Investigation of the function and protein-protein interaction of zebrafish progranulin-A during development

by

David Charles Baranowski

A thesis submitted to the Faculty of Graduate Studies and Research, McGill University, in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Department of Medicine, Division of Experimental Medicine

McGill University

Montreal, Quebec, Canada

What is memory worth....without noble purpose?

-Frank Herbert, Heretics of Dune.

This work is dedicated to my wife Samantha, and my two daughters. Although research is perhaps the most fulfilling job one could hope for, you are the only reason I have accomplished so much and continue to be the true source of my pride, joy and the mirror that ultimately reflects my soul.

To my mother Jennifer for having a patience that I can only hope to achieve, as well as for the odd babysitting duties.

To my sister Ann-marie and my brother Mark for their support and enthusiasm.

To my father Charles for making me stronger.

To my mother and father in-law Doug and Audrey Robertson for always being there whenever we needed help.

To the friends that I have made and lost along the way, maybe now that this is over we can finally go for that beer. Peter Zalai, Matthew Holly, Nancy and Murray Wasley, Peter Rumsey and Katie Alexander.

To Blizzard Corp. for providing the well deserved distraction from the daily grind and for which this thesis suffered continuous setbacks.

ABSTRACT

Progranulin is a secreted glycoprotein consisting of seven and one half tandem repeats of the cysteine-rich granulin motif. Progranulin has modulatory effects on cell proliferation and migration and has been implicated in several biological contexts including epithelial homeostasis and wound healing, immunity, sexual differentiation of the brain and tumorigenesis. Recently, mutations within the human *progranulin* gene have been associated with a form of autosomal dominant frontotemporal dementia. Despite the fact that many reports have been published implicating progranulin in numerous physiological and pathophysiological events particularly within the last decade, no studies have appeared describing the phenotypic consequences of ablation of the single progranulin gene in the mouse knockout model. Furthermore, the cell surface receptors and/or binding partners that mediate progranulin signal transduction have not been fully defined. Work presented in this thesis surrounds the investigation of the physiological functions and putative binding partners of the zebrafish (*Danio rerio*) *progranulin-a* gene product. Using antisense technology we have assessed the phenotypic consequences of *progranulin-a* ablation during zebrafish development, which include anaemia and leukopenia, aberrant neural crest and pharyngeal endoderm patterning that cause craniofacial dysmorphogenesis as well as the failure of the primitive visceral gut and accessory organs to undergo continued growth and differentiation. In conjunction with these developmental studies a new approach to identifying candidate binding partners employing liquid chromatography, surface plasmon resonance and mass spectrometry has been developed. This system of protein-protein interaction screening was employed to indentify putative binding partners for zebrafish progranulin-A. In this manner a single

progranulin binding protein candidate was identified. Based on its primary structure the progranulin-A binding protein is as a member of the class VII C-type lectin domain protein family, which includes the structurally similar fish antifreeze proteins and the regenerating islet-derived protein family in mammals.

RESUME

La progranuline est une glycoprotéine sécrétée se composant de sept répétitions et demi en tandem du motif de granuline riche en cystéine. La progranuline est capable de moduler la prolifération et la migration de cellules et est aussi impliquée dans différents mécanismes biologiques tels que l'homéostasie épithéliale, la cicatrisation, le système immunitaire ainsi que la différentiation sexuelle du cerveau et la formation de tumeurs. Récemment, des mutations du gène de la progranuline chez l'homme ont été associées à une forme de démence autosomale fronto-temporale dominante. Malgré le nombre élevé de publications rapportant l'implication de la progranuline dans de nombreux événements physiologiques et pathophysiologiques, en particulier cette dernière décennie, il n'existe aucune étude décrivant les conséquences phénotypiques de l'ablation de l'unique gène de progranuline dans un modèle de souris transgénique. De plus, les récepteurs situés à la surface de la cellule et/ou les ligands prenant part à la transduction du signal de la progranuline n'ont pas été totalement définis. Le travail présenté dans cette thèse a été développé dans le cadre de la recherche sur les fonctions physiologiques et les ligands supposés du produit du gène de progranuline-a du poisson zèbre (*rerio de Danio*). A l'aide de la technologie d'antisense, nous avons évalué les conséquences phénotypiques de l'ablation de progranuline-a durant le développement du poisson zèbre qui sont: l'anémie, la leucopénie, des anomalies de la crête neurale, l'endoderme pharyngal associé à la dysmorphie crano-faciale ainsi que l'incapacité de l'intestin viscéral et les organes accessoires à engager un processus de croissance de différentiation continu. En parallèle à ces études développementales, une nouvelle approche a été développée pour identifier des ligands, grâce à l'usage de la chromatographie à phase liquide, la résonance de plasmon

de surface et la spectrométrie de masse. Ce système de criblage d'interaction de protéine-protéine a été utilisé afin d'identifier les ligands supposés de la progranuline-a chez le poisson zèbre. Cette méthode a permis l'identification d'une seule protéine potentiellement liée à la progranuline. En se basant sur sa structure primaire, il apparait que cette protéine est un membre de type C de la classe VII de la famille des protéines du domaine lectine, qui comprend les protéines structurellement similaires du poisson antigel et la famille de protéines régénératrices dérivées de l'îlot chez les mammifères.

ACKNOWLEDGEMENTS

This project is a continuation of work that was begun before my presence and will undoubtedly continue in my absence. I would like to thank everyone involved:

Dr. Hugh Bennett, my mentor for being upfront and courageous in your commitment to a project that appeared to forever stand at the cusp of completion and all out implosion. Your friendship, patience and understanding will always be a part of my character.

Benoit Cadieux for teaching everything there is to know about the zebrafish biology and more importantly how to learn. This thesis is a reflection of your own determination and spirit as much as it is mine.

Babykumari Chitramuthu for their help with almost every aspect of this project including, zebrafish husbandry, in situ hybridizations, morpholino validation and conceptual organization. More importantly, for constant moral support and friendship even at the darkest times.

Susan James for introducing a "green" bachelor's student to the right end of the pipettor, protein purification and cell culture.

Dr. Marcos DiFalco for giving good "what not to do" advice and teaching me the mantra of mass spectrometry.

Dr. Pierre Drapeau and students from their equipment and advice for zebrafish microinjection and microscopy.

Dr. Fernando Congote and Dr. Andrew Bateman for always lending an ear when it appeared an impasse had been reached.

Dr. Sylvie Barchéchath and Dr. Bertrand Jean-Claude for help in translations.

Cara Ryan for having the patience and determination to follow in my footsteps and who will build upon and undoubtedly surpass my own efforts in the lab.

Dr. Denis Kay for not only offering advice and assistance but providing an opportunity that comes but once a lifetime. See you soon.

Table of Contents

Section 2 **Materials and methods**

Section 3 Results

Section 4. **Discussion**

ABBREVIATIONS

COMP	Cartilage oligomeric matrix protein
CNS	Central nervous system
CTLD	C-type lectin domain
DLK1	Delta-like homologue 1
EGF	Epidermal growth factor
EMA	Ectomesenchymal arches
FTD	Frontotemporal dementia
FTD-U	Frontotemporal dementia with ubiquitin positive neuronal inclusions
GEP	Granulin-epithelin precursor
HSC	Hematopoietic stem cell
hpf	Hours post-fertilization
ICM	Intermediate cell mass
IGF-IR	Insulin-like growth factor-1 receptor
IFABP	Intestinal fatty acid binding protein
IMAC	Immobilized metal affinity chromatography
LC	Liquid chromatography
MAPK	Mitogen activated protein kinase
MEF	Mouse embryonic fibroblasts
MO	Morpholino oligonucleotide
MS	Mass spectrometry
NC	Neural crest
NMD	Nonsense-mediated decay
PBP	Progranulin binding protein
PCDGF	PC12 derived growth factor
PEPI	Proepithelin
PI_3K	Phosphoinositol-3 kinase
RA	Retinoic acid
REG	Regenerating-islet derived
RP-HPLC	Reversed-phase high performance liquid chromatography

SCL	Stem cell leukemia factor
SLPI	Secretory leukocyte protease inhibitor
SPR	Surface plasmon resonance
TAP	Tandem affinity purification
TFG-β	Transforming growth factor-beta
TNF-α	Tumor necrosis factor-alpha
TUNEL	Terminal deoxytransferase nick end labeling
Y2H	Yeast two-hybrid

MEASURES

bp	base-pair
cR	centirad
Ci	Curie
Da	dalton
dpf	days post-fertilization
g	gram
hpf	hours post-fertilization
kb	kilobase
kDa	kilodalton
M	molar
mCi	millicurie
mg	milligram
min	minute
mM	millimolar
mya	million years ago
ng	nanogram
nl	nanolitre
nM	nanomolar
sec	second
µg	microgram
µl	microlitre
^{o}C	degree celcius

LIST OF FIGURES

Section 1. General Introduction

1.1 Granulins

1.1.1 General Overview – The formation of cellular structures and tissues as well as their continued maintenance relies on a multitude of cell-cell communication pathways (Potter, 2007). A large proportion of these signalling events involve growth factors and their downstream effectors. Underlining the critical role played by growth factors in cellular homeostasis is their well characterized involvement in numerous disease states. Alterations in overall growth factor abundance and/or the potency of their signal transduction pathways can manifest in a plethora of pathophysiological consequences (Rivera et al., 2005; Storkebaum et al., 2004). In general growth factor over-stimulation is often associated with tumourigenesis, whereas insufficiency is frequently related to degenerative processes (Albini and Sporn, 2007; Rivera et al., 2005; Storkebaum et al., 2004). Consequently, the characterization of growth factor functions and elucidation of the mechanisms that regulate their biological activity are of particular interest.

Progranulin (also referred to as PC-cell derived growth factor, granulin-epithelin precursor, proepithelin and acrogranin; herein referred to as progranulin) is a secreted glycoprotein that has been associated with tumourigenic proliferation and/or invasion for a multitude of cancers, including prostate, breast, hepatic, ovarian, bladder, laryngeal and renal carcinomas as well as leiomyosarcoma, glioblastoma and multiple myeloma (Asaka et al., 2006; Cheung et al., 2004; Donald et al., 2001; Jones et al., 2003; Kong et al., 2007; Liau et al., 2000; Matsumura et al., 2006; Monami et al., 2006; Pan et al., 2004; Serrero and Ioffe, 2003; Wang et al., 2003). In all tumourigenic circumstances progranulin is presumed to affect a pathological response due to gene over expression.

Recently, a series of mutations in the *progranulin* gene have been determined to be the etiological factor underlying the development of autosomal dominant frontotemporal dementia (FTD) with ubiquitin-positive inclusions (FTD-U) (Baker et al., 2006; Behrens et al., 2007; Benussi et al., 2006; Boeve et al., 2006; Bronner et al., 2007; Cruts et al., 2006; Davion et al., 2007; Gass et al., 2006; Huey et al., 2006; Josephs et al., 2007; Le Ber et al., 2007; Leverenz et al., 2007; Masellis et al., 2006; Mesulam et al., 2007; Mukherjee et al., 2006; Schymick et al., 2007; Snowden et al., 2006; Spina et al., 2007; van der Zee et al., 2007; Whitwell et al., 2007). The *progranulin* mutations are found throughout the gene and despite the variation in mutation type (nonsense, frameshift, splice-site mutations, etc...) it is theorized that most alterations result in activation of the nonsense-mediated RNA decay (NMD) pathway, culminating in haploinsufficiency (Baker et al., 2006). At present, *progranulin* haploinsufficiency through the NMD pathway has been confirmed *in vivo* for seven mutations, all of which clearly introduce a premature termination codon (Baker et al., 2006; Cruts et al., 2006; Gass et al., 2006). Of particular interest is the discovery of familial FTD-U cases that are associated with mutations that lead to non-synonymous amino acid alterations (Bronner et al., 2007; Gass et al., 2006; Mukherjee et al., 2006; Schymick et al., 2007; van der Zee et al., 2007). The Ala9Asp *progranulin* mutation, the only non-synonymous alteration assessed thus far, is located within the signal peptide and appears to result in decreased expression of the affected allele through an unknown mechanism (Gass et al., 2006). In general familial FTD-U appears to be commonly associated with decreased *progranulin* expression, however the aberrant molecular mechanism that ultimately leads to the manifestation of this pathology is unknown.

The majority of investigative reports regarding progranulin demonstrate that this protein can modulate the growth and/or maintenance of numerous tissues when its expression becomes aberrant. There is little information, however, regarding the physiological functions of progranulin within most tissues. The primary exception to this generalization is the role of progranulin in the innate immune response and re-epithelialization mechanisms involved in wound healing (He et al., 2003; Zhu et al., 2002). Within the wound healing context, progranulin has been revealed to influence several biological functions, in part due to the functional dichotomy that exists between secreted progranulin and its constituent granulin peptides.

During wound healing, progranulin is expressed and secreted as the full-length precursor by endothelial cells and fibroblasts (He et al., 2003; Zanocco-Marani et al., 1999; Zhu et al., 2002). Progranulin promotes re-epithelialization, angiogenesis and acts as a chemotactic molecule for infiltrating fibroblasts and innate immune cells. Although progranulin recruits neutrophils to the wound, it prevents the inflammatory response by inhibiting TNF-α signalling (Zhu et al., 2002). The inhibition of progranulin precursor degradation by proteolysis and maintenance of its associated functions is, in part, due to the protein-protein interaction between progranulin and secretory protease leukocyte inhibitor (SLPI) (Zhu et al., 2002). Within neutrophils, progranulin is processed into individual granulin peptides, which are released during degranulation (Bateman et al., 1990). Neutrophil-derived granulins promote the innate immune response by inducing epithelial cells to secrete interleukin-8 that acts as a neutrophil chemokine (Zhu et al., 2002). In addition, granulins inhibit epithelial cell growth and these affects are further enhanced through the proteolytic degradation of extracellular progranulin to its granulin peptide constituents by neutrophil-derived elastase (Zanocco-Marani et al., 1999; Zhou et

al., 1993; Zhu et al., 2002). Thus, the interplay between the progranulin precursor and its constituent peptides influence the balance between wound healing and the innate immune response to injury.

Despite the significant role that progranulin has been suggested to play in tumourigenesis, FTD-U and wound healing, our understanding of its mechanism of action is limited. The full-length precursor has been shown to activate classical intracellular signalling pathways including those associated with the mitogen activated protein kinase (MAPK), phosphoinositol-3 kinase (PI_3K) and focal adhesion kinase (FAK); however, the majority of these investigations relate to progranulin cell signalling within various tumourigenic cell lines (He et al., 2002; Kamrava et al., 2005; Liu et al., 2007b; Lu and Serrero, 2001; Monami et al., 2006; Tangkeangsirisin et al., 2004; Tangkeangsirisin and Serrero, 2004; Wang et al., 2006). The most significant analysis with regard to progranulin signal transduction involves the use of R^- cells. These are mouse embryonic fibroblasts (MEF) derived from the insulin-like growth factor-I receptor (IGF-IR) knockout animal (Liu et al., 1993; Sell et al., 1994). Using R^- cells, Baserga and colleagues demonstrated that progranulin activated the Shc/MAPK and PI_3K/Akt pathways (Xu et al., 1998; Zanocco-Marani et al., 1999). Perhaps the most fundamental barrier to the additional dissection of the progranulin signal transduction mechanism(s) is the failure to identify a well defined cognate transmembrane receptor. Many attempts to study progranulin protein-protein interactions using a yeast two-hybrid (Y2H) system have been reported (Gonzalez et al., 2003; Shoham et al., 2003; Sui and Wilson, 2000; Thornburg et al., 2004; Trinh et al., 1999; Xu et al., 2007; Zhu et al., 2002). The majority of progranulin binding proteins indentified using the Y2H system are localized to the cytosol and/or nucleus and it remains to be determined if progranulin gains entry into

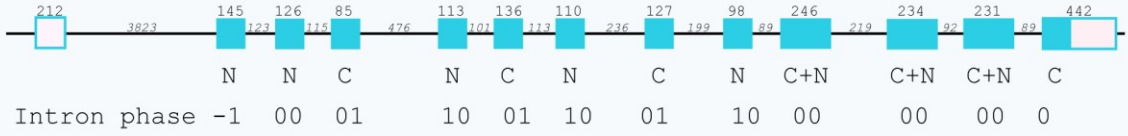
these cell compartments. Of the extracellular progranulin binding proteins indentified by using the Y2H, including SLPI, perlecan and cartilage oligomeric matrix protein (COMP), none have been shown to affect progranulin signal transduction (Gonzalez et al., 2003; Xu et al., 2007; Zhu et al., 2002).

The shortcomings in our understanding of the progranulin signal transduction mechanism is paralleled by the lack of knowledge regarding its physiological function in mouse models of development and disease. Despite the length of time since the characterization of the gene some fifteen years ago and mounting evidence supporting progranulins involvement in numerous pathologies, no report of the phenotype resulting from the knockout of the murine *progranulin* gene has yet been made. Research by Gerton and colleagues has provided evidence to suggest that progranulin is involved in early pre- and post-implantation embryogenesis. Recombinant progranulin administered to 8-cell stage murine embryos *in vitro*, resulted in increased trophectoderm cell number and accelerated the onset of cavitation (Diaz-Cueto et al., 2000). Furthermore, a progranulin function-blocking antibody inhibited development and decreased cell numbers. The same reciprocal affects of ligand versus function-blocking antibody has been demonstrated with regard to blastocyst adhesion and outgrowth (Qin et al., 2005a). These *in vitro* results suggest a role for progranulin during early embryonic development and uterine implantation and therefore, a *progranulin* null phenotype is expected to be embryonic lethal. Furthermore, the identification of progranulin mutations underlying autosomal dominant FTD and the associated multi-generational pedigree maps has not identified any individuals that are homozygous for disease-related mutations (Bronner et al., 2007; Bruni et al., 2007).

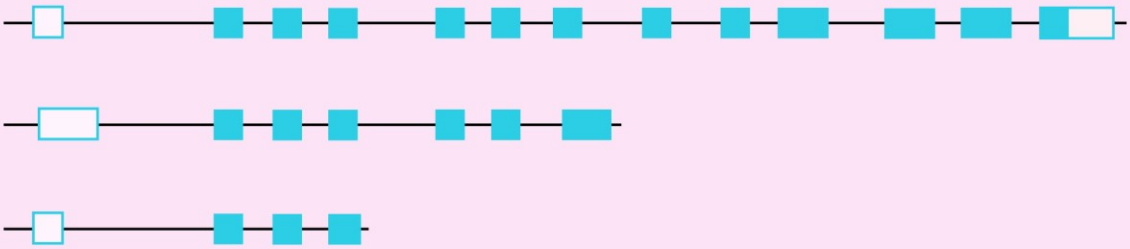
The focus of experimental investigation described in this thesis is to address these two deficiencies in our understanding of progranulin biology, specifically the identification of progranulin binding proteins involved in cell signalling and the functional relevance of this growth factor during embryonic development and organogenesis.

1.1.2 Original isolation and structural determinations – The granulins were originally discovered as a series of 6Kda cysteine-rich peptides derived from the granule fraction of human peripheral leukocytes (Bateman et al., 1990). Subsequent isolation of the coding mRNA as well as identification and structural determination of the human gene demonstrated that the previously isolated peptides originated from a single protein precursor, progranulin (Bhandari and Bateman, 1992; Bhandari et al., 1992). The mammalian progranulin growth factor is a secreted glycoprotein precursor composed of seven and one half tandem repeats of the granulin motif: $X_{2-3}CX_{5-6}CX_5CCX_8CCX_6CCX_5CCX_4CX_{5-6}CX_2$ (Hrabal et al., 1996). Based on the *progranulin* genomic architecture individual granulin motifs are encoded in two formats, either as two tandem exons where the splice site junctions bisect the domain or as a sequence of exons that harbour a C-terminal and N-terminal architecture (Figure 1) (Bhandari and Bateman, 1992). Notably, exons for each classification based on coding sequence (N-terminal, C-terminal or C- and N-terminal) also segregate into the same groups based on intronic phase, therefore theoretically allowing for numerous *cis*-splicing variations that ensure that the granulin motif will presumably remain intact (Figure 1). However, such exon-skipping has not been experimentally defined thus far. Based on available expressed sequence-tagged (EST) libraries there are at least twelve splice variants of *progranulin*,

Human Progranulin Genomic Architecture



Transcripts



Protein

Granulin Motif Consensus Sequence

$$.....X_{2\text{-}3}CX_{5\text{-}6}CX_{5\text{-}8}CCX_6CCX_5CCX_5CX_{5\text{-}6}CX_2.....$$

Figure 1. The human *progranulin* gene architecture and associated transcripts and proteins. *Upper panel*; The human *progranulin* gene is composed of thirteen exons (blue boxes) spanning ~8 kilobase of chromosome 17 (Bhandari and Bateman, 1992). Length of introns and exons are indicated and each exon is further labelled according to its open reading frame as encoding either an N-terminal half (N) or a C-terminal half (C) of the granulin domain. *Center panel*; The full-length human *progranulin* transcript and splice variants derived from the Alternative Splicing Database (ASD) (Bhandari et al., 1992; Stamm et al., 2006). Of those alternative transcripts listed in ASD only those with multiple independent ESTs to confirm validity are illustrated. *Lower panel*; The progranulin protein is comprised of one half and seven granulin motif repeats (purple boxes) with nomenclature associated to each granulin domain (A-G) based on the order in which they were characterized (Bateman et al., 1990; Bhandari and Bateman, 1992; Bhandari et al., 1992). P (paragranulin) represents the N-terminal half domain. The granulin motif consensus sequence is also illustrated (Bateman and Bennett, 1998).

however only two of these alternative gene products have been corroborated through isolation of multiple ESTs (Figure 1) (Stamm et al., 2006). In addition, it is unknown whether these splice variants encode a biologically active product since the bioactive domain(s) within *progranulin* have not been characterized.

1.1.3 The zebrafish *progranulin* family – *Progranulin* is represented as a single gene in mammals that is highly conserved across all species investigated, including human, mice, rats and the pig (Baba et al., 1993; Bhandari and Bateman, 1992; Bhandari et al., 1993; Yang et al., 2006). However, this is not true of all vertebrates. Of particular interest is the *teleost* clade of ray-finned fish that have been shown to harbour multiple *progranulin* genes. The zebrafish (*Danio rerio*) *progranulin* gene family has been well characterized and is represented by four independent genes of variable length (Figure 2) (Cadieux et al., 2005). In addition, Bennett and colleagues isolated a unique antisense transcript that encompassed the *progranulin-1* and *–2* genes and demonstrated the possibility of an antisense transcript associated with the 3' end of *progranulin-a* (Cadieux et al., 2005). Furthermore, a rare *trans*-splicing product between *progranulin-1* and *-2* transcripts was identified.

A large proportion of species for which sufficient EST and genomic sequencing data are available harbour a single *progranulin* gene, suggesting an evolutionary predisposition toward the maintenance or restricted representation of the *progranulin* gene (Figure 3) (Cadieux et al., 2005). The emergence of a *progranulin* gene family within *teleost* species such as the zebrafish is not necessarily unique, however it does provide a window of opportunity to investigate mechanisms of gene expansion and perhaps gain insight into the functional importance of this growth factor.

