
7. Laziuk, G.I., IV, L.e., Ostrovskaia, T.I. & Cherstvoi, E.D. [Bowen-Conrad syndrome]. *Arkh.Patol.* **42**, 60-64 (1980).
8. Beemer, F.A. & Gerards, L.J. [A newborn infant with the Bowen-Conradi syndrome; a fatal hereditary disorder]. *Tijdschr.Kindergeneeskd.* **50**, 54-57 (1982).
9. Aynaci, F.M., Mocan, H., Erduran, E. & Gedik, Y. Hypospadias as a new congenital anomaly in Bowen-Conradi syndrome. *Genet.Couns.* **5**, 369-371 (1994).
10. Gupta, A. & Phadke, S.R. Bowen-Conradi syndrome in an Indian infant: first non Hutterite case. *Clin.Dysmorphol.* **10**, 155-156 (2001).
11. Mange, A.P. Growth and Inbreeding of a Human Isolate. *Hum Biol* **36**, 104-33 (1964).
12. Hostetler, J.A. History and relevance of the Hutterite population for genetic studies. *Am J Med Genet* **22**, 453-62 (1985).
13. Boycott, K.M. *et al.* Clinical genetics and the Hutterite population: a review of Mendelian disorders. *Am J Med Genet A* **146A**, 1088-98 (2008).
14. Chong, J.X. *et al.* A common spinal muscular atrophy deletion mutation is present on a single founder haplotype in the US Hutterites. *Eur J Hum Genet* **19**, 1045-51 (2011).
15. Frosk, P. *et al.* Limb-girdle muscular dystrophy type 2H associated with mutation in TRIM32, a putative E3-ubiquitin-ligase gene. *Am.J.Hum.Genet.* **70**, 663-672 (2002).
16. Lamont, R.E. *et al.* A locus for Bowen-Conradi syndrome maps to chromosome region 12p13.3. *Am.J.Med.Genet.A* **132A**, 136-143 (2005).
17. Graham, J.M., Jr. *et al.* Cerebro-oculo-facio-skeletal syndrome with a nucleotide excision-repair defect and a mutated XPD gene, with prenatal diagnosis in a triplet pregnancy. *Am J Hum Genet* **69**, 291-300 (2001).

18. Meira, L.B. *et al.* Manitoba aboriginal kindred with original cerebro-oculo-facio-skeletal syndrome has a mutation in the Cockayne syndrome group B (CSB) gene. *Am J Hum Genet* **66**, 1221-8 (2000).
19. Laugel, V. *et al.* Cerebro-oculo-facio-skeletal syndrome: three additional cases with CSB mutations, new diagnostic criteria and an approach to investigation. *J Med Genet* **45**, 564-71 (2008).
20. Eichler, D.C. & Craig, N. Processing of eukaryotic ribosomal RNA. *Prog Nucleic Acid Res Mol Biol* **49**, 197-239 (1994).
21. Fromont-Racine, M., Senger, B., Saveanu, C. & Fasiolo, F. Ribosome assembly in eukaryotes. *Gene* **313**, 17-42 (2003).
22. Henras, A.K. *et al.* The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci.* **15**, 2359 (2008).
23. Ban, N., Nissen, P., Hansen, J., Moore, P.B. & Steitz, T.A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* **289**, 905-20 (2000).
24. Nissen, P., Hansen, J., Ban, N., Moore, P.B. & Steitz, T.A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920-30 (2000).
25. Harms, J. *et al.* High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* **107**, 679-88 (2001).
26. Laughrea, M., Latulippe, J., Filion, A.M. & Boulet, L. Mistranslation in twelve Escherichia coli ribosomal proteins. Cysteine misincorporation at neutral amino acid residues other than tryptophan. *Eur J Biochem* **169**, 59-64 (1987).
27. Rosenberger, R.F. & Foskett, G. An estimate of the frequency of in vivo transcriptional errors at a nonsense codon in Escherichia coli. *Mol Gen Genet* **183**, 561-3 (1981).
28. Kramer, E.B. & Farabaugh, P.J. The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. *RNA* **13**, 87-96 (2007).
29. Poehlsgaard, J. & Douthwaite, S. The bacterial ribosome as a target for antibiotics. *Nat Rev Microbiol* **3**, 870-81 (2005).
30. Ben-Shem, A., Jenner, L., Yusupova, G. & Yusupov, M. Crystal structure of the eukaryotic ribosome. *Science* **330**, 1203-9 (2010).
31. Luft, F. The rise of a ribosomopathy and increased cancer risk. *J Mol Med (Berl)* **88**, 1-3 (2010).
32. Narla, A. & Ebert, B.L. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**, 3196-205 (2010).
33. Dixon, J., Trainor, P. & Dixon, M.J. Treacher Collins syndrome. *Orthod Craniofac Res* **10**, 88-95 (2007).
34. Choesmel, V. *et al.* Impaired ribosome biogenesis in Diamond-Blackfan anemia. *Blood* **109**, 1275-83 (2007).
35. Ebert, B.L. *et al.* Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. *Nature* **451**, 335-339 (2008).
36. Ridanpaa, M. *et al.* Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell* **104**, 195-203 (2001).

37. Ganapathi, K.A. *et al.* The human Shwachman-Diamond syndrome protein, SBDS, associates with ribosomal RNA. *Blood* **110**, 1458-1465 (2007).