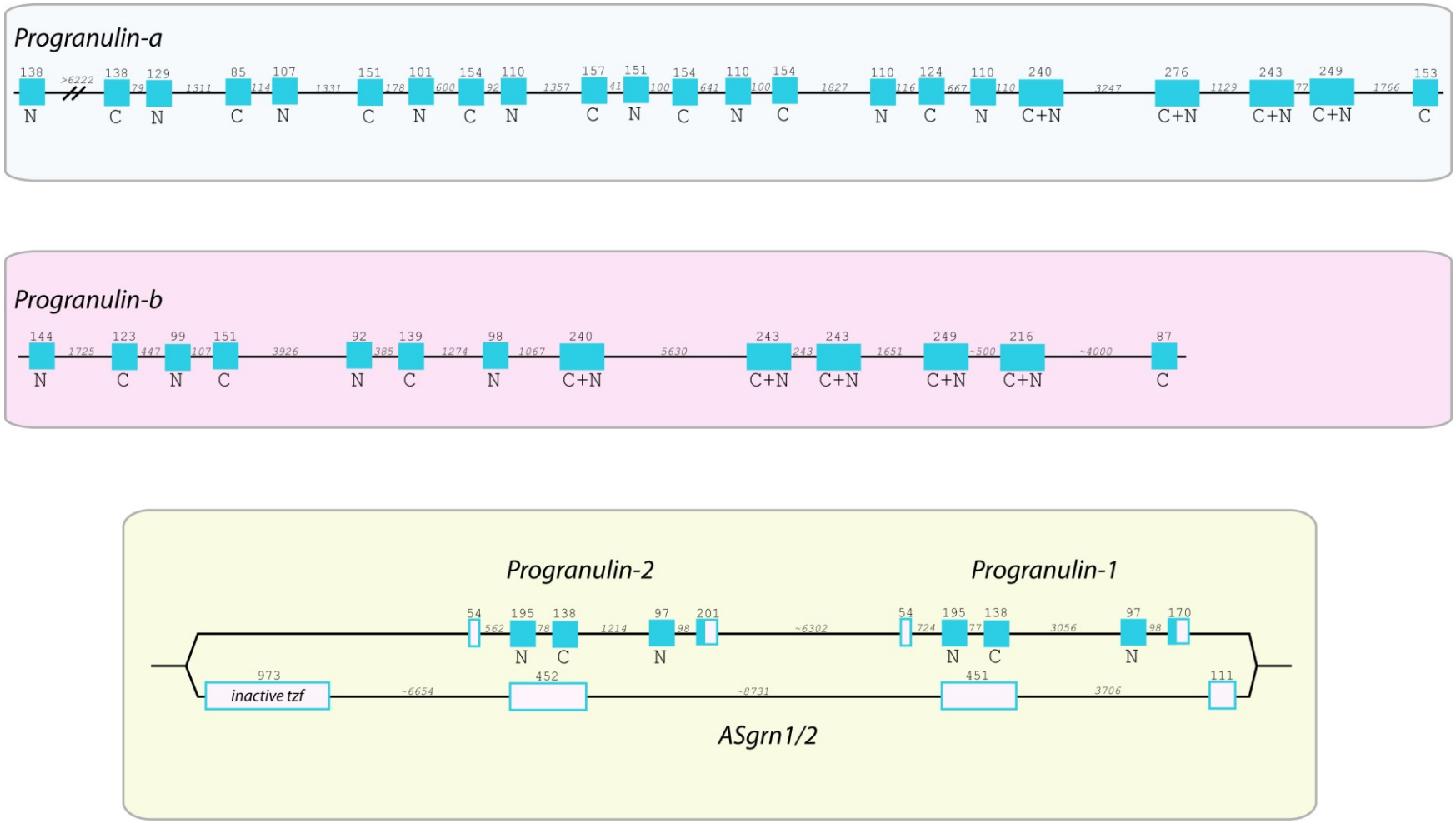

Progranulin-a
138 >6222 138 79 129 1311 85 114 107 1331 151 178 101 600 154 92 110 1357 157 41 151 100 154 641 110 100 154 1827 110 116 124 667 110 110 240 3247 276 1129 243 77 249 1766 153
N C N C N C N C N C N C N C N N C
Progranulin-b
144 1725 123 447 99 107 151 3926 92 385 139 1274 98 1067 240 5630 243 243 243 1651 249 ~500 216 ~4000 87
N C N C N C N C+N C+N C+N C+N C+N C+N C
Progranulin-2
Progranulin-1
54 562 195 78 138 1214 97 98 201 ~6302
N C N
54 724 195 77 138 3056 97 98 170
N C N
973 inactive tzf ~6654 452 ~8731 451 3706 111
ASgrn1/2

Figure 2. The zebrafish *progranulin* gene family. The zebrafish (*Danio rerio*) harbours four progranulin genes (Cadieux et al., 2005). The structure of each zebrafish progranulin gene was determined through the comparison of their published full-length cDNA sequences and the computational organization each gene from the Ensembl Project (Hubbard et al., 2007). The zebrafish Ensembl build Zv6 and the gene identifiers used are ENSDARG00000004954 (*Progranulin-a*), ENSDARG00000025081 (*Progranulin-b*), ENSDARG00000035818 (*Progranulin-1*) and ENSDARG00000035817 (*Progranulin-2*). *Upper panel*; *Progranulin-a* is located on linkage group 3 and is composed of twenty two coding exons that harbour thirteen granulin domains. *Progranulin-b* is located on linkage group 24 and is composed of thirteen coding exons that harbour nine granulin domains. *Progranulin-1* and *progranulin-2* are located on linkage group 19 and are composed of five exons that encode one full and one N-terminal granulin domain. Encompassing the tandem *progranulin-1* and *-2* is the antisense *ASgrn1/2* gene that lacks an open reading frame and harbours an inactive DNA transposon of the tzf subclass of the *Tc1/mariner* superamily (Cadieux et al., 2005).

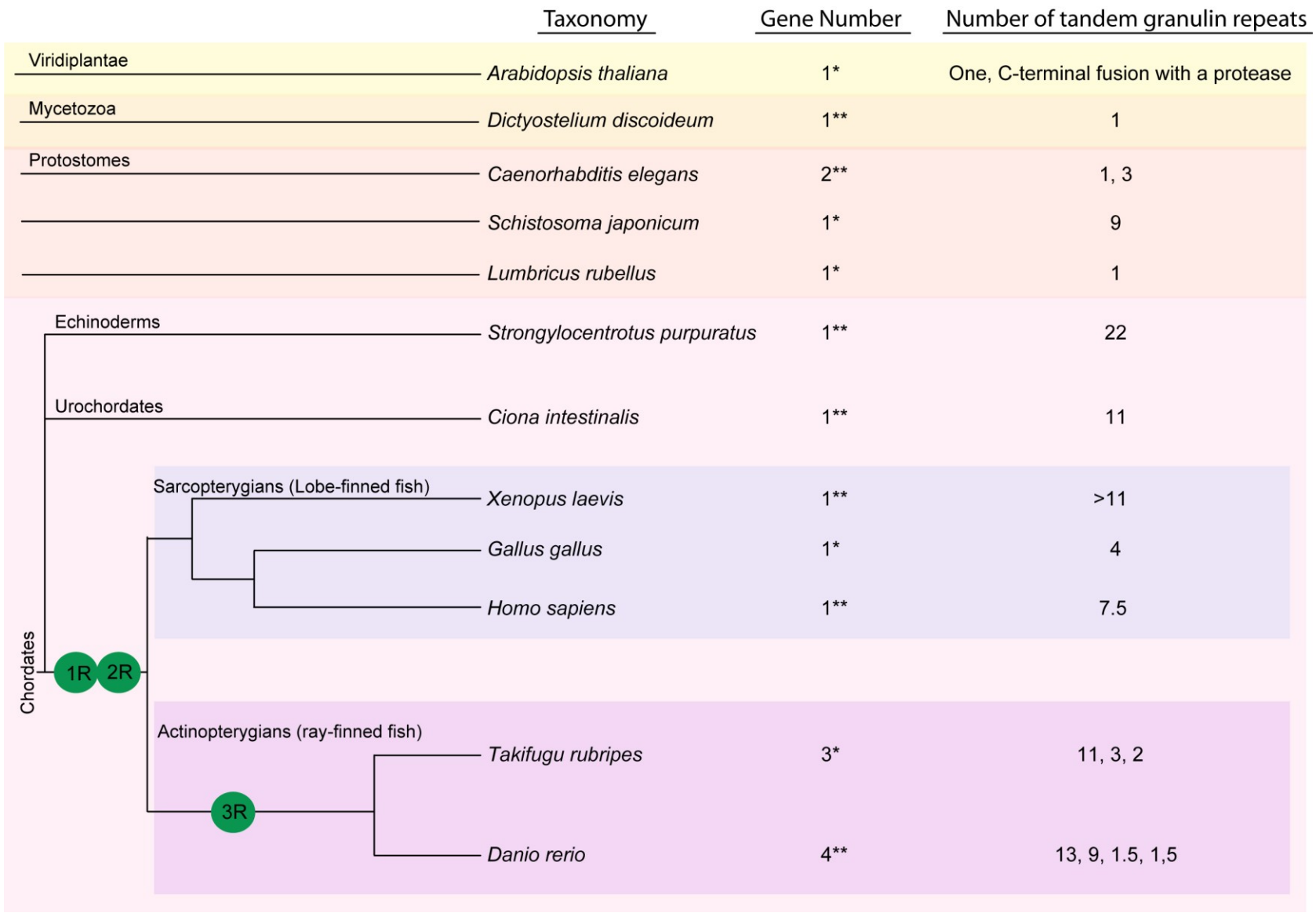

Taxonomy
Gene Number
Number of tandem granulin repeats
Viridiplantae
Arabidopsis thaliana
1*
One, C-terminal fusion with a protease
Mycetozoa
Dictyostelium discoideum
1**
1
Protostomes
Caenorhabditis elegans
2**
1, 3
Schistosoma japonicum
1*
9
Lumbricus rubellus
1*
1
Echinoderms
Strongylocentrotus purpuratus
1**
22
Urochordates
Ciona intestinalis
1**
11
Sarcopterygians (Lobe-finned fish)
Xenopus laevis
1**
>11
Gallus gallus
1*
4
Homo sapiens
1**
7.5
Chordates
1R
2R
Actinopterygians (ray-finned fish)
Takifugu rubripes
3*
11, 3, 2
3R
Danio rerio
4**
13, 9, 1.5, 1,5

Figure 3. Diagrammatic representation of species which harbour granulin genes. The estimated rounds of vertebrate genome duplication events are indicated (1R, 2R, 3R). For each species number of *progranulin* genes and the number of granulin domains are illustrated. (*) All *progranulin* genes and encoded granulin domains were derived from expressed sequence tag (EST) databases and partial genomic sequences when available as described in Cadieux et al., (Cadieux et al., 2005). (**) When possible the annotated architectures for each species were further characterized from available genome projects. Tabulated databases, genome builds and gene identifiers for each species are listed in section 2.4.

1.2 The Evolution of *Progranulin* Genes

1.2.1 Genome-wide duplications and vertebrate evolution – As outlined above, mammalian genomes harbour a single copy of the *progranulin* gene. This is unusual given the theoretical occurrence of at least two genome-wide duplication events within the deuterostome lineage and the 1:2:4 maxim of gene multiplication (Figure 3) (McLysaght et al., 2002; S., 1970; Wolfe and Li, 2003). This is generally thought to be the basis for the occurrence of growth factors and associated receptors as members of multi-gene families (e.g., the four FGF receptors) (Itoh and Ornitz, 2004). It is believed that the two rounds of genome-wide duplication, commonly referred to as the 2R hypothesis, occurred relatively early within the chordate species, with only a few extant chordate groups (*Urochordates* and *Cephalochordates*) having escaped these primordial events (Figure 3) (Dehal et al., 2002; Panopoulou et al., 2003).

It has been argued that the expansion of chordate gene families as compared to their protostome ancestors could be simply a reflection of multiple independent segmental duplications throughout evolution (Hughes, 1999; Hughes and Friedman, 2003). However, the most compelling evidence for genome-wide duplications is the maintenance of the syntenic relationships between paralogous gene clusters, such that large chromosomal blocks appear to have originated from a common root (Abi-Rached et al., 2002; Lundin et al., 2003; Vienne et al., 2003). Perhaps the best described example of this phenomenon is provide by the Hox gene clusters, which clearly display evidence of two rounds of genome-wide duplication during early chordate evolution (Figure 4) (Di Gregorio et al., 1995; Garcia-Fernandez and Holland, 1994; McGinnis and Krumlauf, 1992). The construction of large chromosomal phylogenetic trees between founder,

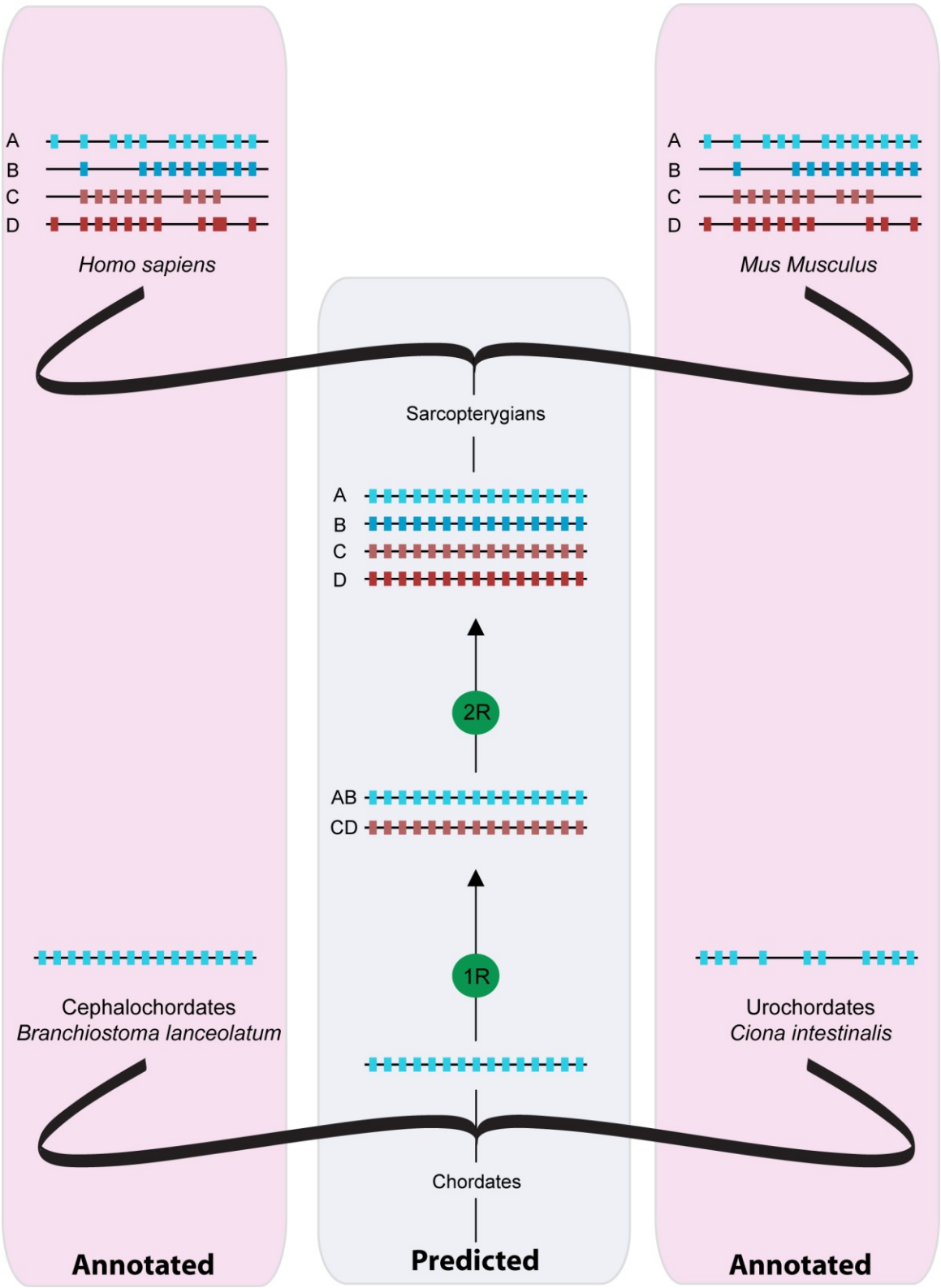

A
B
C
D
Homo sapiens
A
B
C
D
Mus Musculus
Sarcopterygians
A
B
C
D
2R
AB
CD
1R
Cephalochordates
Branchiostoma lanceolatum
Urochordates
Ciona intestinalis
Chordates
Annotated
Predicted
Annotated

Figure 4. The prototypic Hox cluster family as evidence for the occurrence of two rounds of genome-wide duplication during early chordate evolution. *Center panel*; The predicted number of Hox gene clusters found in lobe-finned fish (*sarcopterygians*) that includes mammalian species based on the presence of a single ancestral cluster at the base of the chordate radiation (Garcia-Fernandez and Holland, 1994; McGinnis and Krumlauf, 1992). *Left and right panels*; Annotated Hox gene clusters within two basal chordate phylogenies (Cephalocordates and Urochordates) as illustrated with the architectures of amphioxus (*Branchiostoma lanceolatum*) and the common sea squirt (*Ciona intestinalis*) (Dehal et al., 2002; Panopoulou et al., 2003). Annotated Hox gene clusters of humans (*Homo sapiens*) and mouse (*Mus musculus*) (Bailey et al., 1997; Meyer and Van de Peer, 2005).

extent and chordate species, such as that constructed for the Hox clusters for numerous species, does not always imply clear chromosomal linkage (Horton et al., 2003; Hughes, 1999). However, it is expected that as the chordate ancestors underwent rediploidization as well as the potential for other genomic rearrangements independent of the genome-wide duplications, the clarity of cluster linkage would become obscured (Gu and Huang, 2002; Horton et al., 2003). The role of other mechanisms of genomic rearrangement (e.g. deletion and segmental duplications, ecetera) have altered the representation of many gene families (Hennig, 2004; Martin et al., 2004; Rabbitts et al., 1980). Hence, the 1:2:4 dogma as the basis of gene representation in mammals is not absolute. For example, the four mammalian fibroblast growth factor receptor (FGFR) genes adhere to the 1:2:4 dogma, while in contrast there are currently twenty three known FGFR ligands (Itoh and Ornitz, 2004).

Although the 2R hypothesis has gained some acceptance, it remains a model that continues to be both challenged and supported by ongoing research. A more convincing description (bordering on unequivocal) is the evidence that a recent genome-wide duplication (3R) has occurred within the ray-finned (*teleost*) lineage of actinopterygians, generating a 1:2:4:8 maxim of gene representation. The zebrafish and other members of the Teleostei lineage (including *Takifugu rubripes* and *Tetraodon nigroviridis* pufferfishes) have undoubtedly undergone genome-wide duplication that occurred at the base of the teleost radiation (Postlethwait et al., 1998; Taylor et al., 2003). The duplication of a large number of genes in comparison to mammals has been well characterized for several species of this clade, with the Hox gene clusters representing the prototypic example (Figure 5) (Amores et al., 1998; Amores et al., 2004; Jaillon et al., 2004; Postlethwait et al., 1998; Woltering and Durston, 2006). Overall, it has been

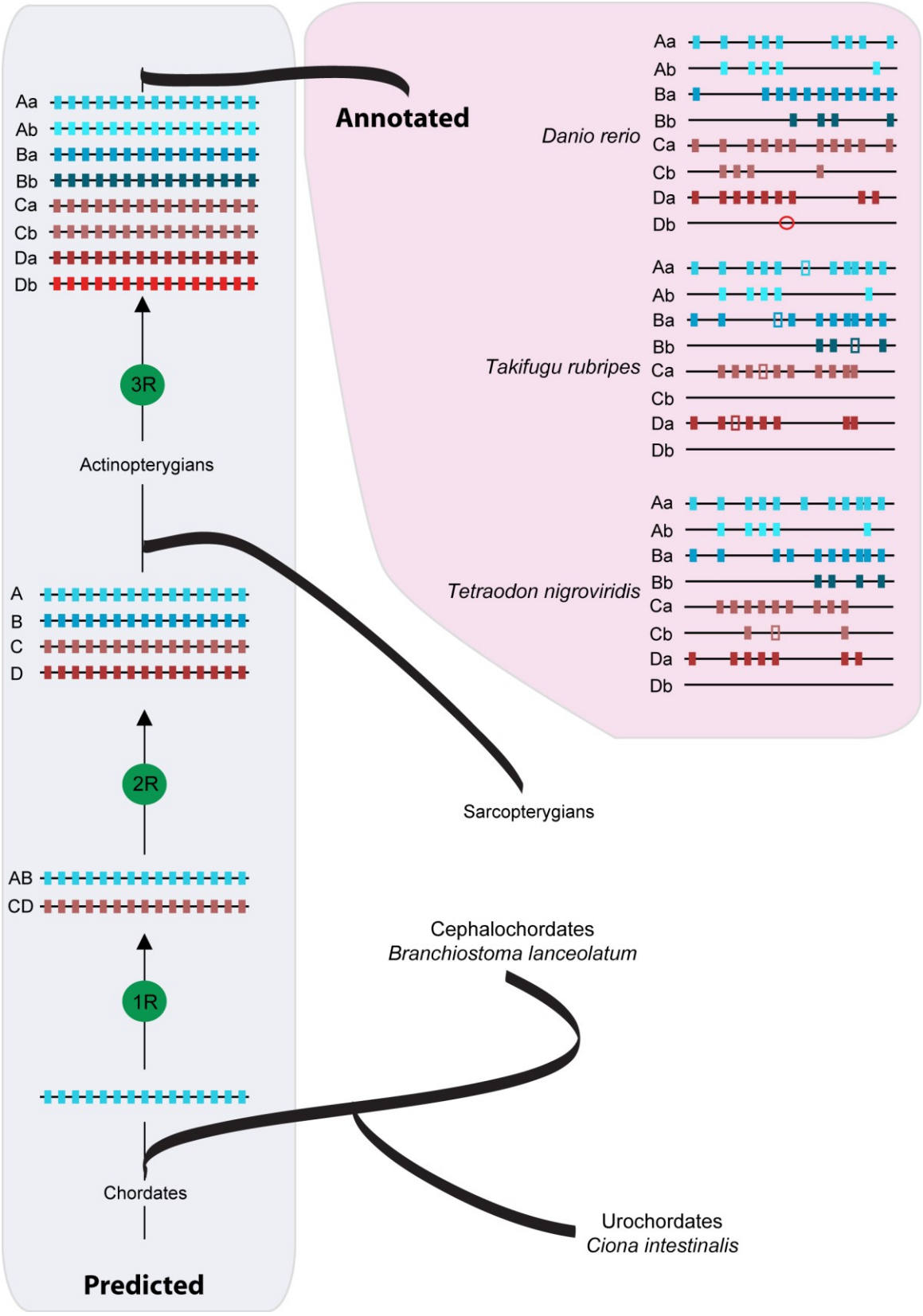

Annotated
Aa
Ab
Ba
Bb
Ca
Cb
Da
Db
Danio rerio
Takifugu rubripes
Tetraodon nigroviridis
3R
Actinopterygians
A
B
C
D
2R
Sarcopterygians
AB
CD
1R
Cephalochordates
Branchiostoma lanceolatum
Chordates
Urochordates
Ciona intestinalis
Predicted

Figure 5. The prototypic Hox cluster family as evidence for the occurrence of a teleost specific genome-wide duplication. *Left panel*; The predicted number of Hox gene clusters found in ray-finned fish (*Actinopterygians*) that includes teleost species and based on the presence of a single ancestral cluster at the base of the chordate radiation (Postlethwait et al., 1998; Taylor et al., 2003). *Right panel*; Annotated Hox gene clusters for the zebrafish (*Danio rerio*) and two pufferfish species (*Takifugu rubripes* and *Tetraodon nigroviridis*) (Amores et al., 1998; Amores et al., 2004; Jaillon et al., 2004; Woltering and Durston, 2006).

estimated that 20%-50% of human genes are represented by two co-orthologues in zebrafish (Figure 5) (Amores et al., 1998; Postlethwait et al., 2000; Woods et al., 2000). The incomplete retention of all the products of 3R is to be expected since a large proportion of gene duplicates were likely to be redundant, acquire deleterious mutations and/or lost as a consequence of the subsequent reversion to the diploid state.

1.2.2 The consequences of genome-wide duplication events and speciation – The concept of gene duplication, including polyploidization events, originated with Susumo Ohno who hypothesized that such mechanisms likely play an integral role in evolution as the original and novel gene function(s) could be partitioned across the two orthologues (Ohno et al., 1968). This concept undoubtedly prompted the research community to investigate and identify such duplication events. Indeed, much research has been devoted to determining the validity (or lack thereof) of ancestral genome-wide duplications. It is difficult to assess the consequences of such phenomena with respect phenotypic and/or evolutionary relevance because of the time required for the effects to become apparent. However, several studies in yeast support the hypothesis that genome-wide duplications can be advantageous to this organism (Kim and Yi, 2006; Tirosh and Barkai, 2007; Wagner, 2002). Although the significance of the 2R hypothesis is of interest for all chordate species, the consequences are not clearly discernable after ~500 million years ago (mya) since genomes are never static entities (Panopoulou et al., 2003). Therefore, the association between genome-wide duplication events as an underlying cause of the expansion of chordate species is qualitatively suggestive. However, the evidence for this link is inconclusive overall. The fact that the teleost specific 3R is a more recent event (~350mya) implies that these species may offer better insight with regard to the impact of

genome-wide duplication during evolution (Postlethwait et al., 1998). It has been noted that the Teleostei lineage of ray-finned fish (including zebrafish) is comprised of ~25,000 extant species, the largest diversification within all vertebrate lineages (Nelson, 1994). In contrast, those ray-finned species that are believed to have diverged prior to the occurrence of the 3R genome-wide duplication (sharks, sturgeons, etc...) have not undergone such wide-spread speciation (Hoegg et al., 2004; Venkatesh et al., 2007). This observation supports Susumo Ohno's original hypothesis, that genome-wide duplication provides the raw material for rapid evolution, assuming that the presence of co-orthologous genes relinquishes one copy from functional constraint and allows for asymmetric molecular evolution to occur between co-orthologues (Ohno et al., 1968). The investigation of the functional equivalence or exclusivity between co-orthologous genes in relation to their mammalian counterpart, is a major experimental advantage for the use of the zebrafish as an experimental model to assess gene function.

1.2.3 Co-orthologous pairs in comparative zebrafish biology - The mechanisms underlying the retention, segmentation or neofunctionalization of gene duplicates are issues that pervade comparative biological investigations involving the zebrafish model of vertebrate development (Force et al., 1999). It is expected that the functions of certain duplicated genes will be novel, given that the experimental model is a species whose phylogeny and morphology diverge significantly from that of mammals. Of particular interest is the degree to which ancestral biological activities are segmented across gene duplicates, with the expectation of identifying multiple functions that would otherwise remain obscure when analyzing the single mammalian orthologue (Force et al., 1999). For example, the murine transforming growth factor-beta (TFG-β) family member Nodal

has been shown to be involved in the development of the mesendoderm during early gastrulation and its biological function is mediated through the formation of a concentration gradient, suggesting that this growth factor may act as a morphogen in this context (Lowe et al., 2001). In zebrafish, Nodal is represented by two co-orthologous genes commonly referred to as *squint* and *cyclops,* a nomenclature that reflects their initial characterization from a large-scale ethylnitrosourea (ENU) mutagenesis screen (Feldman et al., 1998). Both *squint* and *cyclops* are involved in zebrafish mesendoderm development, however only *squint* acts as a morphogen (Chen and Schier, 2001). Further investigation of the molecular distinctions between *squint* and *cyclops* using chimeric constructs demonstrated that the "morphogen" bioactivity is located within a 29 amino acid N-terminal domain of *squint* and the mutation of four acidic residues within this domain is sufficient to eliminate its long-range signalling capacity (Jing et al., 2006). Several other examples of the segmentation or the complementation of ancestral gene function between zebrafish co-orthologous include sonic hedgehog (shh) and tiggywinkle hedgehog (twhh) during primary motoneuron development as well as *sox9a* and *sox9b* in craniofacial morphogenesis (Lewis and Eisen, 2001; Yan et al., 2005).

Several attempts to rationalize the asymmetric functional segmentation between co-orthologues has spawned numerous theoretical concepts including the duplication, divergence and complementation (DDC) model (Force et al., 1999). Although the DDC model became a popular concept during the earlier studies of zebrafish comparative biology, it has not proven to be a worthwhile generality since the complexity of gene function rarely conforms to such imperfect generalities. The major presupposition of the DDC and other models is that the function of the mammalian gene has already been investigated, thereby allowing for an assessment of functional divergence and/or

complementation. This highlights a strong theme in the majority of developmental biology studies using the zebrafish, wherein genes that have been previously investigated in mouse models are often the first to be explored. In contrast, the investigation and assignment of function to novel genes is not nearly as common in zebrafish since issues of functional conservation across species are immediately brought into question. Furthermore, the assignment of orthology or parology may not be clear, as the degree of syntenic correspondence is variable and genetic architectures may diverge widely within large gene superfamilies.

These factors are of particular importance with respect to the zebrafish *progranulin* gene family, which harbours four genes of variable architecture (Cadieux et al., 2005). Therefore, it is important to understand the ancestral origins of each progranulin gene in order to determine which is representative of the common chordate ancestor and, by extension, the orthologue most likely to provide information regarding *progranulin* function in mammals.

1.2.4 Genome-wide duplications and the evolution of *progranulins* – Cadieux et al., previously demonstrated, through EST database mining and *in silico* cloning, that the *progranulin* gene is likely represented in plants, metazoan, protostome and urochordate lineages as single genes (Figure 3) (Cadieux et al., 2005). The prevalence of organisms harbouring a single *progranulin* gene prior to vertebrate evolution and application of the 3R hypothesis with its 1:2:4:8 maxim predicts that the zebrafish may harbour eight *progranulin* orthologues. However, the presence of a single *progranulin* gene in mammalian species would suggest early chordate genome-wide duplications (2R) had ultimately no effect on *progranulin* gene number; therefore, the teleost 3R event would be

theoretically restricted to generating a maximum of two *progranulin* genes. Curiously, the zebrafish possesses four distinct *progranulin* genes (Cadieux et al., 2005). *Progranulin-a* and *-b* encode proteins structurally homologous to the mammalian form, whereas *progranulin-1* and *-2* are *progranulin* paralogues whose structures are found only in teleosts (Figure 2) (Cadieux et al., 2005). The presence of four *progranulin* genes in zebrafish provokes several questions, including the issue as to the number of ancestral *progranulin* genes existing prior to the formation of the teleost radiation. The archetypal ancestral gene is likely represented by the single *progranulin* gene found in the slime mold (*Dictyostelium discoidium*) that consists of a single granulin motif (Figure3). The presence of two *progranulin* genes within the nematode worm *Caenorabiditis elegans* does not necessarily conflict with this argument, since it is the only protostome to harbour *progranulin* co-orthologous. Therefore, it is currently presumed that the common chordate ancestor harboured a single *progranulin* gene. If this ancestral chordate *progranulin* gene is directly related to each of the zebrafish *progranulin* genes, the genetic evidence for this relationship is likely to extend beyond simple gene architecture. Each *progranulin* gene is likely to reside within a genomic milieu reminiscent of their past, therefore the determination of putative syntenic relationships may assist in assigning origin. In addition, this may indicate which zebrafish gene(s) is the true representative of the common ancestral gene.

The zebrafish *progranulin-a* gene is located on linkage group 3 (annotated chromosome 3 based on Sanger Centre nomenclature) which shares a large block of syntenic correspondence surrounding the human *progranulin* located on chromosome 17 (Cadieux et al., 2005; Postlethwait et al., 2000). In contrast to this clear syntenic association, *progranulin-b* is located within a region that shares no phylogenetic cluster

that can be associated with human *progranulin* (Cadieux et al., 2005). Although *progranulin-a* and *–b* share similar genetic architecture (Figure 2 and 6) and their developmental expression patterns are similar to those described for murine *progranulin* the formation of this orthologous pair cannot be resolved as originating from the 3R genome-wide duplication or a limited segmental duplication and translocation event (Cadieux et al., 2005; Daniel et al., 2003). Despite these inconsistencies *progranulin-a* and *–b* most likely represent the co-orthologues of mammalian *progranulin*.

In contrast, *progranulin-1* and *-2* harbour an architecture that is particularly unique with respect to vertebrate *progranulin* genes. Specifically *progranulin-1* and *-2* are located in a tandem head-to-tail fashion along the same strand of zebrafish chromosome 19 (Figure 2) (Cadieux et al., 2005). Both *progranulin-1* and *-2* harbour one and one half granulin domains. These genes are distinguished from all other *progranulin* gene architectures in that the final coding exon consists of an open reading frame of only three amino acids, rather than an archetypal C-terminal granulin motif (Figure 2 and 6). In addition, *progranulin-1* and *-2* are associated with an antisense gene (*ASgrn1/2*), which shares partial exonic complementarity and contains an inactive member of the teleost-specific *Tc1/mariner* transposon superfamily (Cadieux et al., 2005; Lam et al., 1996). Despite the novelty in the genomic architectures of *progranulin-1* and *-2*, these genes do reside within a region harbouring a certain degree of synteny with human *progranulin* (Cadieux et al., 2005). Due to their nearly identical architecture and coding identity it is likely that these paralogues originated from a tandem segmental duplication of a common ancestor. This event is given further credence due to the existence of the *ASgrn1/2* with its archaic class II DNA transposon (Cadieux et al., 2005; Lam et al., 1996). As to the origin of the *progranulin-1/2* ancestor, the presence of some syntenic correspondence to

human chromosome 17 (encompassing the *progranulin* gene) would suggest that this gene may have been present within early chordates although its retention was restricted to actinopterygians. However, the unique architecture of these genes with respect to the final coding exon cannot be ignored. In addition, *progranulin-1* and *-2* have highly restricted expression patterns, unlike *progranulin-a* and *–b* and mammalian *progranulin*.

Although phylogenetic synteny allowed for the clear assignment of *progranulin-a* as a *bona fide* orthologue of the mammalian counterpart, the origins of the other three family members is partly based on common or unique genomic architectures associated with the number of encoded granulin domains, particularly for *progranulin-b*. This presumption is valid only if the genetic architectures of *progranulin* genes remain relatively static. This is not the case for most species as the number of encoded granulin domains contained with each *progranulin* gene varies to a high degree. This suggests that, despite the similar architectures of *progranulin-a* and *–b* the lack of chromosomal synteny between these genes argues that they may not share a common ancestor. The underlying question becomes whether the expansion and/or contraction of *progranulin* genes is a common occurrence and how tolerant is the gene to extreme deletions and/or amplifications of the granulin domain?

1.3 Progranulin protein-protein interactions

1.3.1 Progranulin and the yeast two-hybrid system – The investigation of the nature of protein complexes will provide insights into the mechanisms that translate the intrinsic qualities of a particular protein into a molecular and cellular response. The combination of the information derived from complete genomic sequencing and high-throughput screening techniques has facilitated the definition of some protein-protein complexes.

This is exemplified by the development of several techniques including yeast two-hybrid high-throughput screen (Fields and Song, 1989). The identification of progranulin binding proteins has been focus of many investigations, for which the Y2H system has been utilized almost exclusively. This system has identified a plethora of apparent progranulin binding proteins including secretory leukocyte protease inhibitor (SLPI), cartilage oligomeric matrix protein (COMP), perlecan, cyclin T1 and its associated complex, delta-like 1 (DLK1), hexokinase type III (HK3), family with sequence similarity 131, member C (FAM131C), HoxA1, SMAD9 and topoisomerase IIIβ (Baladron et al., 2002; Gonzalez et al., 2003; Hoque et al., 2005; Hoque et al., 2003; Rual et al., 2005; Trinh et al., 1999; Xu et al., 2007; Zhu et al., 2002). Of these, only SLPI, COMP and perlecan have been confirmed to associate to progranulin using independent co-immunoprecipitation assays. In addition, perlecan and COMP demonstrate a synergistic affect when incubated together with progranulin in cell proliferation assays (Gonzalez et al., 2003; Xu et al., 2007). With respect to the putative interaction between progranulin and HK3, cycltin T1, FAM131C, HoxA1 and topoisomerase IIIβ, these proteins are localized within the cytoplasm and/or nucleus while progranulin has been clearly demonstrated to be a secreted growth factor with a *bona fide* signal peptide (Diaz-Cueto et al., 2000; Qin et al., 2005a; Xu et al., 2007; Zhou et al., 1993). Pe'ery and colleagues, having previously identified a putative interaction between progranulin and the Tat component of the cyclin T1 complex, attempted to assess whther progranulin or truncated constructs localized to the nucleus. Indeed, a truncated progranulin molecule consisting of the three C-terminal motifs C,D and E (see Figure 1) colocalized with Tat to the nucleus (Hoque et al., 2005). However, the construct did not contain the endogenous signal peptide and therefore it remains unclear as to the validity and significance of this

putative protein-protein interaction. Although it remains possible that progranulin does indeed enter these cell compartments before or after secretion, this has yet to be demonstrated. Furthermore, the Y2H results for progranulin should be treated with some caution since the yeast intracellular environment does not ensure proper folding of such a highly disulfide bridged protein. Despite these general incompatibilities the Y2H system has proven effective in identifying putative progranulin binding proteins.

To this date, only one other method beside the Y2H system has been used to identify granulin binding proteins. Shoyab and colleagues used ^{125}I-labeled granulin peptides and cross-linking reagents to identify a putative ~140kDa membrane bound binding protein from human breast carncinoma cells (Culouscou et al., 1993). Fifteen years after its initial identification, the characterization of this putative receptor has not been published. It is not surprising that Y2H studies have failed to isolate a potential membrane bound protein for progranulin as the system is clearly incompatible with transmembrane proteins. Therefore, an alternative system for the identification of protein-protein interactions may be beneficial in uncovering other progranulin interacting species.

1.3.2 Alternative protein-protein interaction reporters – The identification of protein-protein interactions using the Y2H system work optimally for the discovery of intracellular binding partners, but are not appropriate for products of the secretory pathway that include ligands, their binding proteins and cognate receptors. As previously mentioned, several putative progranulin binding proteins have been identified using the yeast two-hybrid system with varying degrees of validation (Gonzalez et al., 2003; Shoham et al., 2003; Sui and Wilson, 2000; Thornburg et al., 2004; Trinh et al., 1999; Xu

et al., 2007; Zhu et al., 2002). However, none of these proteins are membrane bound nor have any been demonstrated to augment the intracellular cascades that are activated by progranulin.

The overall success of the Y2H systems has spawned several variations including the bacterial and mammalian two-hybrid systems that have proven worthwhile due to increased through-put and proper post-translational modification, respectively (Karimova et al., 1998; Vasavada et al., 1991). However, the classical reporter system used in these two-hybrid methodologies remains dependent upon an association between the requisite modules within the nucleus. Slight variations in reporter construction has given rise to Y2H systems that allow for the dectection of cytoplasmic protein-protein interactions that augment Ras and or G-protein coupled receptor signaling (Aronheim et al., 1997; Broder et al., 1998; Ehrhard et al., 2000). Unfortunately, these systems do not allow for the screening of extracellular protein-protein interactions. At present there is no *in vitro* reporter system that is compatible for screening extracellular protein-protein interactions.

The major alternative to two-hybrid screening focuses on a departure from the measurement of an *in vitro* quantitative biological response, for systems that favour purification of protein complexes followed by direct mass spectrometric (MS) identification. Such methodologies did not gain the same early popularity as Y2H due to the complexity of mass spectrometry and the lack of protein sequence databases. The continuing deposition of expressed sequenced tag (EST) information, completion of numerous genome sequencing projects and associated bioinformatics as well as the exponential growth of MS engineering has allowed for a parallel increase in protein-protein interaction screens that rely on these methodologies. Perhaps the most lauded examples of large-scale protein complex purification and subsequent MS characterization

rely on tandem affinity purification tag (TAP-tag), which has been used in a number of contexts including yeast and mammalian cells (Burckstummer et al., 2006; Gavin et al., 2002; Rigaut et al., 1999). However, the TAP-tag system is still reliant upon the addition of epitope tags that can theoretically inactivate or augment the normal function of a protein. Protein microarrays are similarly disadvantageous, due to requisite protein covalent modification and prohibitive costs involved in the large-scale purification and construction of these screens (Talapatra et al., 2002).

Recently, innovations in thin-layer optical sensors have been wedded to microfluidic systems to allow for the quantitative analysis of kinetic parameters between a conjugated "bait" and complex solutes (Canziani et al., 1999). Commonly referred to as surface plasmon resonance (SPR), these biosensors use monochromatic light that is reflected from a thin metallic sheet (typically gold) at a specific angle based upon the surface density. The opposing surface is contained within a microfluidic system and is typically comprised of a dextran coating along with a covalently attached ligand of interest (i.e. progranulin). Upon addition of a solute that contains putative binding proteins, any interaction between solute and the surface produces an increase in local surface density. A change in density translates into a change in the angular reflection of the monochromatic beam. Using SPR for the detection of protein-protein interactions not only assesses specificity, but also distinguishes whether the observed kinetics are typical of high-affinity interactions (Canziani et al., 1999).

Although the use of SPR for the detection of protein-protein interactions provides a highly sensitive reporter method, there is no intrinsic proteome associated with the analysis. Therefore a relevant source material must be chosen that reflects a tissue or cell type that is derived from phyiosogical or pathological state of interest and relevance.

1.3.3 Alternatives to the yeast bioreactor in protein-protein screening – The ideal system for the direct protein-protein interaction screening requires a physiologically relevant source of ligand and interacting partners. The general mechanistics of growth factor signaling are highly conserved across species however, the variation in the elements of signal transduction from species to species argues for a conservative approach when hypothesizing their physiological import in humans. Regardless, experimental investigation of the physiology of several diverse animals, presumably far removed from the evolutionary pathways of mammalian ancestors, has proven directly applicable. For example, the gila monster (*Heloderma horridum*) Exendin-4, a bioactive orthologue of the human glucagon like peptide 1 (GLP-1), has proven naturally resistant to proteolytic inactivation and is currently in phase III clinical trials for treatment of type 2 diabetes (Yoo et al., 2006). Similarly, commercially available calcitonin (Miacalcin), a peptide growth factor commonly used for the treatment of postmenopausal osteoporosis and hypercalcemia for several decades, is derived from salmon (Jowsey et al., 1978; Wisneski, 1990). Furthermore, many growth factors and modulators of signaling cascades were initially characterized using large-scale chemical mutagenesis screens in developmental models. For example, the Wnt family (Wingless), BMPs (Decapentaplegic) and hedgehog family were first characterized as growth factors with morphogenetic properties *in vivo* using the fruit fly (*Drosophila melanogaster*) (Mohler, 1988; Padgett et al., 1987; Rijsewijk et al., 1987). A similar screen using zebrafish (*Danio rerio*) proved equally valuable in defining novel morphogenetic pathways (Tabata and Takei, 2004). The fact that these models of embryonic development have been useful for the identification of signaling components stems from the requirement of numerous

growth factors for germ layer specification and subsequent differentiation. However, a method of directly characterizing the signaling components of a particular orphan growth factor has not been applied to these developmental models.

The zebrafish has become increasingly popular in developmental biology research due to its *ex utero* development, optical clarity, the ability to manipulation gene function, high-throughput, real-time observation and low cost in comparison to similar murine experimental models (Eisen, 1996). Furthermore, it is not necessary to maintain a large zebrafish colony as large numbers (4,000-100,000) of staged zebrafish embryos can be obtained commercially (Aquatica Tropicals, Plant City, CA). Besides zebrafish, the frog (*Xenopus tropicalis*) can also be considered an alternative vertebrate model, although its genome sequencing project, available molecular markers and the applicability of antisense technology for functional experiments are all in their infancy (Carruthers and Stemple, 2006). Clearly, the cost involved in generating a similar source of tissue from mice is not feasible. For this reason the majority of protein-protein interaction studies (including the mammalian two-hybrid system) rely on immortalized cell lines derived from various tissues (Dang et al., 1991; Vasavada et al., 1991). Although the use of *in vitro* cell culture will continually play a role in evaluating signaling mechanisms, the *de novo* identification of protein complexes from immortalized cells is potentially problematic since they may not reflect physiologically relevant interactions. This is not only due to the genetic and/or epigenetic mechanisms that allow cell line proliferation *ex vivo*, but the absence of surrounding tissues which may provide paracrine factors necessary for normal protein function.

The use of zebrafish in large-scale protein-protein interaction screens may be of interest, however the model remains almost exclusively utilized in developmental

biology. Although this may be considered a disadvantage for initiating a proteomic screen using the zebrafish it may prove particularly beneficial for subsequent validation and assessment of the biological relevance of the protein-protein interactions that are identified.

1.4 The zebrafish model

1.4.1 The zebrafish model of vertebrate development – Perhaps the most expansive step forward in our understanding of developmental biology and the mechanisms which govern embryogenesis and organogenesis was the ENU-mutagenesis screening of the fruitfly (*Drosophila melanogaster*), for which the Nobel prize was awarded to Edward B. Lewis, Christiane Nusslein-Volhard and Eric Wieschaus in 1995 (Etcheverry, 1995). Following in the footsteps of this work, the zebrafish has proven equally amenable to chemical mutagenesis and the plethora of data that resulted, along with the development of key *in situ*, histochemical and immunological assays as well as chronological developmental landmarks, was sufficient to fill an entire volume of the journal of Development in 1996 (Haffter et al., 1996). This initial screen of zebrafish developmental mutants identified 4,264 mutants that continue to be a source of ongoing research within the community. The success of chemical mutagenesis has prompted the development of similar large-scale projects in mice, however the overall cost and archiving of mutant strains as well as their subsequent characterization cannot compare with the equivalent high-throughput capacity of small animal models, such as the zebrafish and fruitfly (Hrabe de Angelis et al., 2000; Nolan et al., 2000). The difficult and time consuming nature of isolating chemically-induced mutations through positional cloning has lead to the development of several insertional mutagenesis screens using

various viral vectors from which the identification of the affected gene is possible with simple polymerase chain reaction experiments (Amsterdam et al., 1999; Gaiano et al., 1996; Golling et al., 2002). Interstingly, the high-throughput of the zebrafish viral insertion method and simplicity of subsequent characterization of the insertion site has been commercialized with gene-specific strains available on a "cost-per strain" basis (Znomics Inc., Portland, OR).

In large-scale screens that employ viral or chemical mutagens, the preliminary characterization involves documenting the resultant phenotypic response to the genetic modification, followed by identification of the affected gene. This experimental approach of following phenotype to genotype is commonly referred to as "forward genetics". The reciprocal experiment of "reverse genetics", wherein a specific gene is targeted for inactivation in order to identify its functional relevance, has proven equally feasible in the zebrafish and represents the methodology used in the present study to assess the function of the zebrafish *progranulin* gene family.

1.4.2 Antisense technology and the zebrafish model - Although some progress has been made recently, the zebrafish has yet to benefit fully from gene-specific knockout technology nor have the powerful tissue-specific Cre/Lox systems been adapted to this model (Langenau et al., 2005; Liu et al., 2007a; Thummel et al., 2005). In response to this deficit a technique referred to as targeting induced local lesions in genomes (TILLING) has been developed (Sood et al., 2006). However, the large breeding and sequencing platform that is required to screen for gene-specific mutations using TILLING has limited its usefulness. Due to the *ex utero* development of the zebrafish it is possible to introduce exogenous molecules through microinjection to individual animals, allowing for the

application of antisense technology to investigate gene function. The mainstream alternative to gene inactivation, particularly in cell culture, involves the use short interfering RNA (siRNA) that encode for a specific antisense sequence in relation to the target mRNA (Fire et al., 1998; Zamore et al., 2000). The formation of RNA duplexes causes the activation of the RNaseIII Dicer enzyme and its associated complex pathway resulting in rapid turnover of the targeted mRNA (Zhang et al., 2004). Although the successful use of siRNA in zebrafish has been reported recently, several initial attempts demonstrated a large degree of non-specific targeting (Liu et al., 2005; Mangos et al., 2001; Oates et al., 2000). As a rapid and cost-effective alternative, antisense morpholino oligonucleotides (MO) are used routinely to effectively inhibit protein translation during the first 2−4 days of development to achieve an experimental 'knockdown' that is phenotypically equivalent to a gene knockout (Ekker, 2000; Nasevicius and Ekker, 2000). Within this window of zebrafish development almost all aspects of organogenesis can be assessed, with the exception of sexual differentiation (Ekker, 2000). The metabolic stability of MOs *in vivo* is primarily attributed to their resistance to RNaseH dependent degradation (Summerton, 1999). MOs are designed to be complementary to mRNA sequences close and 5' to the site of translation initiation, presumably blocking the binding of the 40S ribosomal complex to the starter AUG codon (Figure 3A) (Summerton, 1999; Summerton and Weller, 1997). Since MOs do not alter the genome, animals produced in this way are referred to as morphants. For clarity, the text that follows uses lowercase when referring to a gene or transcript (e.g., an MO directed against a *progranulin* transcript, progranulin-a-MO), whereas uppercase refers to the protein (e.g., progranulin-A morphant).