38. Heiss, N.S. *et al.* X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* **19**, 32-8 (1998).
39. Burger, K. *et al.* Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. *J Biol Chem* **285**, 12416-25 (2010).
40. Moss, T. DNA methyltransferase inhibition may limit cancer cell growth by disrupting ribosome biogenesis. *Epigenetics* **6**, 128-33 (2011).
41. McConkey, E.H. & Hopkins, J.W. The Relationship of the Nucleolus to the Synthesis of Ribosomal Rna in Hela Cells. *Proc Natl Acad Sci U S A* **51**, 1197-204 (1964).
42. Zentner, G.E., Saiakhova, A., Manaenkov, P., Adams, M.D. & Scacheri, P.C. Integrative genomic analysis of human ribosomal DNA. *Nucleic Acids Res* **39**, 4949-60 (2011).
43. Stults, D.M., Killen, M.W., Pierce, H.H. & Pierce, A.J. Genomic architecture and inheritance of human ribosomal RNA gene clusters. *Genome Res* **18**, 13-8 (2008).
44. Prokopowich, C.D., Gregory, T.R. & Crease, T.J. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**, 48-50 (2003).
45. Henderson, A.S., Warburton, D. & Atwood, K.C. Location of ribosomal DNA in the human chromosome complement. *Proc Natl Acad Sci U S A* **69**, 3394-8 (1972).
46. Moss, T., Langlois, F., Gagnon-Kugler, T. & Stefanovsky, V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cell Mol.Life Sci.* **64**, 29-49 (2007).
47. Ihara, M., Tseng, H. & Schultz, R.M. Expression of variant ribosomal RNA genes in mouse oocytes and preimplantation embryos. *Biol Reprod* **84**, 944-6 (2011).
48. Gagnon-Kugler, T., Langlois, F., Stefanovsky, V., Lessard, F. & Moss, T. Loss of human ribosomal gene CpG methylation enhances cryptic RNA polymerase II transcription and disrupts ribosomal RNA processing. *Mol Cell* **35**, 414-25 (2009).
49. Stefanovsky, V. & Moss, T. Regulation of rRNA synthesis in human and mouse cells is not determined by changes in active gene count. *Cell Cycle* **5**, 735-9 (2006).
50. Melese, T. & Xue, Z. The nucleolus: an organelle formed by the act of building a ribosome. *Curr Opin Cell Biol* **7**, 319-24 (1995).
51. Scheer, U. & Hock, R. Structure and function of the nucleolus. *Curr Opin Cell Biol* **11**, 385-90 (1999).
52. Handwerger, K.E. & Gall, J.G. Subnuclear organelles: new insights into form and function. *Trends Cell Biol* **16**, 19-26 (2006).
53. Sirri, V., Urcuqui-Inchima, S., Roussel, P. & Hernandez-Verdun, D. Nucleolus: the fascinating nuclear body. *Histochem Cell Biol* **129**, 13-31 (2008).
54. Denissov, S. *et al.* A model for the topology of active ribosomal RNA genes. *EMBO Rep* **12**, 231-7 (2011).

55. Fatica, A. & Tollervey, D. Making ribosomes. *Curr.Opin.Cell Biol* **14**, 313-318 (2002).
56. Hadjiolova, K.V., Nicoloso, M., Mazan, S., Hadjiolov, A.A. & Bachellerie, J.P. Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis. *Eur J Biochem* **212**, 211-5 (1993).
57. Ciganda, M. & Williams, N. Eukaryotic 5S rRNA biogenesis. *Wiley Interdiscip Rev RNA* **2**, 523-33 (2011).
58. Kos, M. & Tollervey, D. Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell* **37**, 809-20 (2010).
59. Bernstein, K.A., Gallagher, J.E., Mitchell, B.M., Granneman, S. & Baserga, S.J. The small-subunit processome is a ribosome assembly intermediate. *Eukaryot.Cell* **3**, 1619-1626 (2004).
60. Dragon, F. *et al.* A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**, 967-70 (2002).
61. Chow, C.S., Lamichhane, T.N. & Mahto, S.K. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. *ACS Chem Biol* **2**, 610-9 (2007).
62. Karijolich, J., Kantartzis, A. & Yu, Y.T. Quantitative analysis of RNA modifications. *Methods Mol Biol* **629**, 21-32 (2010).
63. Maden, B.E. The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog Nucleic Acid Res Mol Biol* **39**, 241-303 (1990).
64. Liang, X.H., Liu, Q. & Fournier, M.J. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* **15**, 1716-28 (2009).
65. Helm, M. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res* **34**, 721-33 (2006).
66. Ofengand, J. *et al.* Pseudouridines and pseudouridine synthases of the ribosome. *Cold Spring Harb Symp Quant Biol* **66**, 147-59 (2001).
67. Liang, X.H., Liu, Q. & Fournier, M.J. rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol Cell* **28**, 965-977 (2007).
68. Baudin-Baillieu, A. *et al.* Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy. *Nucleic Acids Res* **37**, 7665-77 (2009).
69. Maden, B.E., Forbes, J., de Jonge, P. & Klootwijk, J. Presence of a hypermodified nucleotide in HeLa cell 18 S and Saccharomyces carlsbergensis 17 S ribosomal RNAs. *FEBS Lett* **59**, 60-3 (1975).