1.4.3 Comparative Developmental Biology and Progranulins – A large proportion of scientific investigation utilizing the zebrafish involves homologues for which the functional relevance of the mammalian gene is already known. At present there is no definitive statement as to the putative lethality of a *progranulin* null genotype, however several *in vitro* experiments have suggested that this outcome is probable (Diaz-Cueto et al., 2000; Qin et al., 2005a). These experiments demonstrated a role for progranulin during pre- and post-implantation. Whether such a null phenotype would be reproduced in the zebrafish after MO-targeted inactivation of one or more *progranulin* gene products does not necessarily follow since this animal develops *ex utero*.

The use of MOs to effectively inactivate a gene product is strongly dependent upon the efficacy of the MO itself as well as the temporal expression pattern of the gene of interest. With regards to the latter, both *progranulin-a* and *–b* are expressed to varying degrees throughout development in a variety of tissues (Cadieux et al., 2005). In contrast, the paralogous *progranulin-1* and *-2* are not expressed until ~72 hours post-fertilization at which point the 1-cell microinjected MO has likely become inefficient due to dilution (Ekker, 2000; Nasevicius and Ekker, 2000). For this reason only *progranulin-a* and *–b* are amenable to MO-targeted inactivation during zebrafish development and despite the tendency to presume that one or both genes may harbour intrinsically essential biological functions for development it remains to be seen whether these co-orthologues recapitulate an embryonic lethal phenotype in mammals.

1.4.4 Dissection of complex developmental mechanisms using the zebrafish – Developmental biology may be considered the most complex *in vivo* experimental theme within the biological fields. All *in vivo* experimentation must address the tissue-specific

affects of interest as well as the myriad of peripheral responses to the experimental intrusion. However, developmental biology expands upon this complexity due the multipotent nature of the specific and peripheral tissues. The zebrafish has garnered as large degree of respect due to the capability of assessing many cell fates within various contexts including hematopoiesis, craniofacial development as well as visceral and accessory organogenesis.

1.4.5 The zebrafish and comparative hematopoiesis – The human sites of definitive hematopoiesis follow a migratory pattern from the aorta-gonad-mesonephros (AGM) to the fetal liver to the bone marrow (Tavassoli, 1991). However, these sites of blood development are not well conserved in vertebrates (Amatruda and Zon, 1999; Galloway and Zon, 2003). In zebrafish primitive hematopoiesis occurs within the intermediate cell mass (ICM) where nucleated erythroblasts and myeloid cells can be seen circulating at 24hpf (Orkin and Zon, 1997). The ICM is believed to populate the dorsal aorta representing the zebrafish equivalent of the AGM (Amatruda and Zon, 1999). Later, definitive hematopoiesis arises from hematopoietic stem cells (HSC) located within the dorsal aorta and the posterior ICM (Bennett et al., 2001b). Unlike the murine model, definitive hematopoiesis in zebrafish undergoes a migratory transition from the dorsal aorta/ICM to the kidney without the involvement of liver or bone marrow.

Despite these anatomical differences, genes involved in hematopoietic development are well conserved between zebrafish and other vertebrates including the stem cell leukemia (*scl*) factor, *gata-1*, *Pu.1*, *c-myb*, and others (Amatruda and Zon, 1999; Bennett et al., 2001b; Galloway and Zon, 2003; Orkin and Zon, 1997). The zebrafish has become a popular model to study the functional contribution of genes regulating

vertebrate hematopoiesis and cardiovascular development (Amatruda and Zon, 1999). Large-scale mutagenesis screens have produced several models of human hematopoietic diseases including congenital sideroblastic anemia and hemochromatosis type IV (Brownlie et al., 1998; Donovan et al., 2000; Thisse and Zon, 2002). A transgenic zebrafish model of acute T-cell lymphoblastic leukemia has been generated by expression of the mouse c-myc gene under the control of the zebrafish Rag2 promoter (Langenau et al., 2003). Recently, Eckfeldt and colleagues produced a differential display of human progenitor versus committed hematopoietic stem cells (Eckfeldt et al., 2005). Candidates for fate commitment modulation were verified using the zebrafish. Together, these studies demonstrate the utility of the zebrafish model and of the functional conservation of a wide variety of key regulatory proteins across species.

1.4.6 Comparative craniofacial development – The elucidation of the mechanisms and cell types involved in craniofacial development is particularly complex due to the large interplay between numerous tissues and interdependence of growth factor stimulation (Graham et al., 2005; Trainor and Krumlauf, 2001). Derivatives of the neural and non-neural ectoderm, endoderm and mesoderm are all required for proper craniofacial morphogenesis. A large proportion of our understanding of the mechanisms involved in craniofacial patterning have been derived from experiments using mouse, avian and zebrafish models (Creuzet et al., 2005; Trainor, 2005; Yelick and Schilling, 2002). For simplicity and to provide a clear (but brief) understanding of key morphogenetic events, processes will be described in the context of other secreted growth factors including bone morphogenetic proteins (BMP), Wnts and FGFs without an in depth description of the myriad of transcription factors that are also involved in craniofacial morphogenesis.

The neural crest (NC) refers to a multipotent and migratory cell population that give rise to many derivatives including cartilage, bone and part of the peripheral nervous system. Cranial NC cells will form the majority of these tissues within the facial region and originate from the hindbrain at the interface between the neural and non-neural ectoderm of the neural tube. Induction of the NC from this interface involves a number of positive regulators including Wnt8 signalling from within the neural ectoderm and the developing NC (Lewis et al., 2004). BMP signalling (including BMP-2 and -7) is believed to exist as an activity gradient where increased BMP signalling occurs within the non-neural ectoderm whereas these secreted growth factors are limited by antagonists within the neural ectoderm (Schmid et al., 2000). Cumulatively, it is believed that proper NC formation and maintenance is dependent upon this BMP gradient since it has been shown that increased or decreased levels both alter NC formation (Lewis et al., 2004). Based on the rhombomeric (segmental) organization of the hindrain, migrating NC cells form a series of ectomesenchymal arches (EMA) anterior and posterior to the otic placode (ear primordium) (Eisen and Weston, 1993; Raible et al., 1992). The number of EMAs (seven in zebrafish versus five in mice) and the mechanisms of arch segregation and migration vary between species. For instance, the segregation of ectomesenchymal cells into distinct arches in the chick is dependent upon regional apoptosis through the action of BMP-4, whereas the zebrafish and mouse do not rely on cell death for EMA patterning (Kulesa et al., 2004).

The signalling mechanisms between neural and non-neural cells are intrinsic in initial NC specification and initial patterning. Ventral migration of the cranial NC cells from the hindbrain is independent of the pharyngeal endoderm. However, once these cell populations arrive within the dorsal endodermal layer (now referred to as EMAs due to

ongoing differentiation) the maintenance of their EMA derivatives is closely regulated by the ventral pharyngeal endoderm. Indeed, each EMA segmentation is partially dependent upon the invagination of the pharyngeal endoderm to form pouches (David et al., 2002). The lateral migration of the pharyngeal endoderm to form discrete pouches between EMAs 3-7 is dependent upon FGF signalling (Crump et al., 2004; Walshe and Mason, 2003).

1.4.7 Unique and conserved aspects of visceral and accessory organogenesis in zebrafish – The endoderm gives rise to all of the visceral and accessory organs including the esophagus, pharynx, stomach and intestinal tract as well as the liver, pancreas and swim bladder. The rapid development of the zebrafish endoderm from a loose layer of dispersed cells during gastrulation into cohesive and functional organs occurs in five days is particularly impressive. It is this rapid developmental window that has popularized the zebrafish as a model for organ development, particularly the digestive system.

Several cell fate maps have determined that the progenitors for each digestive organ are organized in a linear fashion at 6 hpf, reflecting their mature anterior-posterior framework (Ho, 1992; Ho and Kimmel, 1993; Warga and Nusslein-Volhard, 1999). During gastrulation the endoderm involutes and migrates medially, reaching the midline at ~18 hpf (Warga and Nusslein-Volhard, 1999). Interestingly, the zebrafish endoderm already expresses markers of digestive organs prior to the completion of midline migration, including pancreatic (insulin, somatostatin and pdx-1) and liver (ceruloplasmin) markers (Biemar et al., 2001; Field et al., 2003; Korzh et al., 2001). Furthermore, the endoderm does not acquire any histological organization until 21 hpf, at which point the endoderm begins to acquire a rod like organization along the anterior-

posterior axis (Wallace and Pack, 2003). This rudimentary rod-like structure begins to acquire characteristic epithelial polarity and lumen formation within the stomach and intestine at ~42 hpf (Horne-Badovinac et al., 2001). As previously mentioned the zebrafish expresses genes that are associated with the liver and pancreas prior to completion of midline migration, an observation that has not been described in mice. In addition to this early specification (16 hpf) the endocrine pancreas primordium migrates posteriorly to form an uniform islet medially (24 hpf) which move laterally into the surrounding mesenchyme (~36 hpf) (Kim et al., 2005; Kim et al., 2006). In mice differentiated pancreatic cells form medially (E8.5) and migrate laterally to form an islet (Miralles et al., 1998; Wilson et al., 2003). Furthermore, the development of the mouse endocrine and exocrine pancreas occurs concomitantly and from a single progenitor pool. In the zebrafish the initial development of the exocrine pancreas occurs independently of its counterpart and is initially located in a ventral and anterior position (aspect, 24 hpf), undergoing lateral migration to encompass the endocrine islet by ~36 hpf (Lin et al., 2004b; Wendik et al., 2004; Zecchin et al., 2004).

Similar to the distinctions between the temporal and spatial development of the pancreas in zebrafish and mice, the conservation of gene functions during pancreatic development are variable. Specifically, in mice and zebrafish, *pdx-1* is required for the development of both endocrine and exocrine lineages and *ptf1a* is required only for the exocrine component (Ahlgren et al., 1996; Krapp et al., 1998; Yee et al., 2001; Zecchin et al., 2004). In contrast mouse *ngn3* is required for endocrine development, whereas the zebrafish orthologue is not expressed within this tissue (Gradwohl et al., 2000; Wang et al., 2001).

In contrast the molecular and morphological mechanisms of hepatic differentiation and liver development are generally conserved between zebrafish and mammals, including requirements for FGF and BMP signalling as well as conserved transcription factors such as *hhex* (Keng et al., 2000; Shin et al., 2007; Wallace et al., 2001; Zaret, 2001).

2. Materials and Methods

2.1 Materials – Agarose LE was obtained from Roche Diagnostics (Laval, Quebec, Canada) and DNA ladders (100 bp and 1 kb) were purchased from MBI Fermentas Inc. (Flamborough, Ontario, Canada). *Taq* DNA polymerase was purchased from GE Healthcare (Baie d'Urfé, Quebec, Canada) and KOD-XL DNA polymerase was purchased from VWR Canlab (Mississaga, Ontario, Canada). Agarose gel extraction of PCR products was achieved using a MinElute gel extraction kit (Qiagen, Mississauga, Ontario). TOPO TA cloning kits (pCRII and pcDNA3.1/V5His) were obtained from Invitrogen (Carlsbad, CA). *E.coli* cells (DH5α strain) chemically-competent for transformation were purchased from from Bio S&T (Montreal, Quebec, Canada); electro-competent (Top10 F') cells were obtained from Invitrogen (Carlsbad, CA) or prepared according to the "Preparation of electrocomp^{TM} cells" protocol, version 2.0, from Invitrogen. T7, Sp6 and T3 polymerases, T4 polynucleotide kinase, shrimp intestinal alkaline phosphatase and T4 DNA ligase were purchased from MBI Fermentas Inc (Flamborough, Ontario, Canada). Synthesis of capped mRNA transcripts was performed using the mMessage mMachine T7 Ultra *in vitro* transcription kit from Applied Biosystems Plasmid DNA isolation was achieved using the High Pure Plasmid Isolation kit, the Geno Pure plasmid midi and maxi kits (Roche Diagnostics, Laval, Quebec, Canada). Oligonucleotide primers were obtained from the Sheldon Biotechnology Centre (McGill University, Quebec, Canada) or AlphaDNA (Montreal, Quebec, Canada). Affinity chromatography was performed using Fast-Flow chelating Sepharose from GE Healthcare (Baie d'Urfé, Quebec, Canada). Transient and stable transfects were

preformed using the Fugene 6 transfection reagent from Roche Diagnostics (Laval, Quebec, Canada).

2.2 DNA reactions with modifying and restriction enzymes – All restriction enzymes and modifying enzymes were purchased from GE Healtchare (Baie D'Urfé, Québec, Canada), MBI Fermentas Inc. (Flamborough, Ontario, Canada), or Roche Diagnostics (Laval, Quebec, Canada). Restriction enzyme digests were carried out at 37^{o}C for 2 hours to overnight, in buffers supplied by the manufacturers. DNA modifications, which include ligations and end-labelling reactions, were performed as outlined in Molecular Cloning, a Laboratory Manual (Sambrook et al., 1989), or according to the manufacturer's instructions.

2.3 DNA sequencing – All sequencing was performed at the Sheldon Biotechnology Centre of McGill University, using a MegaBACE 500 capillary electrophoresis sequencer (GE Healthcare) using primer sequencing or cycle sequencing methods, respectively.

2.4 Determination of *progranulin* genomic architectures – Listings of the various taxonomies investigated for *progranulin* gene architecture. Specific genome builds used for each analysis are indicated along with the reference marker associated with each gene for the build in question.

Common Name	Taxonomy	Genome Build Version	Information Database Reference	Reference
Thale cress	*Arabidopsis thaliana*	TAIR7	TAIR:AT1G09850 (XBCP3)	
Human	*Homo sapiens*	NCBI 36	ENSG00000030582	(Lander et al., 2001; Venter et al., 2001)

Chimpanzee	*Pan troglodytes*	PanTro2.1	ENSPTRG00000009273	(2005)
Rhesus Monkey	*Macaca mulatta*	Mmul 1.0	ENSMMUG00000023751	(Gibbs et al., 2007)
Mouse, strain C57BL/6J	*Mus musculus*	m36	ENSMUSG00000034708	(Waterston et al., 2002)
Rabbit, european	*Oryctolagus cuniculus*	Low-coverage 2X	ENSOCUG00000013567	N/A
Frog,	*Xenopus tropicalis*	JGI 4.1		N/A
Zebrafish	*Danio rerio*	Zv6	Progranulin-a: ENSDARG00000004954	(Woods et al., 2000)
Zebrafish	*Danio rerio*	Zv6	Progranulin-b: ENSDARG00000025081	(Woods et al., 2000)
Zebrafish	*Danio rerio*	Zv6	Progranulin-1: ENSDARG00000035818	(Woods et al., 2000)
Zebrafish	*Danio rerio*	Zv6	Progranulin-2: ENSDARG00000035817	(Woods et al., 2000)
Opposum	*Monodelphis domestica*	monDom 5.0	ENSMODG00000011543	(Samollow et al., 2004)
Dog, boxer breed	*Canis familiaris*	CanFam 2.0	ENSCAFG00000014165	(Lindblad-Toh et al., 2005)
C. Elegans	*Caenorhabditis elegans*	WS170	T22H2.6	(1998)
C. Elegans	*Caenorhabditis elegans*	WS170	T02B11.8	(1998)
Common Name	**Taxonomy**	**Genome Build Version**	**Information Database Reference**	**Reference**
Slime mold, strain AX4	*Dictyostelium discoideum*	DictyBase V2	Entrez Gene ID: 3390519	(Chisholm et al., 2006; Eichinger et al., 2005)
Sea squirt	*Ciona intestinalis*	JGI 2.0	ENSCING00000012906	(Dehal et al., 2002)
Sea Urchin	*Strongylocentrotus purpuratus*	Spur V2.1	Entrez Gene ID: 575170 (LOC575170)	(Sodergren et al., 2006)

2.5 RNA extraction – Total RNA was isolated from approximately 50 embryos staged according to Kimmel (Kimmel et al., 1995) after freezing in liquid nitrogen and subsequent homogenization by passing through a 28G1/2 gauge syringe (Becton Dickson, Oakville, Ontario, Canada) using Trizol from Invitrogen (Carlsbad, CA), according to manufacturer's instructions. Subsequently, an equal volume of water-saturated phenol was added to the homogenate and incubated at 65^{o}C for 10 min. The aqueous phase was re-extracted once with phenol and twice with chloroform/isoamyl-alcohol (24/1). Total

RNA was precipitated by adding one-quarter volume of 10M LiCl at –20^{o}C for 1 hour and collected by centrifugation. The pellet was washed with 75% ethanol, dried, resuspended in DEPC-treated H_2O and stored at –80^{o}C. Amounts of total RNA were assessed by A_{260} measurements and visualization of rRNA bands through TAE-agarose gel electrophoresis next to an RNA ladder (Gibco BRL).

For extraction of total RNA from individual embryos (specifically those injected with the splice-site morpholino and associated controls), single embryos staged according to Kimmel (Kimmel et al., 1995) were frozen in liquid nitrogen and the total RNA extracted using the Illustra RNAspin Mini kit (GE Healthcare), according to manufacturer's instructions.

2.6 RT-PCR – Prior to cDNA synthesis, total RNA was treated with DNase I according to the manufacturer's instructions (MBI Fermentas Inc. (Flamborough, Ontario, Canada). Oligo dT-primed template cDNA for RT-PCR analysis was synthesized from the extraction of individual zebrafish embryos at 24 hours post-fertilization using the RevertAid H Minus First Strand Synthesis Kit (K-1632, MBI Fermentas Inc., Flamborough, Ontario, Canada). To control for genomic DNA contamination, the samples were also incubated in the absence of reverse transcriptase. The resulting cDNA reaction mixture was resuspended in a final volume of 20μl in dH_20, from which 2 μl aliquots were used in each subsequent amplification reaction. Primer sets used to detect an intron inclusion or exon excision event due to splice-morpholino injection along with actin controls are as follows:

Actin (forward) 5'-ATGGATGATGAAATTGCCGC-3'

Actin (reverse) 5'-TGTCATCTTTTCCCTGTTGG-3'

Progranulin-a (forward) 5'-GTTGAGACTGACAGTCTGCCT-3'

Progranulin-a (reverse1) 5'-GACAAGAGAACTCAGCAGGAC-3'

Individual polymerase chain reactions were performed as follows on a ThermoHyBaid (model HBSPO2110) thermal cycler (Ashford, UK) using Taq (Amersham Biosciences): initial denaturing at 94^{o}C for 2 min, followed by 35 cycles of 94^{o}C for 45 sec, annealing at 60^{o}C for 1 min and extension at 72^{o}C for 1 min. A final extension step at 72^{o}C was carried for 7 min. The authenticity of the amplicons was confirmed by extraction using a MinElute Gel extraction kit (Qiagen) and TOPO pCRII cloning (Invitrogen) -according to manufacturer's instructions- followed by sequencing of both strands with T7 and Sp6 priming sites (Sheldon Biotechnology Centre).

2.7 Fish maintenance - Wild type zebrafish were purchased from Scientific Hatcheries (San Diego, CA) and maintained on a 14h/10h light/dark cycle at 28.5^{o}C in an Allentown Aquaneering tank system (New Jersey), fed twice daily, and bred as described elsewhere (Mullins et al., 1994). Embryos for developmental studies were collected from tanks and staged according to Kimmel *et al*. (Kimmel et al., 1995).

2.8 Preparation of constructs for riboprobe synthesis - For the following riboprobe synthesis reactions, RNA polymerases (T3, T7 and Sp6), ribonuclease inhibitor and DNase1 were purchased from MBI Fermentas Inc (Flamborough, Ontario, Canada). Digoxigenin-11-UTP and the anti-digoxigenin-AP (alkaline phosphatase) Fab fragments, for probe preparation and detection, respectively, were purchased from Roche Diagnotics

(Laval, Quebec, Canada). The NBT/BCIP chromagen mix was used for alkaline phosphatase staining. NBT (nitroblue tetrazolium) and BCIP (bromochloroindolyl phosphate) were purchased from MBI Fermentas Inc.

Individual insert fragments were cloned into the TOPO pCRII vector (Invitrogen, Carlsbad, CA) according to manufacturer's instructions: For *progranulin*-a, two clones corresponding to the 5' (1302 base pairs) and 3' (1298 base pairs) ends of the full-length mRNA. For both progranulin-a 5' and 3' plasmids, probe synthesis was carried out using the Sp6 polymerase with NotI-linearized template (antisense) and the T7 polymerase with HindIII-linearized template (sense). As a control for riboprobe synthesis and subsequent *in situ* hybridization a ~2.5kb fragment of zebrafish sonic hedgehog (*Shh*) subcloned into pbluescript (gift from the Akimenko lab, University of Ottawa) was used routinely (Krauss et al., 1993). *Shh* riboprobes were transcribed using T7 polymerase with EcoRI-linearized template (antisense) and T3 polymerase with XhoI-linearized template (sense). The remaining riboprobes were synthesised from fragments cloned in house from cDNA or kindly provided by members of the zebrafish research community.

2.9 Whole-mount *in situ* hybridization - In order to prevent the appearance of melanin pigmentation, embryos at approximately 18-24hpf were grown in egg water supplemented with 0.003% of the tyrosinase inhibitor 1-phenyl-2-thiourea (phenythiocarbamide P-5272, Sigma). Staged embryos were manually dechorionated and fixed for 2 hours at room temperature or overnight at 4^{o}C in 4% paraformaldehyde/PBS. After several washes in PBS, embryos were stored in 100% methanol until needed. Whole-mount in situ hybridization with digoxygenin-labeled RNA probes and antibody staining were essentially done according to the Schulte-Merker *et al*. (Schulte-Merker et

al., 1992) or Thisse *et al*. (Thisse et al., 2001) at a hybridization temperature of 70^{o}C. An extended description of the procedure is appended at the end of this section. In some cases, polyvinyl alcohol was added to the staining solution in order to minimize the occurrence of background, especially when the reaction was required to proceed for several days (De Block and Debrouwer, 1993). The description of anatomical orientations for zebrafish embryos follows suggested conventions (Moorman, 2001). Stained whole-mount and sectioned embryos were mounted in glycerol and visualized under a Leica MZFLIII stereomicroscope. Pictures were taken with a Leica DC350F camera and processed with Adobe Photoshop 7.0 software within the acceptable digital augmentation guidelines described in the Journal of Cell Biology and adopted by The Journal of Biological Chemistry (Rossner and Yamada, 2004).

2.10 Whole-mount immunohistochemistry and immunoflourescence – The immunodectection of antigens within zebrafish whole-mounts was carried out according to Pack et al., (Pack et al., 1996) and Fashena et la., (Fashena and Westerfield, 1999) with modifications. Briefly, staged embryos were fixed for 2 hours in 4% PFA/PBS at room temperature, washed in PBS, manually dechorionated if necessary, and then dehydrated in methanol. Rehydration of embryos was performed for 5 min each in successive solutions of MeOH / PBST (25%/75%, 50%/50%, 75%/25%), then 3 times in PBST. Embryos were then permeabilized with proteinase K diluted in PBST at a final concentration of 10 μg/ml. Duration of treatment increased with age as follows: 0 min for 6hpf embryos, 25 min for 48hpf embryos, 30 min for 72hpf embryos. Post fixation was then carried in 4% PFA/PBS for 20 min at room temperature, and embryos were rinsed 3 times in PBST.

Preincubation in blocking solution was carried in PBST / 5% Goat or Calf serum / 1% DMSO at room temperature for 2-5 hours. Primary antibody was then added at the appropriate dilution and incubated at 4ºC overnight:

anti-Carboxypeptidase-A; 1/1000 (Pack et al., 1996)

anti-Zn8; 1/500 (Fashena and Westerfield, 1999)

The following day, embryos were rinsed 6 times in PBST for 15 min. Detection of the primary antisbody was performed using anti-mouse or anti-rabbit conjugated alkaline phosphatase from Calbiochem (San Diego, CA) at a dilution 1/1000 for two hours in blocking solution followed by several PBST washes. in the ABC Staining System (sc-2018, Santa Cruz Biotechnology, Santa Cruz, CA) according to the manufacturer's instructions. Briefly, embryos were incubated with pre-adsorbed secondary antibody for 2-3h at a dilution of 1/200, then rinsed 6 times in PBST. Staining was performed with diaminobenzidine (DAB) reacting with the horseradish peroxidase-conjugated secondary antibody. Alternatively, immunodetection was carried with NBT/BCIP using a goat anti-rabbit IgG secondary antibody conjugated to alkaline phosphatase (#A-3937, Sigma, St-Louis, Missouri, USA).

2.11 Histochemical staining – Jaw and branchial cartilage staining with Alcian Blue 8GX (Sigma) was carried out using an adaptation of published protocols (Schilling et al., 1996; Schlombs et al., 2003). Embryos raised to 120 hpf were fixed using 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS) for 2 hours at room temperature and then stored in 100% methanol at –20°C. Embryos were stained overnight at 4°C in 0.08% alcian blue / 20% glacial acetic acid / 80% ethanol. After

several washes and rehydration in water embryos were bleached in 0.8% KOH / 0.9%H_2O_2 / 0.2% Triton X-100 for 10 minutes followed by neutralization in several washes of water-saturated sodium tetraborate. Tissues were softened in 1 mg/ml trypsin / 60% water-saturated sodium tetraborate / 0.2% Triton X-100 for 1 hour at room temperature and cleared of background staining in 0.8% KOH / 18% glycerol / 0.2% Triton X-100 for 30 minutes. Whole-mounts and dissected architectures were mounted in 100% glycerol and photographed using a Leica MZ FLIII stereomicrosope.

For histochemical staining of hemoglobin, embryos were placed in freshly prepared *o*-dianisidine stain solution (40% ethanol with 0.01 M sodium acetate, 0.65% H_2O_2, and 0.6 mg/ml *o*-dianisidine (D-9143; sigma) for 15 minutes and then were washed in water. For microangiography individual zebrafish embryos (3 days post-fertilization) were anesthetized in aquarium water containing 90 mg/L 3-aminobenzoic acid ethyl ester (A-5040, Sigma) and placed on injection platforms and overlayed with phosphate buffered saline containing 0.3% methyl cellulose. A ~1uL solution of 1X Danieau buffer (58mM NaCl / 0.7 mM KCl / 0.4mM $MgSO_4$ / 0.6mM $Ca(NO_3)_2$ / 5.0mM Hepes, pH 7.6) containing 0.1% FITC-labelled dextran (Sigma) was microinjected into the sinus venousus located ventral to the heart proper. Individual embryos removed after 5min and immediately placed into 4% PFA/PBS for 2hours in the dark. Whole-mount fluorescence was visualized and photographed using a Leica MZ FLIII stereomicrosope equipped for appropriate GFP filters.

.

2.12 Whole-mount TUNEL apoptosis assay – Embryos were initially prepared as described for whole-mount in situ hybridization up until the point of riboprobe addition. TUNEL staining was carried using the *In Situ* cell death detection kit from Roche

Diagnostics (Laval, Quebec, Canada), according to manufacturer's instructions with minor modifications. Briefly, embryos were preincubated once for 5 minutes and once for 30 minutes in Labeling buffer, followed by incubation in a 50:1 ratio of Labeling mix to TdT enzyme for 1 hour at 37°C. The reaction was stopped using several washes of PBST with 5mM EDTA. For negative and positive controls embryos were incubated without TdT enzyme or previously treated with DNaseI (Fermentas) for 30 minutes at 37°C, respectively. At this point embryos were processed for subsequent immunohistochemistry or in situ hybridization as outlined above, followed by colorimetric visualization of TUNEL positive cells using an anti-fluorescein alkaline phosphatase conjugate and Fast Red substrate (Roche Diagnostics).

2.13 Production of progranulin-A and progranulin-B polyclonal antisera – Rabbit polyclonal antisera were generated against synthetic peptides corresponding to residues 242 to 256 of zebrafish progranulin-A (RAEWEDHKQKKPETQC), and residues 153 to 173 of zebrafish progranulin-B (CGSSPFLRKFAARRRKPLEKNA), respectively. The cysteine residues were introduced into the sequences and used for thiol coupling to keyhole limpet hemocyanin (KLH). Peptide synthesis and production of antibodies were generated through the Sheldon Biotechnology Centre of McGill University. Polyclonal antiserum was purified using the Montage Antibody Purification PROSEP-A Kit (P36486, Millipore, Nepean, Canada). IgG was quantified according to published values where an O.D. of 1.32 at 280nm equals 1.0 mg/ml of IgG (Leslie and Clem, 1969). After desalting and concentrating, the antibodies were resuspended in 1% BSA/PBS. After staining, embryos were rinsed 3 times in PBST, postfixed in 4% PFA/PBS, and stored a

4ºC in PBST. Alternatively, embryos were transferred in increasing concentrations of glycerol as follows: 25%, 50%, 75%, 100% glycerol for 30 minutes each, rinsed in glycerol twice and store at 4ºC. The glycerol treatment renders the embryos with enhanced optical clarity.

Preadsorption of antibodies was performed on the first day. 50-1000 spare embryos (usually 24-48hpf embryos) were quickly rehydrated in PBST, and rinsed 3 times in PBST. Incubation of the antibody in the blocking solution (PBST / 2% BSA / 5% Calf serum) containing the spare embryos was carried out at room temperature for several hours and then stored at 4ºC until used. Preadsorption was done with the primary antibody at a concentration 5-10 times higher than the final concentration needed for the incubation step during detection. Antibodies were diluted at working concentrations prior to use (1/1000 for anti-progranulin-A; 1/1000 for anti-progranulin-B; 1/400 for anti-progranulin-1; 1:200 for anti-acetylated alpha-tubulin used as positive control (mouse monoclonal; T-6793, Sigma).

2.14 SDS-PAGE and Western Blot – At 24hpf embryos were manually de-yolked and frozen in liquid nitrogen then stored at –80C. Samples were thawed and 30µL of 2X Laemmli buffer (Laemmli, 1970) was added and extracts boiled for 5minutes. Embryo extracts were resolved as 2 embryo equivalents per lane on 15% acrylamide gels and transferred to nitrocellulose membrane (GE healthcare, Baie-d'Urfe, QC, Canada). The blots were incubated in PBST with 1:4000 anti-progranulinA for 1hour followed by extensive washing. After incubating with an anti-rabbit or -mouse IgG-horseradish peroxidase (HRP)-conjugated secondary antibody (diluted 1:4,000) at room temperature for 1 hour and blots were visualized using enhanced chemiluminescence (GE Healthcare)

according to the manufacturer's instructions. In addition, mouse monoclonal β-actin antibody (AC-40; Sigma) at a dilution of 1:1000 was used to determine total protein content per lane. The rabbit polyclonal anti-progranulinA and anti-progranulin-B were raised against synthetic peptides corresponding to residues 242 to 256 and 153 to 173, respectively (Sheldon Biotechnology Centre). Epitope-specific immunoglobulins were purified by affinity chromatography. Specificity of Western blot immunoreactive bands were confirmed to be progranulin-A or –B using synthetic peptide blocking incubations.

15 morpholino-injected or wild-type embryos were collected at the 72hpf stage, frozen in liquid nitrogen, and stored at -80^{o}C until needed. Embryos were homogenized in 30 µl 2 x Laemmli buffer (10X Laemmli is 0.25 M Tris, 1.92 M Glycine and 1 % SDS in aqueous solution) by gentle pipetting according to a published protocol. To this, 1/100 volume of reducing agent mercaptoethanol was added. A BroadRange high molecular weight (M.W.) protein standard (Bio-Rad, Hercules, CA, USA) used as migrating markers was similarly prepared using 5µl of M.W. marker stock. The samples were boiled for 5 min at 95^{o}C to facilitate protein denaturation, chilled on ice for 30 seconds and centrifuged at 13 000 rpm for 10 min at 4^{o}C. The samples were then loaded and resolved through migration on a 8% acrylamide gel in running buffer (5x stock running buffer is 1.51% Tris-base, 7.21% Glycine, 0.5% SDS in ddH_2O) for 90 min at 150 volts. Electrophoresis was monitored using pre-stained M.W. markers (Bio-Rad).

Control immunodetection was performed using an anti-mouse actin monoclonal antibody (AC-40, Sigma) as primary antiserum which detects the zebrafish homologue, diluted at 1:5000. For the detection of progranulin-A, the anti-progranulin-A antibody was diluted at 1:10000. Detection was achieved using an anti-rabbit IgG antibody

conjugated to HRP as secondary antibody (Amersham Biosciences), using the ECL detection kit (Amersham Biosciences). Membranes were exposed to Kodak X-OMAT film for signal development, and images were analysed using Adobe PhotoShop 7.0 software after scanning.

2.15 Morpholino microinjection – We microinjected 1- and 2-cell stage embryos essentially as described (Nasevicius and Ekker, 2000). Morpholine-based oligonucleotides were designed according to published criteria (Summerton and Weller, 1997) and purchased from GeneTools, LLC (Philomath, Oregon, USA). Negative controls include individual morpholinos incorporating 4 or 5 nucleotides mis-matches, and a standard morpholino of scrambled sequence. A morpholino complementary to the zebrafish chordin transcript was used as positive control to monitor individual morpholino dose and efficacy of injections. Morpholinos (MO) oligonucleotides were diluted in 1X Danieau buffer (58mM NaCl / 0.7 mM KCl / 0.4mM $MgSO_4$ / 0.6mM $Ca(NO_3)_2$ / 5.0mM Hepes, pH 7.6) with 0.1% Phenol Red (as tracer) (Nasevicius and Ekker, 2000). Morpholinos designed to target the 5'UTR, ATG start codon of progranulin-A or the exon4/intron4 boundary (Splice-MO) and controls (See Figure 10) are as follows:

MO1-UTR, 5'-GATCCTTAGGGTCCAAGTTACCGAT-3'

MO2-UTR, 5'-GAGCAGGTGGATTTGTGAACAGCGG-3'

MO3-ATG, 5'-GAGGCAGACTGTCAGTCTCAACATT-3'

MO2-5b.p. mismatch, 5'-GAACACGTGGATTTCTGAAGAGAGG-3'

Splice-MO 5'- TATAATGTCTCACTTTGGGAAGGTC-3'

Scramble, 5'-CCTCTTACCTCAGTTACAATTTATA-3'.

The splice-MO was a gift from Dr. Peter Hitchcock (University of Michigan, Michigan) MOs were injected into the yolk cell adjacent to the blastomeres of 1- or 2-cell stage embryos using a PLI-100 pico-injector (Harvard Apparatus, Holliston, Massachusetts,USA). Typically, 1 to 5 nl of diluted MO were injected per embryo for a final amount of up to a maximum amount of 10 ng of MO. Concentrations of 15 ng/embryo (7.5ng/nl) was considered the upper limit as the scramble control produces a consistent minority of non-specific developmental defects at this concentration. For asymmetric morpholino injections a concentration of 10 ng/embryo of MO2-UTR (referred to as MO in text) along with 0.05% FITC-dextran (Sigma-Aldrich, Oakville, ON, Canada) was injected into a single marginal blastomere at the 16-cell stage (Peyrieras et al., 1998). Asymmetric morpholino deposition was confirmed by fluorescence analysis during gastrulation using a Leica MZ FLIII stereomicrosope (Leica Microsytems, Richmond Hill, ON, Canada).

2.16 Cell cultures – The TN20 insect cell line (kindly provided by Dr. Luis F. Congote; University of McGill, Quebec) was maintained in EX-CELL 420 insect serum-free medium (JRH Biosciences, Lanex, KS) containing 10ug/ml gentamaycin (Invitrogen) in adherent culture flasks (Starstedt) at 28.5°C. TN20 and stable transfected cells were passaged from 80% to 20% confluency every five days. The Cos-7 cell line (kindly provided by Dr. Andrew Bateman; McGill University, Quebec) was maintained in DMEM (Invitrogen) containing 10ug/ml gentamaycin (Invitrogen) and 10% serum (Invitrogen) in adherent culture flasks (Starstedt) at 37°C with 5% CO_2. Cos-7 cells were passaged from 80% to 20% confluency every five days.

2.17 Stable transfection of TN20 insect cells – Full-length progranulin-A was subcloned into pIB-V5/His TOPO vector (Invitrogen, Carlsbad, CA) to produce a carboxyl-terminally tagged recombinant product containing the V5 epitope and 6xHistidine tags. The pIB-progranulin-A/V5/His vector was transfected into 50% confluent TN20 insect cells using 5ug of DNA and the Fugene 6 (Roche Diagnotics, Laval, QC) transfection reagent at a ratio of 3:1 (reagent: DNA) according to manufacturer's instructions. After 48hours fresh medium was added along with 80ug/ml Blasticidin (Invivogen). Stable transfectants were selected over a three week period using 80ug/ml Blasticidin after which the concentration was decreased and maintained at 10ug/ml. Secretion of full length zebrafish progranulin/V5/His was confirmed by Western blot (see Section 2.12 for details) using an antigen purified anti-progranulin-A polyclonal antibody (1:5000) and an ECL chemiluminescence kit (GE Healthcare) following manufacturer's instructions. goat anti-rabbit Horseradish peroxidase secondary (GE Healthcare) purified using nickel chelating fast flow Sepharose (GE-Healthcare) and buffer exchanged into sterile PBS using Centricon Plus-20 columns (Millipore). Purity was assessed by silver stained SDS-PAGE (Bio-Rad), and N-terminal sequencing (Sheldon Biotechnology Centre).

2.18 Expression vector construction – The full-length *progranulin-a* sequence was purchased from RZPD (Berlin, Germany) as clone UCDMp574E2318Q2. Full-length *progranulin-a* was subcloned using a forward that overlapped the starter AUG a reverse primer that read through the termination codon:

AUG (forward) 5'-AAAATGTTGAGACTGACAGTCTGC-3'

Read-through (reverse) 5'-TAGAGTTAGGGCTCGTTTC-3'

Polymerase chain reaction performed as follows on a ThermoHyBaid (model HBSPO2110) thermal cycler (Ashford, UK) using *Pfu* polymerase (Fermantas): initial denaturing at 94^{o}C for 2 min, followed by 28 cycles of 94^{o}C for 1min, annealing at 58^{o}C for 1 min, 30sec and extension at 72^{o}C for 6 min. An additional two rounds of PCR were formed following the addition of *Taq* polymerase (GE Healthcare) for 3' adenine addition (as required for TOPO cloning), with a final extension step at 72^{o}C for 15 min. The PCR product was confirmed as full-length by DNA eletrophoresis, extracted using the MinElute Gel Extraction kit (Qiagen) and sublconed into pIB-V5/His TOPO and pcDNA3.1-V5/His vector (Invitrogen, Carlsbad, CA), according to manufacturer's instructions. The final vector constructs consisted of ful-length progranulin-A with a carboxyl-terminal tag consisting of the V5 epitope and 6xHistidine.

2.19 Purification of zebrafish progranulin-A/V5/His – A total of three T_{125} adherent cultures flasks were used to grow stably transfected TN20 cells from 20% to 80% confluence over five days using 20ml of serum free medium containing 10ug/ml gentamaycin (Invitrogen) and 10ug/ml Blasticidin (Invivogen) per flask. After five days of growth the medium was removed, combined and centrifuged at 2,000rpm for 15min to remove suspended cells. The total volume was diluted 3:1 (vol:vol) with running buffer (10mM HEPES, 500mM NaCl, pH 7.8) and applied to a previously equilibrated Nickel sepharose column (GE Healthcare; according to manufacturer's instructions) with a bed volume of ~6mL at a flow rate of ~2ml/min. The column was washed with three column volumes (18ml) of running buffer followed by three column volumes (18ml) of wash buffer (10mM HEPES, 500mM NaCl, 40mM Imidazole, pH 7.8). Bound proteins were

eluted with three column volumes of elution buffer (10mM HEPES, 500mM NaCl, 500mM Imidazole, pH 7.8) directly into a Centricon PL-20 (Centricon) with a 30,000 nominal molecular weight cut off and centrifuged at 2,000rpm at 4°C for 20min. The concentrate was diluted and re-centrifuged under the same conditions twice using sterile phosphate buffered saline pH 7.2 until a volume of ~200µL was achieved. To assess for purity, confirm identity and presence of full-length protein, four aliquots of 5% of the total volume (10µL) were used for (A) Western blot using the antigen purified anti-progranulin-A polyclonal antibody (1:5000), (B) Western blot using an anti-V5 polyclonal antibody (Invitrogen), (C) SDS-PAGE and protein transfer followed by coomoassie blue staining, excision and N-terminal sequencing (Sheldon Biotechnology Centre) and (D) SDS-PAGE and silver staining followed by proteolytic digestion and mass spectrometric identification (Genome Quebec). Purified progranulin-A/V5/His was quantified by amino acid analysis (Dr. Claude Lazure, Institute des recherches du Montreal).

2.20 Surface plasmon resonance; ligand immobilization – For SPR analysis, the zebrafish progranulin-A/V5/His ligand (15 mg/mL in 10 mM sodium acetate pH 4.0) was amine-coupled to BIACORE CM4 sensor chips according to the manufacturer's recommendations to produce an active surface of ~2500 basal response units. Confirmation of binding and structural integrity was assessed by following the kinetics of anti-progranulin-A and the newly prepared progranulin-A/V5/His surface (data not shown).

2.21 Surface plasmon resonance; protein-protein interaction screening – Protein-protein interactions were monitored using BIACORE 2000 and 3000 optical biosensors set at 25°C with an active progranulin-A/V5/His surface of 2500-3000 response units. All interactions were conducted at a flow rate of 10μL/min for 2 minutes. Between sample injections, regeneration was achieved by two 5 μL pulses of solution A (10mM glycine pH 2.0 containing 0.5M NaCl), followed by two 5 μL pulses of solution B (10 mM HEPES pH 7.4, 150mM NaCl, 3.4mM EDTA, 0.005% Surfactant P20). All fractions assessed for progranulin interactions were conducted in duplicate. SPR data was prepared according to the double referencing method (Myszka, 1997; Myszka, 2000), however large non-specific interactions were problematic. To visually assess the validity of an interacting species the kinetics of the interaction were monitored positive slope values during the sample application, suggesting an additive interaction that occurred independently of the non-specific saturation effect.

Following each round of reversed phase high performance liquid chromatography (RP-HPLC), aliquots of individual fractions (120μL following primary and secondary fractionations; 400μL following tertiary fractionation) were lyophilized and resuspended in 120μL of running buffer (10 mM HEPES pH 7.4, 150mM NaCl, 3.4mM EDTA, 0.005% Surfactant P20) in the presence or absence of 5 mM mannose, glucose, or galactose). For data analysis a merge of chromatographic data with SPR signal results was produced by plotting absorbance at 214nm and the ratio of SPR progranulin-A/V5/His interactive binding to non-specific association to the chip surface over time for those fractions whose progranulin binding kinetics demonstrated a positive slope.

2.22 Identification of progranulin binding proteins by mass spectrometry – Following three rounds of RP-HPLC and subsequent SPR analysis those fraction that demonstrated a positive interaction as well as two outlying negative control fractions were analyzed for protein content by mass spectrometry. Individual fractions were reduced, sulfur-alkylated, and digested with trypsin using a robotic digester (Micromass) at the Protein Unit Facility of the Genome Quebec Proteomics Platform. The resultant peptide mixtures were analyzed by direct on-line liquid chromatography tandem mass spectrometry (LC-MS/MS) using a quadrupole time-of flight (QToF) mass spectrometer (Waters). The MS/MS spectra were evaluated using Mascot software (Matrix Science) to identify tryptic peptide sequences matched to the National Center for Biotechnology Information (NCBI) nonredundant protein and nucleotide data bases (dbEST) with a confidence level of 95% or greater (Mann and Wilm, 1994; Perkins et al., 1999). False positives were filtered by negative control fractions and by mining the dbEST database of human (*Homo sapiens*) rather than the zebrafish (*Danio rerio*) database.

2.23 Zebrafish embryo protein extraction and fractionation – Approximately 8,000 live 28-30 hours post-fertilization embryos were snap frozen in an acetone/dry ice bath courtesy of Scientific Hatcheries (Huntington Beach, CA) and stored at -80°C. Embryos were thawed and homogenized in 80ml of acidic extraction buffer (1M HCl, 5% formic acid, 1% (w/v) NaCl and 1% TFA). The homogenate was centrifuged at 2,400rpm at 4°C for 20min and the supernatant subjected to reversed-phase enrichment using C_{18} SepPak cartridges (Waters). Cartridges were washed with 4 column volumes of buffer A (0.1%TFA) and proteins eluted using 2 column volumes of buffer B (0.1% TFA,

80%ACN) followed by immediate lyophilization to 20% of initial volume to remove all organic solvent and applied immediately to a pre-equilibrated reversed-phase C18 μBondaPak column at 1ml/min and washed in buffer A for 20minutes at 1ml/min. The separation of enriched proteins was achieved using reversed-phase high performance liquid chromatography (RP-HPLC) using a Waters Breeze HPLC system and a C18 μBondaPak column. Samples were separated using a linear gradient of increasing acetonitrile (ACN) at a constant flow-rate of 1.5ml/min and collected as 1min fractions. Proteins were purified using three rounds of RP-PHLC and subsequent SPR analysis.

Gradient conditions for each RP-HPLC fractionation were 0.1% (v/v) Trifluoroacetic acid (TFA) with a 10-80% acetonitrile gradient over 50min (Primary separation), 0.13% (v/v) heptafluorobutyric acid with a 10-80% acetonitrile gradient over 50min (Secondary separation) and 0.1%TFA with a 25-65% acetonitrile gradient over 50min (Tertiary separation).

2.24 Knock-down phenotype analyses – To monitor the phenotypic consequences of zebrafish progranulin-a knock-down, several marker gene expression and histochemical techniques are used at appropriate developmental timepoints. Cell proliferation is monitored using the mouse monoclonal anti-phospho-histone H3 antibody (06-570; Upstate, Charlottesville, VA) and proliferating cell nuclear antigen (PCNA) mRNA expression. Apoptosis is monitored using the *In Situ* Cell Death Detection Kit (Roche Diagnostics). The following riboprobes have been constructed by amplifying a segment of the full-length mRNA using Taq DNA polymerase, and cloned into pCR2.1 (Invitrogen). After sequencing of individual amplicons for confirmation of their

authenticity and verification of their orientation, corresponding antisense riboprobes were synthesized using conventional methods (see above) using either the Sp6 or T7 RNA polymerases. A PCNA riboprobe (1089 nucleotides) corresponding to nucleotides 21-1110 of the PCNA mRNA (accession BC049535); a flk-1 riboprobe (1209 nucleotides) corresponding to nucleotides 17-1226 of the flk-1 mRNA (accession AF487829); an angiopoietin receptor (Tie-2) riboprobe (1072 nucleotides) corresponding to 2471-3543 of the Tie2 mRNA (accession AF053632); a GATA-2 riboprobe (1159 nucleotides) corresponding to nucleotides 3613-4772 of the GATA-2 mRNA (accession AF001220); an ikaros riboprobe (1109 nucleotides) corresponds to nucleotides 1171-2280 of the ikaros mRNA (accession AF092175); a rag1 riboprobe (1062 corresponds) corresponding to nucleotides 71-1133 of the rag1 mRNA (accession U71093); a sox17 riboprobe (1102 nucleotides) corresponding to nucleotides 11-1113 the sox17 mRNA (accession AF168614); a preproinsulin riboprobe (437 nucleotides) corresponding to nucleotides 1-437 of the preproinsulin mRNA (accession AF036326); a pdx-1 riboprobe (914 nucleotides) corresponding to nucleotides 1-914 of the pdx-1 mRNA (accession AF036325); a trypsin riboprobe (767 nucleotides) corresponding to nucleotides 34-801 of the trypsin precursor mRNA (AF541952); a E-cadherin (cdh1) riboprobe (1093 nucleotides) corresponding to nucleotides 2111-3204 of the cdh1 mRNA (accession AF364811); a N-cadherin (cdh2) riboprobe (1157 nucleotides) corresponds to nucleotides 18-1175 of the cdh2 mRNA (accession AF418565); a neural adhesion molecule L1.1 riboprobe (1240 nucleotides) corresponding to nucleotides 241-1481 of the L1.1 mRNA (accession AY376855); a neural adhesion molecule L1.2 riboprobe (1258 nucleotides) corresponding to nucleotides 86-1344 of the L1.2 mRNA (accession AY376856); a zash1

riboprobe (1855 nucleotides) corresponding to nucleotides 83-1938 of the zash1 (neurod) mRNA (accession NM130978).