70. Brand, R.C., Klootwijk, J., Planta, R.J. & Maden, B.E. Biosynthesis of a hypermodified nucleotide in Saccharomyces carlsbergensis 17S and HeLa-cell 18S ribosomal ribonucleic acid. *Biochem J* **169**, 71-7 (1978).
71. Youvan, D.C. & Hearst, J.E. A sequence from Drosophila melanogaster 18S rRNA bearing the conserved hypermodified nucleoside am psi: analysis by reverse transcription and high-performance liquid chromatography. *Nucleic Acids Res* **9**, 1723-41 (1981).

72. Meyer, B. *et al.* The Bowen-Conradi syndrome protein Nep1 (Emg1) has a dual role in eukaryotic ribosome biogenesis, as an essential assembly factor and in the methylation of {Psi}1191 in yeast 18S rRNA. *Nucleic Acids Res.* **39**, 1526-1537 (2011).
73. Phipps, K.R., Charette, J. & Baserga, S.J. The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip Rev RNA* **2**, 1-21 (2011).
74. Robledo, S. *et al.* The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *RNA.* **14**, 1918-1929 (2008).
75. Mayer, C. & Grummt, I. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**, 6384-91 (2006).
76. Reiter, A.K., Anthony, T.G., Anthony, J.C., Jefferson, L.S. & Kimball, S.R. The mTOR signaling pathway mediates control of ribosomal protein mRNA translation in rat liver. *Int J Biochem Cell Biol* **36**, 2169-79 (2004).
77. Grandori, C. *et al.* c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nat Cell Biol* **7**, 311-8 (2005).
78. Barna, M. *et al.* Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. *Nature* **456**, 971-5 (2008).
79. Pestov, D.G., Strezoska, Z. & Lau, L.F. Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: effects of nucleolar protein Bop1 on G(1)/S transition. *Mol.Cell Biol.* **21**, 4246-4255 (2001).
80. Rubbi, C.P. & Milner, J. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO J.* **22**, 6068-6077 (2003).
81. Chen, D. *et al.* Ribosomal protein S7 as a novel modulator of p53-MDM2 interaction: binding to MDM2, stabilization of p53 protein, and activation of p53 function. *Oncogene* **26**, 5029-37 (2007).
82. Fumagalli, S. *et al.* Absence of nucleolar disruption after impairment of 40S ribosome biogenesis reveals an rpL11-translation-dependent mechanism of p53 induction. *Nat.Cell Biol.* **11**, 501-508 (2009).
83. Holzel, M. *et al.* Defects in 18 S or 28 S rRNA processing activate the p53 pathway. *J Biol Chem* **285**, 6364-70 (2010).
84. Bernstein, K.A., Bleichert, F., Bean, J.M., Cross, F.R. & Baserga, S.J. Ribosome biogenesis is sensed at the Start cell cycle checkpoint. *Mol Biol Cell* **18**, 953-64 (2007).
85. Azuma, M., Toyama, R., Laver, E. & Dawid, I.B. Perturbation of rRNA synthesis in the bap28 mutation leads to apoptosis mediated by p53 in the zebrafish central nervous system. *J.Biol.Chem.* **281**, 13309-13316 (2006).
86. Simmons, T. & Appel, B. Mutation of pescadillo disrupts oligodendrocyte formation in zebrafish. *PLoS One* **7**, e32317 (2012).
87. Watkins-Chow, D.E. *et al.* Mutation of the diamond-blackfan anemia gene Rps7 in mouse results in morphological and neuroanatomical phenotypes. *PLoS Genet* **9**, e1003094 (2013).

88. Higa-Nakamine, S. *et al.* Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Res* **40**, 391-8 (2012).
89. Freed, E.F., Bleichert, F., Dutca, L.M. & Baserga, S.J. When ribosomes go bad: diseases of ribosome biogenesis. *Mol.Biosyst.* **6**, 481-493 (2010).
90. Ganapathi, K.A. & Shimamura, A. Ribosomal dysfunction and inherited marrow failure. *Br J Haematol* **141**, 376-87 (2008).
91. Khan, S. *et al.* Do ribosomopathies explain some cases of common variable immunodeficiency? *Clin Exp Immunol* **163**, 96-103 (2011).
92. Dixon, M.J. Treacher Collins syndrome. *Hum Mol Genet* **5 Spec No**, 1391-6 (1996).
93. Dauwerse, J.G. *et al.* Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. *Nat.Genet.* **43**, 20-22 (2011).
94. Dixon, J. *et al.* Tcof1/Treacle is required for neural crest cell formation and proliferation deficiencies that cause craniofacial abnormalities. *Proc Natl Acad Sci U S A* **103**, 13403-8 (2006).
95. McGowan, K.A. *et al.* Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat.Genet.* **40**, 963-970 (2008).
96. Mauro, V.P. & Edelman, G.M. The ribosome filter hypothesis. *Proc Natl Acad Sci U S A* **99**, 12031-6 (2002).
97. Kondrashov, N. *et al.* Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383-397 (2011).
98. Yoon, A. *et al.* Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* **312**, 902-6 (2006).
99. Kongsuwan, K. *et al.* A Drosophila Minute gene encodes a ribosomal protein. *Nature* **317**, 555-8 (1985).
100. Marygold, S.J. *et al.* The ribosomal protein genes and Minute loci of Drosophila melanogaster. *Genome Biol* **8**, R216 (2007).