We are indebted to several colleagues for their generous riboprobe construct gifts: Jennifer Rhodes (the A. T. Look lab, Harvard University, Boston) for cebpα, Pu.1, mpo, l-plastin, scl, gata-1, alpha globin, and flk-1; Didier YR Stainier (UCSF, San Francisco) for fkd7 (foxA1) and fkd2 (foxA3); Hitoshi Okamoto and Hiroshi segawa (RIKEN Brain Science Institute, Japan) for islet-1 and islet-2; Michael Pack (University of Pennsylvania, Philadelphia) for pdx-1 and hhex; Peter Leach (Johns Hopkins Medical Institutions, Baltimore) for ptf1a/p48. Zhiyuan Gong(National University of Singapore, Singapore) for elastaseB and transferrin. Alain Ghysen (Université Montpellier II, France) for the lateral liner markers CB403 and CB701. All other markers were purchased from RZPD (Germany).

3. Results

3.1 Progranulin gene evolution - The illustration of the architectures for *progranulin* genes from the human, mouse and zebrafish *progranulin-1* and *-2* were inferred from published information (Baba et al., 1993; Bhandari and Bateman, 1992; Bhandari et al., 1992; Cadieux et al., 2005). Architecture of the thale cress cysteine peptidase gene *XBCP3* that contains a granulin domain was derived The Arabidopsis Information Resource (TAIR) database (Huala et al., 2001). Identification of the sea urchin *progranulin* gene was partially derived from the Sea Urchin Genome Project located at the Baylor College of Medicine using the GNOMON gene prediction method developed by the National Center for Biotechnology Information (NCBI) (Sodergren et al., 2006). Manual translation of the genomic sequence upstream of LOC575170 uncovered the intronless repeat of twenty-five granulin domains. Confirmation of these architectures (when possible) and the evaluation of other *progranulin* genomic architectures for all other species were derived from the Ensembl Genome Browser (Curwen et al., 2004; Hubbard et al., 2002; Stalker et al., 2004). See section 2.4 for details. For those genes which did not demonstrate a conserved *progranulin* genomic architecture (*Canis familiaris* and *Macaca mulatta*) the putatively non-homologous region was manually translated and confirmed to lack full or partial granulin domain sequences. Minor inconsistencies were noted for the majority of *in silico*-generated genomic architectures, including the frog *progranulin* gene which is presumably incomplete as it lacks a signal peptide and stop codon. See section 2.4 for details

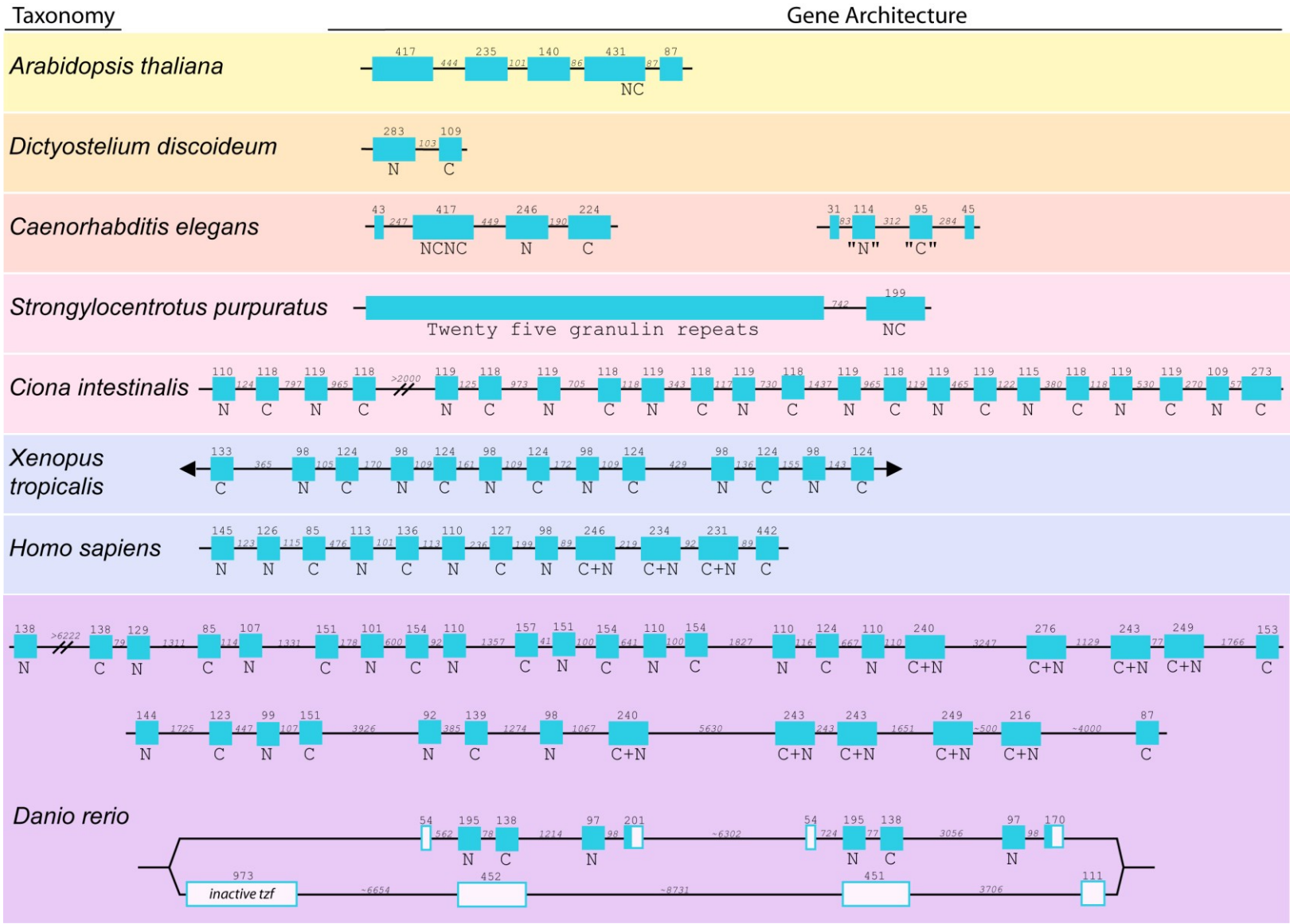

Taxonomy
Gene Architecture
Arabidopsis thaliana
417 444 235 101 140 86 431 87
NC
Dictyostelium discoideum
283 103 109
N C
Caenorhabditis elegans
43 247 417 449 246 190 224
NCNC N C
31 83 114 312 95 284 45
"N" "C"
Strongylocentrotus purpuratus
742 199
Twenty five granulin repeats
NC
Ciona intestinalis
110 124 118 797 119 965 118 >2000 119 125 118 973 119 705 118 118 119 343 118 117 119 730 118 1437 119 965 118 119 119 465 119 122 115 380 118 118 119 530 119 270 109 57 273
N C N C N C N C N C N C N C N C N C N C N C N C N C
Xenopus tropicalis
133 365 98 105 124 170 98 109 124 161 98 109 124 172 98 109 124 429 98 136 124 155 98 143 124
C N C N C N C N C N C N C
Homo sapiens
145 123 126 115 85 476 113 101 136 112 110 236 127 199 98 89 246 219 234 92 231 89 442
N N C N C N C N C+N C+N C+N C
138 >6222 138 79 129 1311 85 114 107 1331 151 178 101 600 154 92 110 1357 157 41 151 100 154 641 110 100 154 1827 110 116 124 667 110 110 240 3247 276 1129 243 77 249 1766 153
N C N C N C N C N C N C N C N C N C+N C+N C+N C+N C
144 1725 123 447 99 107 151 3926 92 385 139 1274 98 1067 240 5630 243 243 243 1651 249 ~500 216 ~4000 87
N C N C N C N C+N C+N C+N C+N C+N C+N C
Danio rerio
54 562 195 78 138 1214 97 98 201 ~6302 54 724 195 77 138 3056 97 98 170
N C N N C N
973 ~6654 452 ~8731 451 3706 111
inactive tzf

Figure 6. Comparison of the genetic representation and architectures of *progranulin* genes of non-chordate species in relation to human and zebrafish *progranulin* genes. The *Progranulin* gene number and architectures for plants (*Arabidopsis thalania*), slime mold (*Dictyostelium discoideum*), worms (*Caenorhabditis elegans*), sea urchin (*Strongylocentrotus purpuratus*), sea squirt (*Ciona intestinalis*), the diploid frog (*Xenopus tropicalis*), humans (*Homo sapiens*) and the zebrafish (*Danio rerio*) *progranulin* gene family are illustrated. For each exon the encoded domain fragment(s) is indicated below as either N-terminal (N), C-terminal (C) or a combination. Tabulated databases, genome builds and gene identifiers for each species are listed in section 2.4. For the smaller C.elegans (*Caenorhabditis elegans*) *progranulin* gene (reference identifier T02B11.8, see Section 2.4) the use of quotations surrounding the the N-terminal ("N") and C-terminal ("C") is used to indicate that the exon/intron boundaries within this gene do not uniformly bisect the granulin domain, an architecture unique to this *progranulin* gene. Introns for which the genomic sequencing is incomplete are indicated (e.g. *Ciona intestinalis* intron four; >2000). For the diploid frog (*Xenopus tropicalis*) the use of arrows at the 5' and 3' extremities indicates that an exon containing a *bona fide* signal peptide nor a final stop codon were indentified.

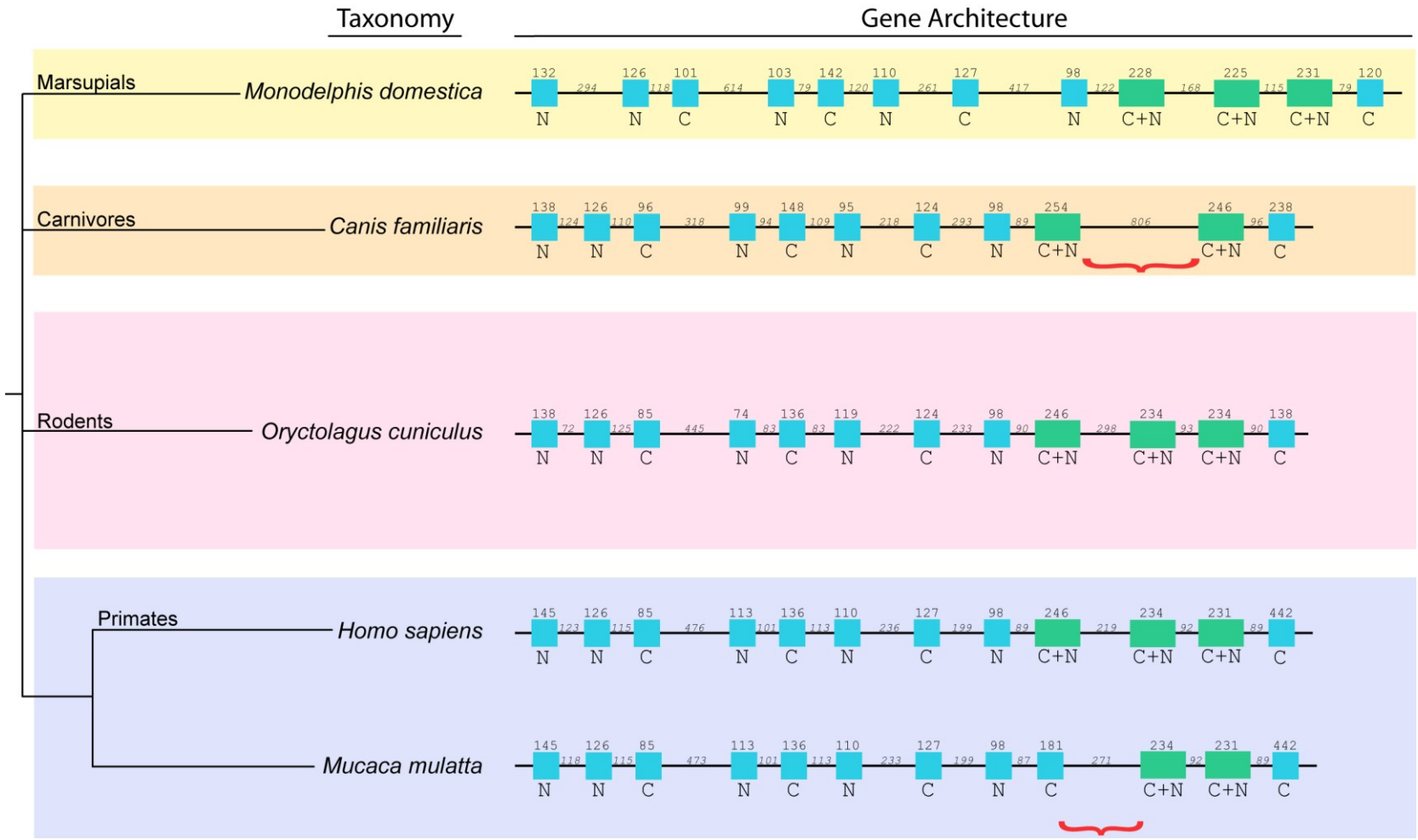

Taxonomy
Gene Architecture
Marsupials
Monodelphis domestica
Carnivores
Canis familiaris
Rodents
Oryctolagus cuniculus
Primates
Homo sapiens
Mucaca mulatta

Figure 7. The architecture of *progranulin* genes in mammals. The *progranulin* gene architectures for several species originating from a common lobe-finned (*sarcopterygian*) ancestor including the lesser hedgehog (*Monodelphis domestica*), the boxer dog breed (*Canis familiaris*), rabbit (*Oryctolagus cuniculus*), humans (*Homo sapiens*) and rhesus monkey (*Mucaca mulatta*) are illustrated. For each exon the encoded domain fragment(s) is indicated below as either N-terminal (N), C-terminal (C) or a combination. Tabulated databases, genome builds and gene identifiers for each species are listed in section 2.4. For those species whose *progranulin* gene architectures diverge significantly for that of humans, the region of interest is highlighted (underlying bracket).

3.2. Progranulin knockdown

3.2.1 Progranulin-B is necessary for early developmental events

Microinjection of progranulin-b-MO at 5ng/embryo (Figure 8) resulted in developmental arrest at ~4 hpf at the initiation of gastrulation with evidence of only partial dorsoventral axis formation. This is consistent with the functional ablation of *progranulin* in mice that leads to developmental arrest prior to implantation (Qin et al., 2005b; Xu et al., 1998). Furthermore, these data both, reinforce the prediction of early embryonic lethality in a mammalian knockout, and highlight the cross-species preservation of progranulin function.

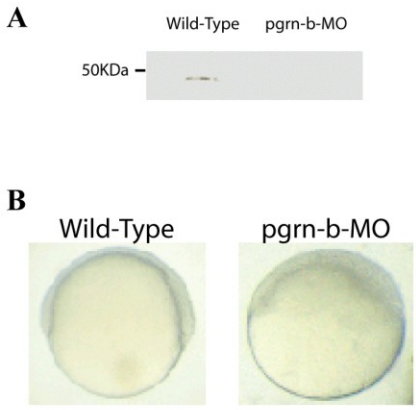
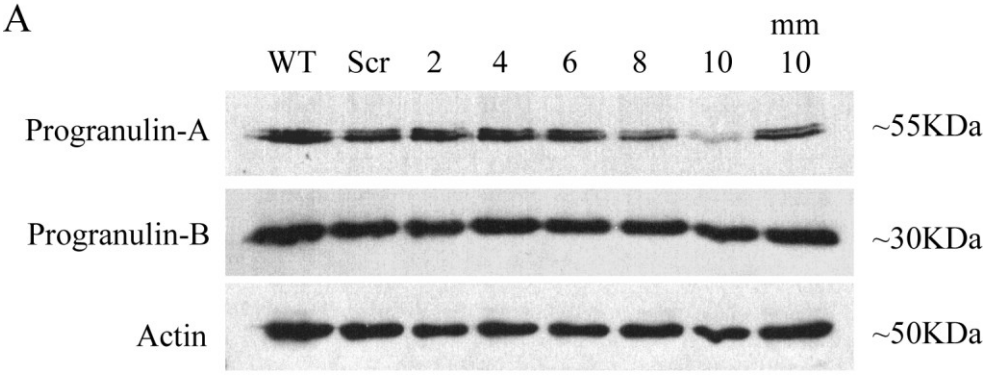


Figure 8. Preliminary phenotypic assessment and validation of zebrafish progranulin-b knockdown. Morpolino-based translation inhibition of progranulin-B results in developmental arrest at ~4hpf. (A) Effective translation inhibition of a morpholino targeting the 5' UTR of progranulin-b at a concentration of 5ng/embryo as assessed by Western blot analysis. Progranulin-B is expressed as a single band ~30 kDa. (B) Wild-type (WT) and a progranulin-b morpholino (pgrn-b-MO; 5 ng/embro) injected embryo at 8hpf. Whereas Wild-type embryos progress through gastrulation pgrn-b-MO injected embryos appear morphologically equivalent to that expected for a 4hpf embryo.

3.2.2 Validation of progranulin-A translation inhibition during zebrafish development – Of the four zebrafish family members, *progranulin-a* represents the syntenic orthologue to mammalian *progranulin* and recapitulates a subset of known murine and human expression patterns (Cadieux et al., 2005). To determine the functional relevance of progranulin-A during zebrafish development we employed antisense morpholino oligonucleotides (MOs) directed against the 5' untranslated region (UTR) and start (ATG) codon of the *progranulin-a* mRNA. Of three independent MOs targeting the 5'UTR, only one (hereafter referred to as UTR-MO) effectively reduced the translation of *progranulin-a in vivo*. A concentration of 10 ng/embryo of *progranulin-a* MO effectively decreased *progranulin-a* translation (Figure 9A) without affecting the expression of the progranulin-B homologue. At higher doses embryos displayed neural necrosis at mid-segmentation stages (data not shown) which we considered a result of toxicity, therefore a concentration of 10ng/embryo was used for all further experimentation. Injection of the equivalent amount of a 5 base-pair mismatch morpholino (mm) only slightly decreased *progranulin-a* translation (Figure 9A). Progranulin-A morphants exhibit few morphological manifestations prior to 72hpf, with the exception of an overall decrease in animal size. By 120hpf progranulin-A morphants display prominent craniofacial dysmorphogenesis, pericardial oedema and gut dysplasia (Figure 9B).

A second morpholino was designed to affect pre-RNA maturation, specifically exon-intron splicing. The splice-MO (Sp-MO) targets the fourth exon/intron boundary of progranulin-a (Figure 10A). The microinjection of Sp-MO at 10ng/embryo faithfully recreated the overall visible morphological defects including craniofacial dysmorphogenesis, gut dysplasia and pericardial edema (Figure 10B).

A


WT
Scr
2
4
6
8
10
mm 10
Progranulin-A
~55KDa
Progranulin-B
~30KDa
Actin
~50KDa

B

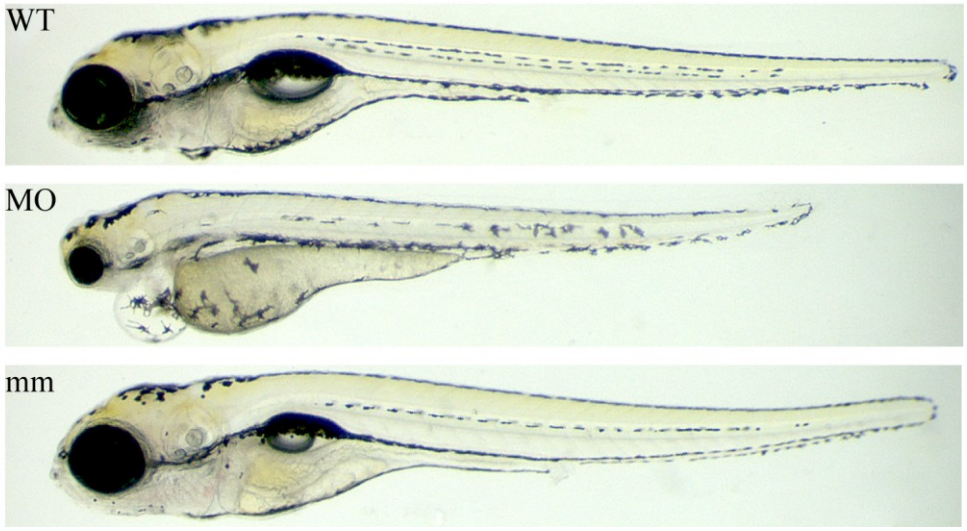
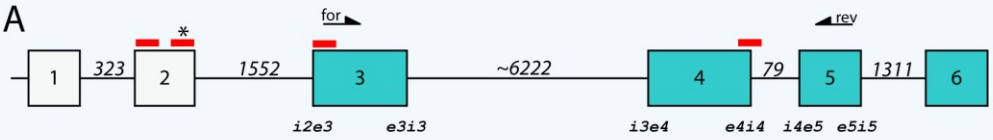

WT
MO
mm

Figure 9. Morpolino-based translation inhibition of progranulin-A results in craniofacial dysmorphogenesis and gut dysplasia. (A) Dose-dependent efficacy of a morpholino targeting the 5' UTR of *progranulin-a* was assessed by Western blot analysis in comparison to its co-orthologue progranulin-b and actin. Both progranulin-A and –B are represented by a single band roughly 55 and 30 kDa, respectively. (B) Live zebrafish embryos at 120 hpf, comparing wild-type (WT), progranulin-a morpholino (MO; 10 ng/embryo) and a 5 base-pair mismatch (mm; 10 ng/embryo) injected embryos. MO injected larvae exhibit an overall decrease in size as well as pericardial oedema, craniofacial and gut defects. Western blot analysis of MO efficacy was determined for three independent injection sets.

A



Morpholino	Exon Excision Effect
i2e3	Loss of signal peptide
e3i3	" "
i3e4	Secreted product with N-terminal half domain and 12 full domains
e4i4	" "
i4e5	Secreted prodcut with C-terminal half domain and 12 full domains
e5i5	" "

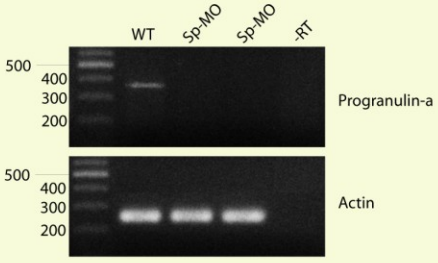
Morpholino	Intron Inclusion Effect
i2e3	Possible alternative starter MET or No effect
e3i3	Early termination; Secreted N-terminal half domain only
i3e4	" "
e4i4	Early termination; Secreted as single domain
i4e5	" "
e5i5	Early termination; Secreted one and one N-terminal half domains

B

Excision of exon 4

Inclusion of intron 4

C



D

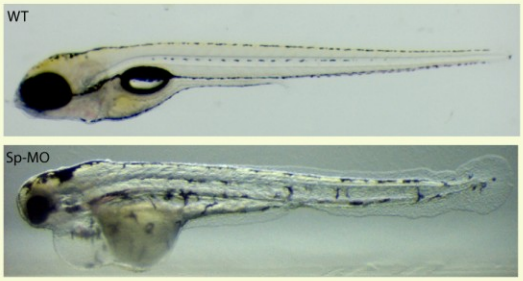


Figure 10. Design and validation of a splice-morpholino directed against progranulin-a. (A) Architecture of the 5' region of zebrafish *progranulin-a* derived from the automated Ensembl organization ENSDARG00000004954 and published cDNA sequence (Cadieux et al., 2005; Hubbard et al., 2007). Positions of the four progranulin-a targeted MOs (red boxes), primers used in RT-PCR (arrows), nomenclature for each intron/exon (e.g. i3e3) and exon/intron (e3i3), and lengths of introns are indicated. Open (white) boxes indicate untranslated exons while those which contain part of the open reading frame are illustrated in blue. The functional MO used throughout this study is targeted against the 3' end of the second untranslated exon. (B) Description of the presumed protein products generated by various splice-directed MOs. The putative transcripts generated by a MO directed against exon4/intron4 (e4i4) are illustrated. (C) Live zebrafish embryos at 120 hpf comparing wild-type (WT) and Sp-MO (Sp-MO; 10 ng/embryo) injected embryos. (D) RT-PCR results for the expression of *progranulin-a* and actin from individual WT or Sp-MO (10ng/embryo) injected embryos at 24hpf.

However, assessment of the presumed alteration in *progranulin*-a transcript size was not possible, since the morphants do not appear to express progranulin-a (Figure 10C).

3.2.3 Progranulin-A knockdown and hematopoietic alterations – The expression of progranulin-a within the ICM occurs relatively late (24 hpf) compared to other hematopoietic markers such as embryonic *globins*, *gata-1* and *scl*, precluding its involvement in hemangioblast formation (~11 hpf) or initial HSC specification (11−18 hpf) (Amatruda and Zon, 1999). As the embryo matures, *progranulin-a* expression increases progressively within the dorsal aorta, posterior ICM, and peripheral neutrophils and macrophages (Cadieux et al., 2005). This pattern of *progranulin-a* expression within definitive HSCs and peripheral leukocytes continues throughout all developmental stages analyzed (24−120 hpf). Preliminary results obtained with the progranulin-A knockdown morphant highlight the functional requirement of this growth factor during HSC development.

As expected the primitive HSC progenitors (which form prior to 24hpf) are unaffected in 83% (n=19/23) of progranulin-A morphants (Figure 11A). However, the definitive HSC population within the ICM appears absent in 83% (n=35/41, Figure 11B), as do myelo-erythroid progenitors in 67% (n=12/18, Figure 11C) of progranulin-A morphants. In conjunction with the absence of HSCs, downstream markers of macrophages, neutrophils and erythrocytes are substantially decreased (Figure 11D,F). Based on these results, we suggest that progranulin-A is required for the maintenance of definitive HSCs (Figure 11G). Whether this is the result of alterations in a common multipotent progenitor, a distinct progranulin-A requirement in each lineage, or both is currently unknown. It is

worthwhile investigating the expression patterns of other hematopoietic genes such as *c-myb*, *runx1* and *Pu.1* to further characterize the progranulin-A morphant phenotype.

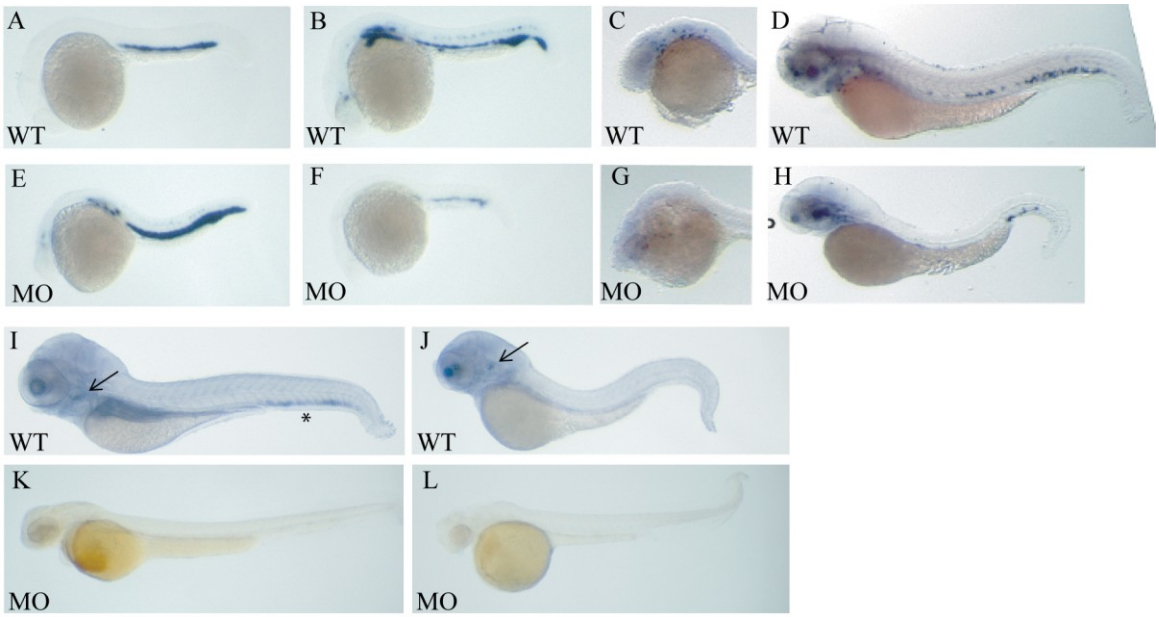


Figure 11. Hematopoietic defects in progranulin-A morphant embryos. Zebrafish were microinjected with a morpholino targeting the 5' UTR of *progranulin-a* (MO; 10ng/embryo) and assessed for gene expression alterations HSCs and downstream lineages. Wild-type (A-D,I,J) and progranulin-A morphant (E-H,K,L) expression patterns of HSCs using *Scl* at 24 hpf (A,E), (B,F) myelo-erythroid progenitors using *Gata-1* at 24hpf (B,F), monocyte/macrophages using *L-plastin* at 24 hpf (C,G), neutrophils using *Mpo* at 72 hpf (D,H), definitive HSCs using *ikaros* at 72 hpf (I,J) and red blood cells using the haemoglobin histochemical stain O-dianisidine at 48 hpf (K,L). For the expression pattern of *ikaros* the position of the thymus (arrows) and a second HSC pool within the ventral tail region (asterisk) are illustrated. Results for each marker were generated from a single microinjected clutch.

3.2.4 Progranulin-A knockdown and craniofacial dysmorphogenesis – To characterize the extent to which the craniofacial architecture is perturbed in progranulin-A morphants the neuro- and viscerocranial cartilage at 120 hpf using alcian blue staining was analyzed (Figure 12) (Neuhauss et al., 1996). In conjunction with the decrease in overall body size, cartilaginous structures were similarly decreased while some viscerocranial elements were missing. Specifically, the Meckel's, palatoquadrate, ceratohyal and hyosympletic structures were all decreased (Figure 12B,F). The orientation of the ceratohyal bones were further affected in tandem with the loss of the symplectic rod region (Figure 12B,H). The neurocranial cartilage was similarly decreased in size but remained unaffected with respect to orientation and architecture (Figure 12D). In contrast, the ventral ceratobranchial elements were absent in progranulin-A morphants (Figure 12B). The loss of ceratobranchial cartilages 1-5 occurred in 83% of MO injected embryos (n=127) whereas 10% of mismatch injections (n=96) displayed this craniofacial defect.

3.2.5 Progranulin-A is necessary for the maintenance and morphogenesis of ectomesenchymal arches – The formation of the jaw structure during development relies on signals originating from various tissue sources. Craniofacial dysmorphogenesis typically reflects direct or indirect disruption of the cranial neural crest (NC) cells or alterations in the patterning of their ectomesenchymal arch (EMA) derivitives. To identify the developmental stage at which potential NC defects appear in progranulin-A morphants we utilized the transgenic *fli1:EGFP* zebrafish to monitor NC migration in live embryos (Figure 13A-F) (Lawson and Weinstein, 2002). Defects uncovered using this

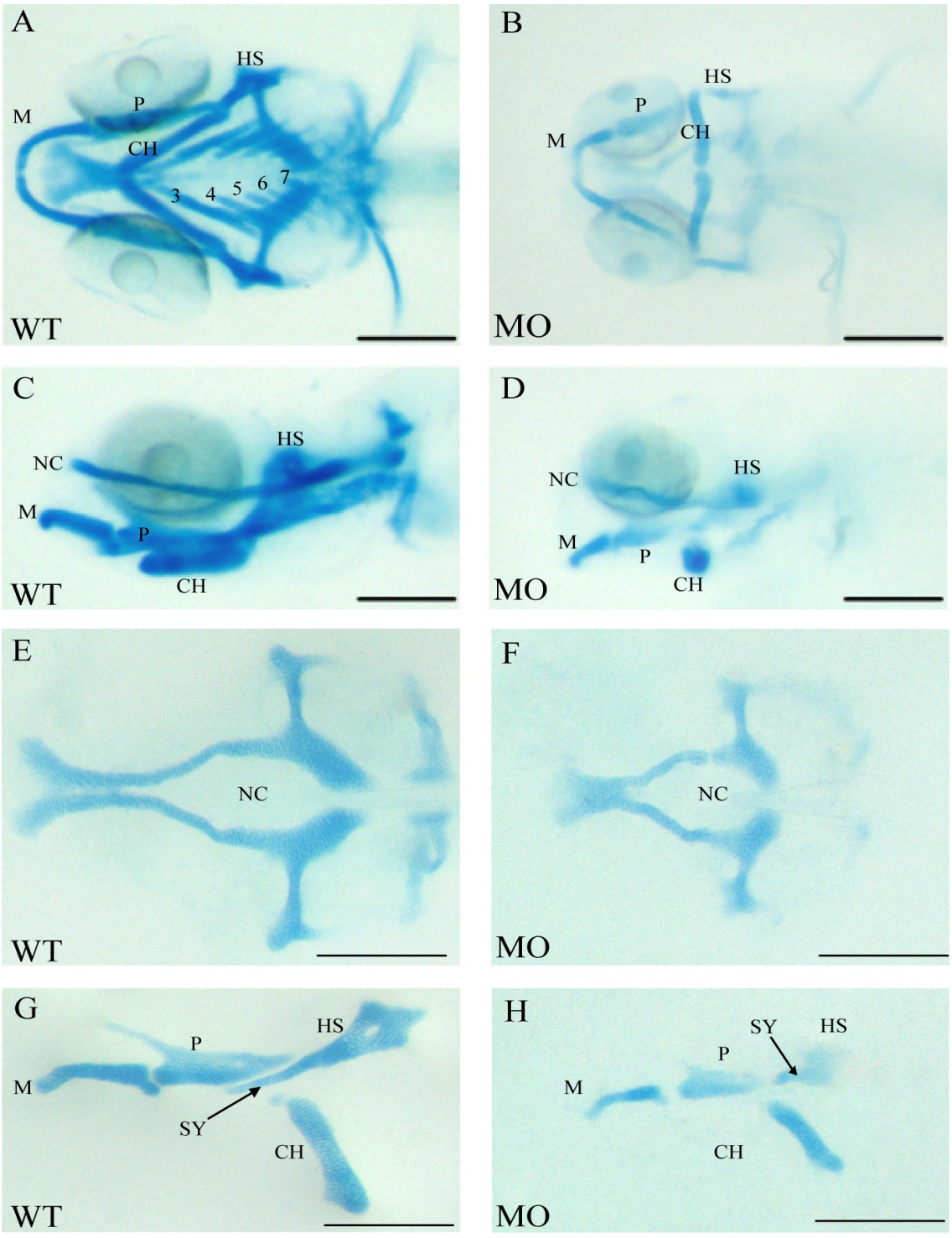

A
HS
M
P
CH
3
4
5
6
7
WT
B
HS
P
M
CH
MO
C
HS
NC
M
P
WT
CH
D
NC
HS
M
P
MO
CH
E
NC
WT
F
NC
MO
G
HS
P
M
SY
CH
WT
H
SY
HS
P
M
CH
MO

Figure 12. Progranulin-A morphants display craniofacial defects affecting the ventral viscerocranium. Alcian blue staining of wild-type (A,C,E,G) and progranulin-A morphant (B,D,F,H) larvae at 120 hpf positioned in ventral (A,B) or lateral orientations (E,F). The Meckel's (m) and palatoquadrate (p) as well as the ceratohyal (ch) and hyosymplectic (hs) elements that originate from the primary and secondary pharyngeal arch mesenchyme respectively, are present but reduced in progranulin-A morphants (B,F). The neurocranium (nc) demonstrates a similar overall decrease in size (D). The alteration in ceratohyal orientation (B,F) originates from the loss of the symplectic rod region (sy) of the hyosymplectic jaw joint (H). Cartilagenous elements originating from the ventral pharyngeal arches that include the ceratobranchial series (cb1-5) are absent in progranulin-A morphants (B).

method were confirmed based on alterations in *dlx2* expression (Figure 13G-L) (Akimenko et al., 1994). The formation and initial ventral migration of the first three NC cell streams from the hindbrain was unaffected in progranulin-A morphants at 24 hpf (Figure 13D,J). Additionally, the formation of four distinct EMAs (streams 3 and 4; Figure 13D,J) was also unaffected. However, further division of the most ventral EMA population between 28 and 34 hpf did not occur in progranulin-A morphants (Figure 13E,F,K,L). In addition, all EMA segments are reduced at 34 hpf, particularly populations 3 and 4, in keeping with the lack of ceratobranchial cartilage which originates from these cell populations.

3.2.6 Progranulin-A is expressed within the pharyngeal endoderm and is required for morphogenesis – It has been previously demonstrated that the division of the third most posterior EMA segments is dependent on the lateral migration and infiltration of this ectomesenchyme by the pharyngeal endoderm (Crump et al., 2004; Piotrowski and Nusslein-Volhard, 2000; Reiter et al., 1999). The expression of several endodermal markers to evaluate potential alterations in the pharyngeal endoderm or lateral pouch formation was investigated (Figure 14A-G,J). At 28 hpf *nkx2.3* is expressed within the pharyngeal endoderm destined to form the craniofacial pouches (Lee et al., 1996). In progranulin-A morphants the expression of *nk2.3* is generally unaffected, however the most posterior pouch is absent, in keeping with the loss of continued posterior EMA segmentation (Figure 14D). The absence of continued endodermal pouch formation was confirmed based on Zn8 immunolocalization at 34 hpf (Figure 14E). To determine if the alterations in posterior arch development was due to apoptosis we labelled fragmented DNA using TUNEL and undertook immunohistochemical detection of Zn8 to provide

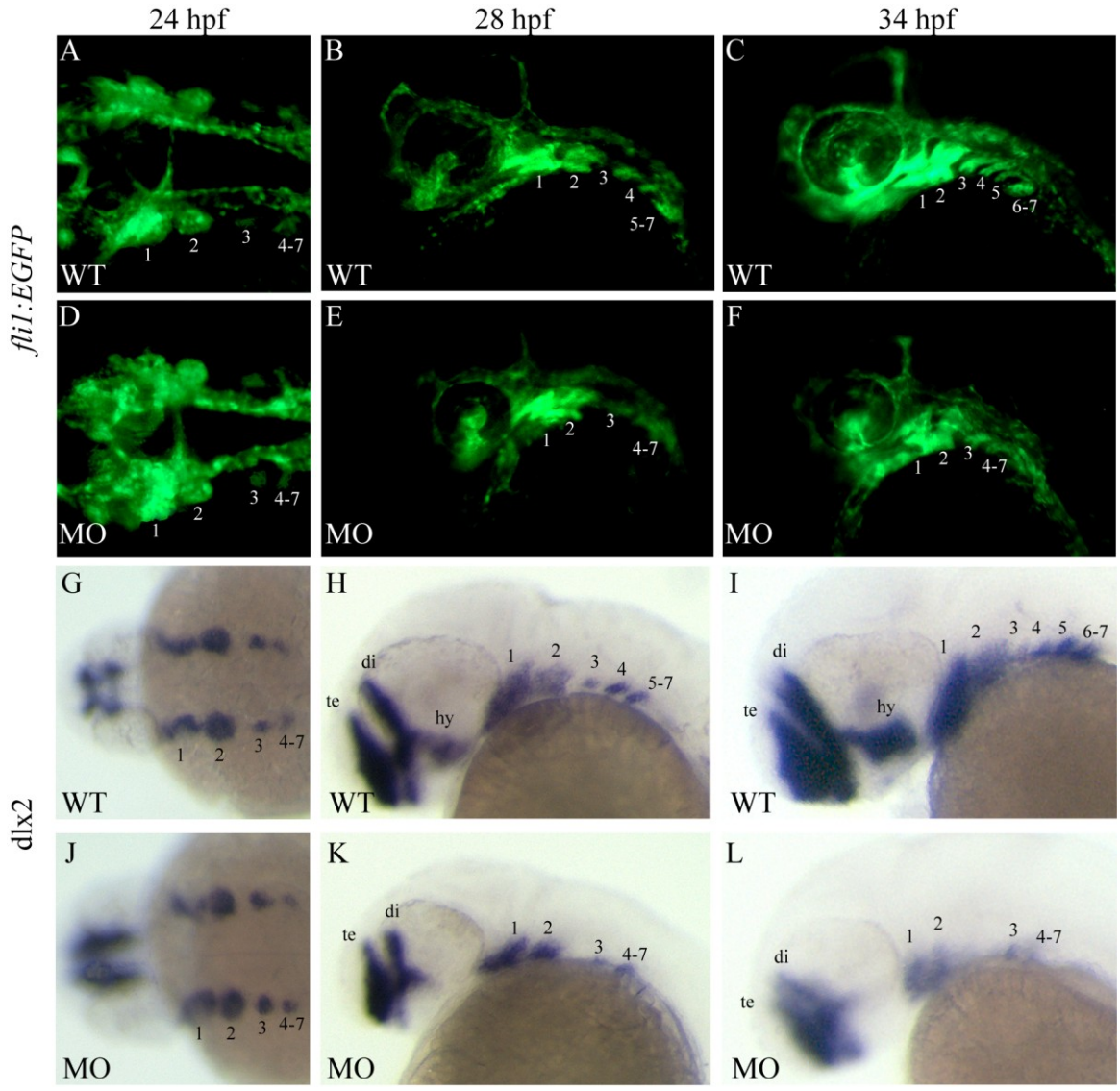

24 hpf
28 hpf
34 hpf
fli1:EGFP
dlx2
A
B
C
D
E
F
G
H
I
J
K
L
WT
MO
1
2
3
4-7
4
5-7
5
6-7
di
te
hy

Figure 13. Maintenance and subdivision of the ventral ectomesenchymal cells is perturbed in progranulin-A morphants. Analysis of cranial ectomesenchymal arch (EMA) migration using the expression of the *fl1i:EGFP* transgene in live embryos (A-F) and in situ hybridization for *dlx2* (G-L) at 24 (A,D,G,H), 28 (B,E,H,K) and 34 hpf (C,F,I,L). The cranial EMAs (numbered anterior to posterior, 1-7) appear unaffected at 24 hpf in progranulin-A morphants, dorsal view (D,J). Continued expression of both the *fl1i:EGFP* transgene and *dlx2* decreases particularly for the ventral EMA 3-7 from 28 to 34 hpf, lateral views (E,F,K,L). Although the initial subdivision of EMAs is normal (D,J) continued partitioning of the posterior 4th stream fails to develop (F,L). The expression pattern of *dlx2* in progranulin-A mophants is otherwise unaffected in the telencephalon (te) and diencephalon (di), however the hypothalamus (hy) is absent.

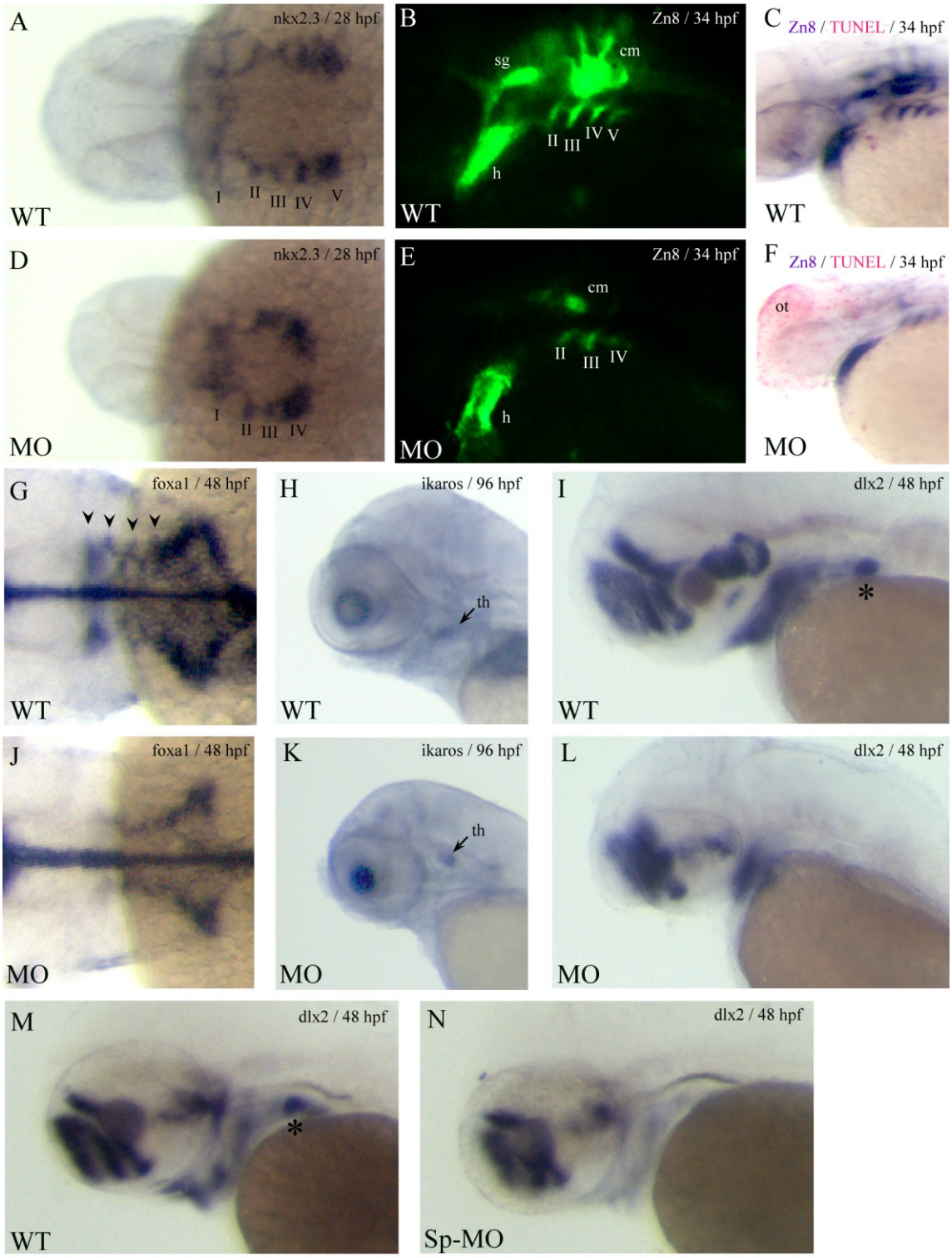

A
nkx2.3 / 28 hpf
I II III IV V
WT
B
Zn8 / 34 hpf
sg
cm
II III IV V
h
WT
C Zn8 / TUNEL / 34 hpf
WT
D
nkx2.3 / 28 hpf
I II III IV
MO
E
Zn8 / 34 hpf
cm
II III IV
h
MO
F Zn8 / TUNEL / 34 hpf
ot
MO
G
foxa1 / 48 hpf
WT
H
ikaros / 96 hpf
th
WT
I
dlx2 / 48 hpf
*
WT
J
foxa1 / 48 hpf
MO
K
ikaros / 96 hpf
th
MO
L
dlx2 / 48 hpf
MO
M
dlx2 / 48 hpf
*
WT
N
dlx2 / 48 hpf
Sp-MO