101. Matsson, H. *et al.* Targeted disruption of the ribosomal protein S19 gene is lethal prior to implantation. *Mol Cell Biol* **24**, 4032-4037 (2004).
102. Zhang, S., Shi, M., Hui, C.C. & Rommens, J.M. Loss of the mouse ortholog of the shwachman-diamond syndrome gene (Sbds) results in early embryonic lethality. *Mol.Cell Biol* **26**, 6656-6663 (2006).
103. He, J. *et al.* Targeted disruption of Dkc1, the gene mutated in X-linked dyskeratosis congenita, causes embryonic lethality in mice. *Oncogene* **21**, 7740-7744 (2002).
104. Dixon, J., Brakebusch, C., Fassler, R. & Dixon, M.J. Increased levels of apoptosis in the prefusion neural folds underlie the craniofacial disorder, Treacher Collins syndrome. *Hum Mol Genet* **9**, 1473-80 (2000).
105. Jones, N.C. *et al.* Prevention of the neurocristopathy Treacher Collins syndrome through inhibition of p53 function. *Nat.Med.* **14**, 125-133 (2008).
106. Oliver, E.R., Saunders, T.L., Tarle, S.A. & Glaser, T. Ribosomal protein L24 defect in belly spot and tail (Bst), a mouse Minute. *Development* **131**, 3907-20 (2004).

107. Willig, T.N. *et al.* Mutations in ribosomal protein S19 gene and diamond blackfan anemia: wide variations in phenotypic expression. *Blood* **94**, 4294-306 (1999).
108. Venkatasubramani, N. & Mayer, A.N. A zebrafish model for the Shwachman-Diamond syndrome (SDS). *Pediatr Res* **63**, 348-52 (2008).
109. Provost, E. *et al.* Ribosomal biogenesis genes play an essential and p53-independent role in zebrafish pancreas development. *Development* **139**, 3232-3241 (2012).
110. Danilova, N., Sakamoto, K.M. & Lin, S. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. *Blood* **112**, 5228-5237 (2008).
111. Torihara, H. *et al.* Erythropoiesis failure due to RPS19 deficiency is independent of an activated Tp53 response in a zebrafish model of Diamond-Blackfan anaemia. *Br J Haematol* **152**, 648-54 (2011).
112. Hakuno, F., Hughes, D.A. & Yamamoto, M. The Schizosaccharomyces pombe mra1 gene, which is required for cell growth and mating, can suppress the mating inefficiency caused by a deficit in the Ras1 activity. *Genes Cells* **1**, 303-315 (1996).
113. Liu, P.C. & Thiele, D.J. Novel stress-responsive genes EMG1 and NOP14 encode conserved, interacting proteins required for 40S ribosome biogenesis. *Mol.Biol Cell* **12**, 3644-3657 (2001).
114. Eschrich, D., Buchhaupt, M., Kotter, P. & Entian, K.D. Nep1p (Emg1p), a novel protein conserved in eukaryotes and archaea, is involved in ribosome biogenesis. *Curr.Genet.* **40**, 326-338 (2002).
115. Buchhaupt, M., Meyer, B., Kotter, P. & Entian, K.D. Genetic evidence for 18S rRNA binding and an Rps19p assembly function of yeast nucleolar protein Nep1p. *Mol.Genet.Genomics* **276**, 273-284 (2006).
116. Buchhaupt, M., Kotter, P. & Entian, K.D. Mutations in the nucleolar proteins Tma23 and Nop6 suppress the malfunction of the Nep1 protein. *FEMS Yeast Res.* **7**, 771-781 (2007).
117. Leulliot, N., Bohnsack, M.T., Graille, M., Tollervey, D. & van, T.H. The yeast ribosome synthesis factor Emg1 is a novel member of the superfamily of alpha/beta knot fold methyltransferases. *Nucleic Acids Res.* **36**, 629-639 (2008).
118. Tkaczuk, K.L., Dunin-Horkawicz, S., Purta, E. & Bujnicki, J.M. Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. *BMC.Bioinformatics.* **8**, 73 (2007).
119. Taylor, A.B. *et al.* The crystal structure of Nep1 reveals an extended SPOUT-class methyltransferase fold and a pre-organized SAM-binding site. *Nucleic Acids Res.* **36**, 1542-1554 (2008).
120. Wurm, J.P. *et al.* The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. *Nucleic Acids Res.* **38**, 2387-2389 (2010).
121. Thomas, S.R., Keller, C.A., Szyk, A., Cannon, J.R. & Laronde-Leblanc, N.A. Structural insight into the functional mechanism of Nep1/Emg1 N1-specific

pseudouridine methyltransferase in ribosome biogenesis. *Nucleic Acids Res.* **39**, 2445-2457 (2011).
122. Wu, X., Sandhu, S., Patel, N., Triggs-Raine, B. & Ding, H. EMG1 is essential for mouse pre-implantation embryo development. *BMC.Dev.Biol.* **10**, 99 (2010).
123. Barlow, J.L. *et al.* A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q- syndrome. *Nat.Med.* **16**, 59-66 (2010).
124. Barkic, M. *et al.* The p53 tumor suppressor causes congenital malformations in Rpl24-deficient mice and promotes their survival. *Mol.Cell Biol.* **29**, 2489-2504 (2009).
125. Lamont, R.E. University of Manitoba (2003).
126. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J.Mol.Biol.* **215**, 403-410 (1990).
127. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl.Acids.Res.* **22**, 4673-4680 (1994).
128. Nagy, A., M, G., K, V. & R, B. *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2003).
129. Southern, E. Southern blotting. *Nat Protoc* **1**, 518-25 (2006).
130. Towbin, H., Staechelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc.Natl.Acad.Sci.USA* **76**, 4350-4354 (1979).
131. Bonifacino, J.S., Dell'Angelica, E.C. & Springer, T.A. Immunoprecipitation. *Curr Protoc Protein Sci* **Chapter 9**, Unit 9 8 (2001).
132. Bonifacino, J.S. Metabolic labeling with amino acids. *Curr Protoc Protein Sci* **Chapter 3**, Unit 3 7 (2001).
133. Prophet, E.B. Tissue processing, dehydration, clearing, and infiltration. in *Laboratory methods in histotechnology* (eds. Prophet, E.B., Mills, B., Arrington, J.B. & Sobin, L.H.) 29-31 (American Registry of Pathology, Washington, D.C, 1994).
134. Pestov, D.G., Lapik, Y.R. & Lau, L.F. Assays for ribosomal RNA processing and ribosome assembly. *Curr Protoc Cell Biol* **Chapter 22**, Unit 22 11 (2008).
135. Farley, F.W., Soriano, P., Steffen, L.S. & Dymecki, S.M. Widespread recombinase expression using FLPeR (flipper) mice. *Genesis* **28**, 106-10 (2000).
136. Lander, E.S. & Botstein, D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* **236**, 1567-70 (1987).
137. Sheffield, V.C., Nishimura, D.Y. & Stone, E.M. Novel approaches to linkage mapping. *Curr.Opin.Genet.Dev.* **5**, 335-341 (1995).
138. Nakamuta, M. *et al.* Complete phenotypic characterization of apobec-1 knockout mice with a wild-type genetic background and a human apolipoprotein B transgenic background, and restoration of apolipoprotein B mRNA editing by somatic gene transfer of Apobec-1. *J Biol Chem* **271**, 25981-8 (1996).

139. Mallam, A.L. & Jackson, S.E. The dimerization of an alpha/beta-knotted protein is essential for structure and function. *Structure* **15**, 111-22 (2007).
140. Narla, A., Hurst, S.N. & Ebert, B.L. Ribosome defects in disorders of erythropoiesis. *Int J Hematol* **93**, 144-9 (2011).
141. Liu, J.M. & Ellis, S.R. Ribosomes and marrow failure: coincidental association or molecular paradigm? *Blood* **107**, 4583-8 (2006).
142. Parenti, G. Treating lysosomal storage diseases with pharmacological chaperones: from concept to clinics. *EMBO Mol Med* **1**, 268-79 (2009).
143. Sakahira, H., Breuer, P., Hayer-Hartl, M.K. & Hartl, F.U. Molecular chaperones as modulators of polyglutamine protein aggregation and toxicity. *Proc Natl Acad Sci U S A* **99 Suppl 4**, 16412-8 (2002).
144. Edwards, J.H., Harnden, D.G., Cameron, A.H., Crosse, V.M. & Wolff, O.H. A new trisomic syndrome. *Lancet* **1**, 787-90 (1960).
145. Innes, A.M. & Lowry, R.B. Bowen-Conradi syndrome in non Hutterite infant. *Clin.Dysmorphol.* **11**, 147-148 (2002).
146. Lemire, E.G. Bowen-Conradi syndrome. *Clin.Dysmorphol.* **11**, 149 (2002).
147. Armistead, J. *et al.* Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome. *Am.J.Hum.Genet.* **84**, 728-739 (2009).
148. Kopito, R.R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol* **10**, 524-30 (2000).
149. Tyedmers, J., Mogk, A. & Bukau, B. Cellular strategies for controlling protein aggregation. *Nat Rev Mol Cell Biol* **11**, 777-88 (2010).
150. Jadhav, T. & Wooten, M.W. Defining an Embedded Code for Protein Ubiquitination. *J Proteomics Bioinform* **2**, 316 (2009).
151. Tai, H.C. & Schuman, E.M. Ubiquitin, the proteasome and protein degradation in neuronal function and dysfunction. *Nat Rev Neurosci* **9**, 826-38 (2008).
152. Lee, D.H. & Goldberg, A.L. Selective inhibitors of the proteasome-dependent and vacuolar pathways of protein degradation in Saccharomyces cerevisiae. *J Biol Chem* **271**, 27280-4 (1996).
153. Lam, Y.W., Lamond, A.I., Mann, M. & Andersen, J.S. Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Curr Biol* **17**, 749-60 (2007).
154. Dino Rockel, T. & von Mikecz, A. Proteasome-dependent processing of nuclear proteins is correlated with their subnuclear localization. *J Struct Biol* **140**, 189-99 (2002).
155. Stavreva, D.A. *et al.* Potential roles for ubiquitin and the proteasome during ribosome biogenesis. *Mol Cell Biol* **26**, 5131-45 (2006).
156. Cmarko, D. *et al.* Ultrastructural analysis of nucleolar transcription in cells microinjected with 5-bromo-UTP. *Histochem Cell Biol* **113**, 181-7 (2000).
157. Hozak, P., Cook, P.R., Schofer, C., Mosgoller, W. & Wachtler, F. Site of transcription of ribosomal RNA and intranucleolar structure in HeLa cells. *J Cell Sci* **107 (Pt 2)**, 639-48 (1994).
158. Marsh, K.L., Dixon, J. & Dixon, M.J. Mutations in the Treacher Collins syndrome gene lead to mislocalization of the nucleolar protein treacle. *Hum Mol Genet* **7**, 1795-800 (1998).

159. Angelini, M. *et al.* Missense mutations associated with Diamond-Blackfan anemia affect the assembly of ribosomal protein S19 into the ribosome. *Hum Mol Genet* **16**, 1720-7 (2007).
160. Da Costa, L. *et al.* Nucleolar localization of RPS19 protein in normal cells and mislocalization due to mutations in the nucleolar localization signals in 2 Diamond-Blackfan anemia patients: potential insights into pathophysiology. *Blood* **101**, 5039-45 (2003).
161. Grandi, P. *et al.* 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol Cell* **10**, 105-15 (2002).
162. Krogan, N.J. *et al.* High-definition macromolecular composition of yeast RNA-processing complexes. *Mol Cell* **13**, 225-39 (2004).
163. Wilson, L., Gale, E. & Maden, M. The role of retinoic acid in the morphogenesis of the neural tube. *J Anat* **203**, 357-68 (2003).
164. Ching, Y.H. *et al.* High resolution mapping and positional cloning of ENU-induced mutations in the Rw region of mouse chromosome 5. *BMC Genet* **11**, 106 (2010).
165. Milkereit, P. *et al.* A Noc complex specifically involved in the formation and nuclear export of ribosomal 40 S subunits. *J Biol Chem* **278**, 4072-4081 (2003).
166. Rout, M.P. *et al.* The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol* **148**, 635-651 (2000).
167. Gall, J.G., Bellini, M., Wu, Z. & Murphy, C. Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol Biol Cell* **10**, 4385-402 (1999).
168. Cretien, A. *et al.* Study of the effects of proteasome inhibitors on ribosomal protein S19 (RPS19) mutants, identified in patients with Diamond-Blackfan anemia. *Haematologica* **93**, 1627-1634 (2008).
169. Austin, K.M., Leary, R.J. & Shimamura, A. The Shwachman-Diamond SBDS protein localizes to the nucleolus. *Blood* **106**, 1253-8 (2005).
170. Misteli, T., Caceres, J.F. & Spector, D.L. The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**, 523-7 (1997).
171. Spector, D.L. Nuclear organization of pre-mRNA processing. *Curr Opin Cell Biol* **5**, 442-7 (1993).
172. von Mikecz, A. The nuclear ubiquitin-proteasome system. *J Cell Sci* **119**, 1977-84 (2006).
173. Wojcik, C. & DeMartino, G.N. Intracellular localization of proteasomes. *Int J Biochem Cell Biol* **35**, 579-89 (2003).
174. Schubert, U. *et al.* Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **404**, 770-4 (2000).
175. Ciechanover, A. The ubiquitin proteolytic system: from a vague idea, through basic mechanisms, and onto human diseases and drug targeting. *Neurology* **66**, S7-19 (2006).
176. Chan, Y.L., Suzuki, K. & Wool, I.G. The carboxyl extensions of two rat ubiquitin fusion proteins are ribosomal proteins S27a and L40. *Biochem Biophys Res Commun* **215**, 682-90 (1995).

177. Redman, K.L. & Rechsteiner, M. Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature* **338**, 438-40 (1989).
178. Finley, D., Bartel, B. & Varshavsky, A. The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* **338**, 394-401 (1989).
179. Warner, J.R. The economics of ribosome biosynthesis in yeast. *Trends Biochem.Sci.* **24**, 437-440 (1999).
180. Wery, M., Ruidant, S., Schillewaert, S., Lepore, N. & Lafontaine, D.L. The nuclear poly(A) polymerase and Exosome cofactor Trf5 is recruited cotranscriptionally to nucleolar surveillance. *RNA* **15**, 406-19 (2009).
181. Alawi, F., Lin, P., Ziober, B. & Patel, R. Correlation of dyskerin expression with active proliferation independent of telomerase. *Head Neck* **33**, 1041-51 (2011).
182. Gu, B.W., Ge, J., Fan, J.M., Bessler, M. & Mason, P.J. Slow growth and unstable ribosomal RNA lacking pseudouridine in mouse embryonic fibroblast cells expressing catalytically inactive dyskerin. *FEBS Lett* **587**, 2112-7 (2013).
183. Chen, F.W. & Ioannou, Y.A. Ribosomal proteins in cell proliferation and apoptosis. *Int Rev Immunol* **18**, 429-48 (1999).
184. Gill, T., Cai, T., Aulds, J., Wierzbicki, S. & Schmitt, M.E. RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation. *Mol Cell Biol* **24**, 945-53 (2004).
185. Jack, K. *et al.* rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol Cell* **44**, 660-6 (2011).
186. Ohno, S., Shibayama, M., Sato, M., Tokunaga, A. & Yoshida, N. Polypyrimidine tract-binding protein regulates the cell cycle through IRES-dependent translation of CDK11(p58) in mouse embryonic stem cells. *Cell Cycle* **10**, 3706-13 (2011).
187. Sachs, A.B. Cell cycle-dependent translation initiation: IRES elements prevail. *Cell* **101**, 243-5 (2000).
188. Horos, R. *et al.* Ribosomal deficiencies in Diamond-Blackfan anemia impair translation of transcripts essential for differentiation of murine and human erythroblasts. *Blood* **119**, 262-72 (2012).
189. Bidou, L. *et al.* Nonsense-mediated decay mutants do not affect programmed -1 frameshifting. *RNA* **6**, 952-61 (2000).
190. Leung, R., Cuddy, K., Wang, Y., Rommens, J. & Glogauer, M. Sbds is required for Rac2-mediated monocyte migration and signaling downstream of RANK during osteoclastogenesis. *Blood* **117**, 2044-53 (2011).
191. Devlin, E.E., Dacosta, L., Mohandas, N., Elliott, G. & Bodine, D.M. A transgenic mouse model demonstrates a dominant negative effect of a point mutation in the RPS19 gene associated with Diamond-Blackfan anemia. *Blood* **116**, 2826-35 (2010).
192. Gu, B.W., Bessler, M. & Mason, P.J. A pathogenic dyskerin mutation impairs proliferation and activates a DNA damage response independent of telomere length in mice. *Proc Natl Acad Sci U S A* **105**, 10173-8 (2008).

193. Nagy, A. Cre recombinase: the universal reagent for genome tailoring. *Genesis* **26**, 99-109 (2000).
194. Friedrich, G. & Soriano, P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5**, 1513-23 (1991).
195. Schaft, J., Ashery-Padan, R., van der Hoeven, F., Gruss, P. & Stewart, A.F. Efficient FLP recombination in mouse ES cells and oocytes. *Genesis.* **31**, 6-10 (2001).
196. Jensen-Seaman, M.I. *et al.* Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res* **14**, 528-38 (2004).
197. Copp, A.J., Greene, N.D. & Murdoch, J.N. The genetic basis of mammalian neurulation. *Nat Rev Genet* **4**, 784-93 (2003).
198. Meier, P., Finch, A. & Evan, G. Apoptosis in development. *Nature* **407**, 796-801 (2000).
199. Penaloza, C., Lin, L., Lockshin, R.A. & Zakeri, Z. Cell death in development: shaping the embryo. *Histochem Cell Biol* **126**, 149-58 (2006).
200. *Fundamental Neuroscience*, 1256 (Elesevier/ Academic Press, Amsterdam, 2008).
201. Curtis, M.A., Low, V.F. & Faull, R.L. Neurogenesis and progenitor cells in the adult human brain: a comparison between hippocampal and subventricular progenitor proliferation. *Dev Neurobiol* **72**, 990-1005 (2012).
202. Eriksson, P.S. *et al.* Neurogenesis in the adult human hippocampus. *Nat Med* **4**, 1313-7 (1998).
203. Kuhn, H.G., Dickinson-Anson, H. & Gage, F.H. Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *J Neurosci* **16**, 2027-33 (1996).
204. Spalding, K.L. *et al.* Dynamics of hippocampal neurogenesis in adult humans. *Cell* **153**, 1219-27 (2013).
205. O'Farrell, P.H., Stumpff, J. & Su, T.T. Embryonic cleavage cycles: how is a mouse like a fly? *Curr Biol* **14**, R35-45 (2004).
206. Copp, A.J. Death before birth: clues from gene knockouts and mutations. *Trends Genet* **11**, 87-93 (1995).
207. Copp, A.J. Neurulation in the cranial region--normal and abnormal. *J Anat* **207**, 623-35 (2005).
208. Greene, N.D. & Copp, A.J. Development of the vertebrate central nervous system: formation of the neural tube. *Prenat Diagn* **29**, 303-11 (2009).
209. Harris, M.J. & Juriloff, D.M. Mini-review: toward understanding mechanisms of genetic neural tube defects in mice. *Teratology* **60**, 292-305 (1999).
210. Fleming, A. & Copp, A.J. A genetic risk factor for mouse neural tube defects: defining the embryonic basis. *Hum Mol Genet* **9**, 575-81 (2000).
211. Ybot-Gonzalez, P. & Copp, A.J. Bending of the neural plate during mouse spinal neurulation is independent of actin microfilaments. *Dev Dyn* **215**, 273-83 (1999).
212. Wilson, L. & Maden, M. The mechanisms of dorsoventral patterning in the vertebrate neural tube. *Dev Biol* **282**, 1-13 (2005).

213. Zhang, Y., Morimoto, K., Danilova, N., Zhang, B. & Lin, S. Zebrafish models for dyskeratosis congenita reveal critical roles of p53 activation contributing to hematopoietic defects through RNA processing. *PLoS One* **7**, e30188 (2012).
214. van Gijn, D.R., Tucker, A.S. & Cobourne, M.T. Craniofacial development: current concepts in the molecular basis of Treacher Collins syndrome. *Br J Oral Maxillofac Surg* **51**, 384-8 (2013).
215. Le Merrer, M. & Maroteaux, P. Cartilage hair hypoplasia in infancy: a misleading chondrodysplasia. *Eur J Pediatr* **150**, 847-51 (1991).
216. Dixon, J. & Dixon, M.J. Genetic background has a major effect on the penetrance and severity of craniofacial defects in mice heterozygous for the gene encoding the nucleolar protein Treacle. *Dev Dyn* **229**, 907-14 (2004).
217. Gonzales, B. *et al.* The Treacher Collins syndrome (TCOF1) gene product is involved in pre-rRNA methylation. *Hum.Mol.Genet.* **14**, 2035-2043 (2005).
218. Di, R.A., Rogers, J.D., Brown, R., Werner, P. & Bottiglieri, T. S-Adenosyl-Methionine improves depression in patients with Parkinson's disease in an open-label clinical trial. *Mov Disord.* **15**, 1225-1229 (2000).
219. Thompson, M.A. *et al.* Dietary supplement S-adenosyl-L-methionine (AdoMet) effects on plasma homocysteine levels in healthy human subjects: a double-blind, placebo-controlled, randomized clinical trial. *J Altern Complement Med* **15**, 523-9 (2009).
220. Burns, K.A., Murphy, B., Danzer, S.C. & Kuan, C.Y. Developmental and post-injury cortical gliogenesis: a genetic fate-mapping study with Nestin-CreER mice. *Glia* **57**, 1115-29 (2009).
221. Dehay, C. & Kennedy, H. Cell-cycle control and cortical development. *Nat Rev Neurosci* **8**, 438-50 (2007).
222. Agarwal, M.L., Agarwal, A., Taylor, W.R. & Stark, G.R. p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts. *Proc Natl Acad Sci U S A* **92**, 8493-7 (1995).
223. Taylor, W.R. *et al.* Mechanisms of G2 arrest in response to overexpression of p53. *Mol Biol Cell* **10**, 3607-22 (1999).
224. Taylor, W.R. & Stark, G.R. Regulation of the G2/M transition by p53. *Oncogene* **20**, 1803-15 (2001).
225. Donati, G., Montanaro, L. & Derenzini, M. Ribosome biogenesis and control of cell proliferation: p53 is not alone. *Cancer Res* **72**, 1602-7 (2012).
226. Lesnyak, D.V. *et al.* Methyltransferase that modifies guanine 966 of the 16 S rRNA: functional identification and tertiary structure. *J Biol Chem* **282**, 5880-7 (2007).
227. Kemp, J.P. *et al.* Nuclear factors involved in mitochondrial translation cause a subgroup of combined respiratory chain deficiency. *Brain* **134**, 183-95 (2011).
228. Miller, C. *et al.* Defective mitochondrial translation caused by a ribosomal protein (MRPS16) mutation. *Ann Neurol* **56**, 734-8 (2004).
229. O'Brien, T.W., O'Brien, B.J. & Norman, R.A. Nuclear MRP genes and mitochondrial disease. *Gene* **354**, 147-51 (2005).

230. Shutt, T.E. & Shadel, G.S. A compendium of human mitochondrial gene expression machinery with links to disease. *Environ Mol Mutagen* **51**, 360-79 (2010).
231. Guan, M.X. Mitochondrial 12S rRNA mutations associated with aminoglycoside ototoxicity. *Mitochondrion* **11**, 237-45 (2011).
232. Sylvester, J.E., Fischel-Ghodsian, N., Mougey, E.B. & O'Brien, T.W. Mitochondrial ribosomal proteins: candidate genes for mitochondrial disease. *Genet Med* **6**, 73-80 (2004).
233. Pavitt, G.D. & Proud, C.G. Protein synthesis and its control in neuronal cells with a focus on vanishing white matter disease. *Biochem Soc Trans* **37**, 1298-310 (2009).
234. Schiffmann, R. & Elroy-Stein, O. Childhood ataxia with CNS hypomyelination/vanishing white matter disease--a common leukodystrophy caused by abnormal control of protein synthesis. *Mol Genet Metab* **88**, 7-15 (2006).
235. Haud, N. *et al.* rnaset2 mutant zebrafish model familial cystic leukoencephalopathy and reveal a role for RNase T2 in degrading ribosomal RNA. *Proc.Natl.Acad.Sci.U.S.A* **108**, 1099-1103 (2011).
236. Henneke, M. *et al.* RNASET2-deficient cystic leukoencephalopathy resembles congenital cytomegalovirus brain infection. *Nat.Genet.* **41**, 773-775 (2009).
237. Ding, Q., Markesbery, W.R., Chen, Q., Li, F. & Keller, J.N. Ribosome dysfunction is an early event in Alzheimer's disease. *J Neurosci* **25**, 9171-5 (2005).
238. Ruggero, D. & Pandolfi, P.P. Does the ribosome translate cancer? *Nat Rev Cancer* **3**, 179-92 (2003).
239. Chakraborty, S., Nagashri, M.N., Mohiyuddin, S.M., Gopinath, K.S. & Kumar, A. Gene expression profiling of oral squamous cell carcinoma by differential display rt-PCR and identification of tumor biomarkers. *Indian J Surg Oncol* **1**, 284-93 (2010).
240. Zhou, B., Wu, Q., Chen, G., Zhang, T.P. & Zhao, Y.P. NOP14 promotes proliferation and metastasis of pancreatic cancer cells. *Cancer Lett* **322**, 195-203 (2012).
241. Amsterdam, A. *et al.* Many ribosomal protein genes are cancer genes in zebrafish. *PLoS Biol* **2**, E139 (2004).