Figure 14. Posterior pharyngeal endoderm fails to migrate laterally and form distinct pouches in progranulin-A morphants. (A,D) Dorsal view of the expression of *nkx2.3* at 28 hpf highlights endodermal pharyngeal pouches (numbered anterior to posterior I-V) where the most posterior pouch (IV) fails to subdivide in progranulin-A morphants. (B,E) Lateral view of the endodermal pouches based on Zn8 immunofluorescence at 34 hpf demonstrates the continued prohibition of posterior pouch subdivision. In addition, there is a marked decrease in Zn8 expression within sensory ganglia (sg) and hindbrain commissural neurons (cm) of progranulin-A morphants. (C,F) Dual Zn8 immunohistochemistry and TUNEL detection at 34 hpf indicate the apparent loss of ventral EMAs and pharyngeal pouches is not associated with aberrant apoptosis within these elements, however TUNEL staining was detected throughout the optic tectum (ot). (G,J) Dorsal view of the expression of *fkd7* within the pharyngeal endoderm (pe) at 48 hpf demonstrates that the pe is maintained within progranulin-a morphants as bilateral streams, however these cells fails to migrate laterally to form pouches. (H,K) Lateral view of the expression of *ikaros* at 96 hpf demonstrates that the pe is still capable of migrating to contribute to the formation of the thymus (th). (I,L-N) Lateral view of *dlx2* expression at 48 hpf reveals that progranulin-A morphants derived from either 5'UTR targeting (MO; L) or the use of a splice MO (Sp-MO; N) fail to maintain the posterior cranial EMAs in addition to defective migration.

orientation with respect to endodermal pouch positioning (Figure 14C,F). Despite a clear upregulation of apoptosis within the optic tectum, progranulin-A morphants did not demonstrate increased programmed cell death within or surrounding the endodermal pouches (Figure 14F). Similarly, detection of apoptosis using TUNEL at 48 and 72 hpf did not highlight any regions within the craniofacial arches (data not shown). The absence of apoptosis and posterior arch segmentation within both pharyngeal endoderm and EMAs may have reflected a developmental delay. To determine if pouch formation occurred in progranulin-A morphants at later stages we analyzed the expression of *foxa1* at 48 hpf (Figure 14G,J). Although Progranulin-A morphants express *foxa1* within the bilateral pharyngeal streams, the lack of lateral branches was evident, arguing against developmental delay (Figure 14G). Despite the apparent failure of the pharyngeal endoderm to migrate properly, other organs dependent on these cells for development, specifically the thymus, were unaffected (Figure 14K) (Reiter et al., 2001). In contrast to the apparent morphogenetic defect within the pharyngeal endoderm, the ventral EMAs are absent in progranulin-A morphants at 48 hpf based on *fli1:EGFP* and *dlx2* expression (Figure 14L, data not shown). This suggests that progranulin-A may be necessary for the maintenance as well as the segmentation of ventral EMA populations. The loss of *dlx2* at 48 hpf within fifth EMA segment occurred in 91% of progranulin-A morphants (n=48) whereas 20% and 50% of mismatch injected embryos (n=44) demonstrated complete or partial loss, respectively.

The pharyngeal endoderm and epidermis have been previously shown to express *progranulin-a* between 24 and 72 hpf (Cadieux et al., 2005). Both of these tissues generate signals required for normal viscerocranial arch formation (David et al., 2002; Knight et al., 2005). To determine if the expression of *progranulin-a* within the

pharyngeal endoderm is specifically necessary for craniofacial morphogenesis, *progranulin-a* targeted MO was injected into a single marginal blastomere of embryos at the 16-cell stage, thus restricting progranulin-A knockdown to mesodermal and endodermal derivates (Peyrieras et al., 1998). Subsequent alcian blue staining demonstrates that these chimeric embryos asymmetrically replicate the craniofacial defects of progranulin-A morphants, suggesting that progranulin-A derived from the pharyngeal endoderm is necessary for proper craniofacial morphogenesis (Figure 15).

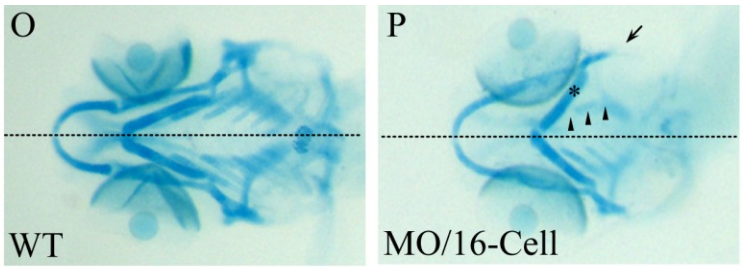


Figure 15. Progranulin-a expressed within the pharyngeal endoderm is responsible for proper development of the posterior viscerocranium. (A,B) Ventral view of alcian blue staining of cartilaginous elements at 120 hpf in wild-type (A) and embryos injected with progranulin-a directed MO within a single marginal blastomere at the 16-cell stage (B). The asymmetric MO distribution results in defects within the left side only, including alterations in ceratohyal (*) and hyosymplectic (arrow) development as well as absence of ceratobranchial elements (arrowheads).

3.2.7 Progranulin-A is required for continued growth and morphogenesis of the liver and exocrine pancreas – At 120hpf progranulin-A morphants clearly exhibit gut abnormalities which include the absence of the digestive tract and inflated swim bladder (Figure 9B). To investigate the status of the developing digestive system and associated organs, including the liver and pancreas, we followed the expression of a series of molecular markers at various developmental stages was followed (Figures 16 and 17). The expression of *pdx-1* and *isl-1* within the developing endocrine pancreas are generally unaffected progranulin-A morphants at 48 hpf (Figure 16G,H). Analysis of the general endodermal marker *foxa1* at 48 hpf reiterates this decrease in the liver size where the hepatic bud appears to maintain a close association with the developing intestine, failing to infiltrate the lateral mesenchyme (Figure 16I). The continued development of the liver rudiment is not observable since the expression of *foxa1* and the differentiation marker *transferrin* are absent at 72 hpf (Figure 16J,K). The loss of *transferrin* expression at 72 hpf occurred in 82% of progranulin-A morphants (n=120) whereas 25% and 21% of mismatch injected embryos (n=60) demonstrated complete or partial loss, respectively. Similar to the late onset defect in liver development, *foxa1* expression indicates that the intestine is present and has undergone correct left-right asymmetry by 48 hpf (Figure 16I). However its continued development, including cytodifferentiation appears absent in progranulin-A morphants based on the lack of *ifabp* expression (Figure 16L).

With respect to development of the pancreas, the exocrine lineage appears unaffected with respect *pdx-1* and *isl-1* expression and its presence remains observable at 72 hpf based on *insulin* expression (Figure 16D). In contrast, the exocrine pancreas is absent at

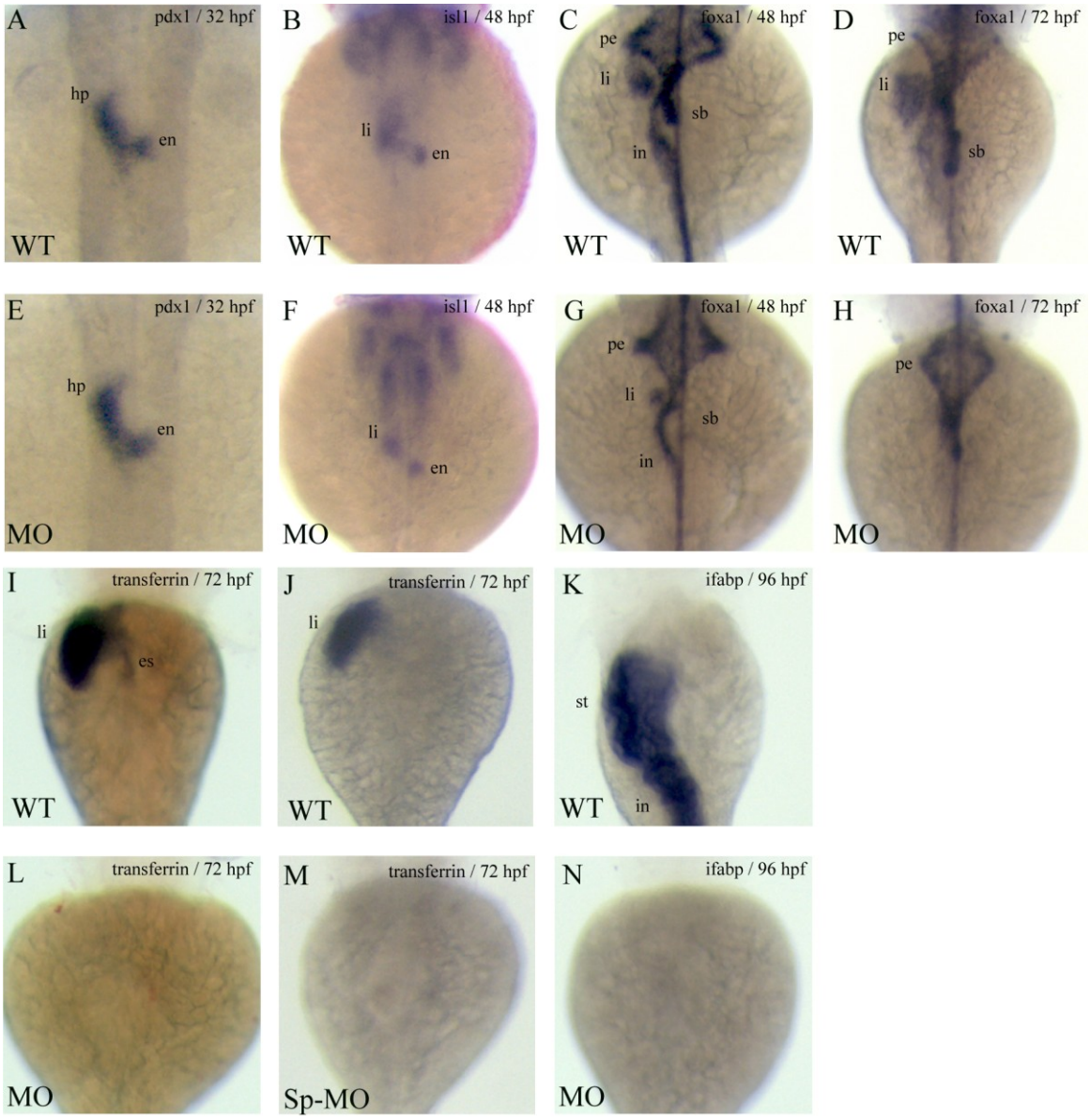

A
pdx1 / 32 hpf
hp
en
WT
B
isl1 / 48 hpf
li
en
WT
C
foxa1 / 48 hpf
pe
li
sb
in
WT
D
foxa1 / 72 hpf
pe
li
sb
WT
E
pdx1 / 32 hpf
hp
en
MO
F
isl1 / 48 hpf
li
en
MO
G
foxa1 / 48 hpf
pe
li
sb
in
MO
H
foxa1 / 72 hpf
pe
MO
I
transferrin / 72 hpf
li
es
WT
J
transferrin / 72 hpf
li
WT
K
ifabp / 96 hpf
st
WT
in
L
transferrin / 72 hpf
MO
M
transferrin / 72 hpf
Sp-MO
N
ifabp / 96 hpf
MO

Figure 16. Progranulin-A morphants display alterations in continued development of the digestive system and liver. (A,E) Expression of *pdx1* at 34 hpf within the endocrine pancreas (en) is unaffected in progranulin-A morphants. (B,F) Expression of *isl1* at 48hpf within the endocrine pancreas is unaffected. (C,D,G,H) Expression of the broad endodermal marker *foxa1* that demarcates the pharyngeal endoderm (pe), liver (li), swim bladder (sb) and intestine (in) at 48 (C,G) and 72 hpf (D,H). The expression of *foxa1* in progranulin-A morphants is decreased within the liver and confined to a region directly lateral to the developing gut at 48 hpf (G) and completely absent within this organ by 72 hpf (H). (I,L) Expression of the liver differentiation marker *transferrin* at 72hpf recapitulates the observed late onset hepatic defect and is absent in the esophagus of progranulin-A morphants. (J,M) The liver defect was confirmed using the splice MO (Sp-MO). (K,N) Expression of *ifabp* as a marker of intestinal differentiation within the stomach (st) and intestine (in) at 96 hpf is absent in progranulin-A morphants.

72 hpf based on *trypsin* and carboxypeptidase-A expression (Figure 17E,F). The loss of *trypsin* expression at 72 hpf occurred in 89% of progranulin-A morphants (n=49) whereas 17% and 34% of mismatch injected embryos (n=59) demonstrated complete or partial loss, respectively. Although *progranulin-a* was detected within the epithelial lining of the pharynx, intestine and swim bladder this pattern is not evident until 120 hpf (Cadieux et al., 2005). To relate the observed visceral defects to *progranulin-a* expression the *in situ* hybridization studies were repeated in order to identify patterns previously unidentified. *Progranulin-a* is expressed within the endocrine pancreas at 36 hpf, however it is not found within this tissue or its surroundings at 28 or 48 hpf implicating (data not shown) a temporally confined requirement for this growth factor in pancreas development (Figure 17G).

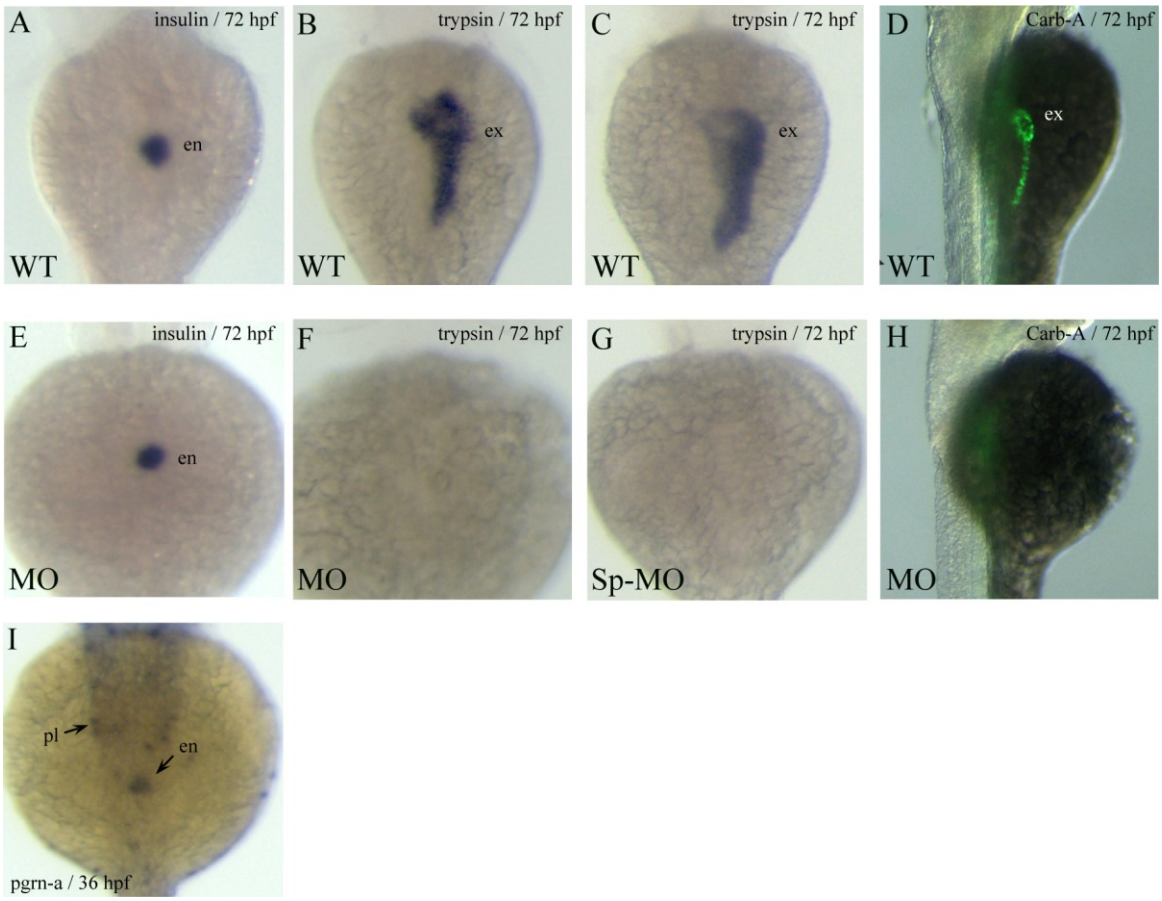


Figure 17. Progranulin-A is necessary for the development of the exocrine pancreas. (A,E) Expression of *insulin* within the endocrine pancreas (en) at 72hpf is unaffected in progranulin-A morphants. (B,F) Expression of *trypsin* within the exocrine pancreas (ex) at 72hpf is absent in progranulin-A morphants. (C,G) The exocrine pancreas defect was confirmed using the splice MO (Sp-MO). (D,H) Overlap of brightfield and immunofluorescence for Carboxypeptidase-A at 72hpf within the exocrine pancreas is absent in progranulin-A morphants. (I) Expression of progranulin-a within peripheral leukocytes (pl) and the endocrine pancreas (en) at 36 hpf.

3.3 Progranulin binding protein identification

3.3.1 Development of a novel protein-protein interaction screen; LC-SPR-MS – A system of proteomic fractionation by RP-HPLC which is coupled offline to the SPR protein-protein interaction screening method with final identification of the putative binding protein by mass spectrometry (LC-SPR-MS) was developed for this study. Recombinant zebrafish progranulin-A was used as bait to detect potential binding partners in an extract of zebrafish embryos. Recombinant carboxyl-terminally tagged zebrafish progranulin-A (progranulin-A/V5/His) was generated using the baculovirus expression system and purified to homogeneity from culture medium using a combination of immobilized metal ion affinity (IMAC) and size exclusion chromatography (Section 2.17-2.19). Recombinant progranulin-A/V5/His was covalently bound through amine chemistry to an SPR chip surface (CM4-type) in parallel with an activation/deactivation control surface. Confirmation of overall structural integrity was assessed by monitoring the antigenicity of the bound ligand using rabbit anti-zpgrna (Figure 18). Eight thousand staged zebrafish embryos (29 hours post-fertilization) were processed using a reversed-phase extraction procedure and subjected to RP-HPLC fractionation (Belcourt et al., 1993; Bennett et al., 1981). Column fractions were screened for binding activity by SPR through three rounds of RP-HPLC to yield a homogenous candidate binding protein (Figure 19). This putative progranulin-A binding partner was digested with trypsin and the resultant peptides analyzed using a SCIEX/Applied Biosystems LC/MS/MS mass spectrometer. For each LC/MS/MS analysis the resulting raw data was analyzed using Mascot against the zebrafish EST database with the taxonomy specific as *Danio rerio*. Sumarized MS/MS results for one SPR positive fraction are illustrated in Figure 20. Based on its primary structure, the progranulin binding protein (PBP) can be classified as

a member of the type VII secreted C-type lectin domain (CTLD) proteins which includes the fish type II antifreeze proteins and the mammalian regernerating islet-derived (REG) family (Figure 21) (Zelensky and Gready, 2004; Zelensky and Gready, 2005).

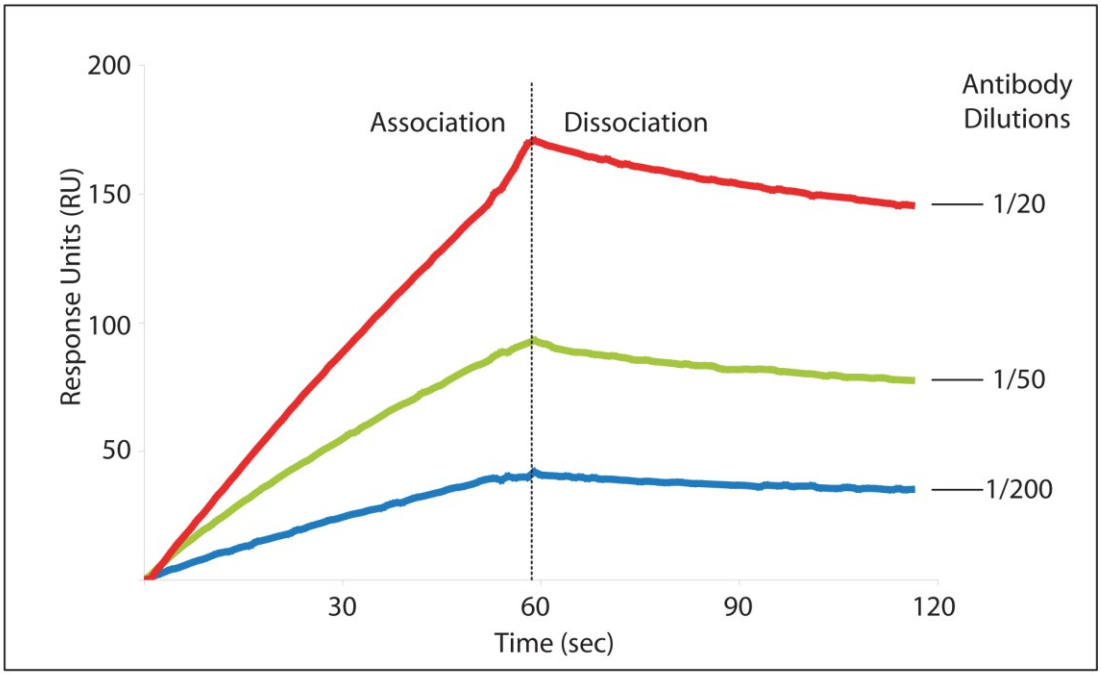
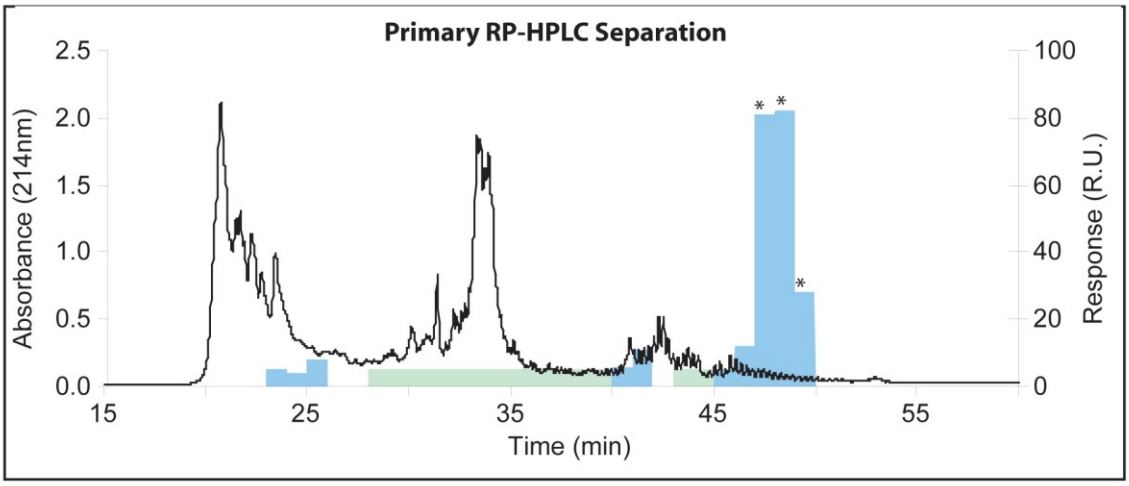
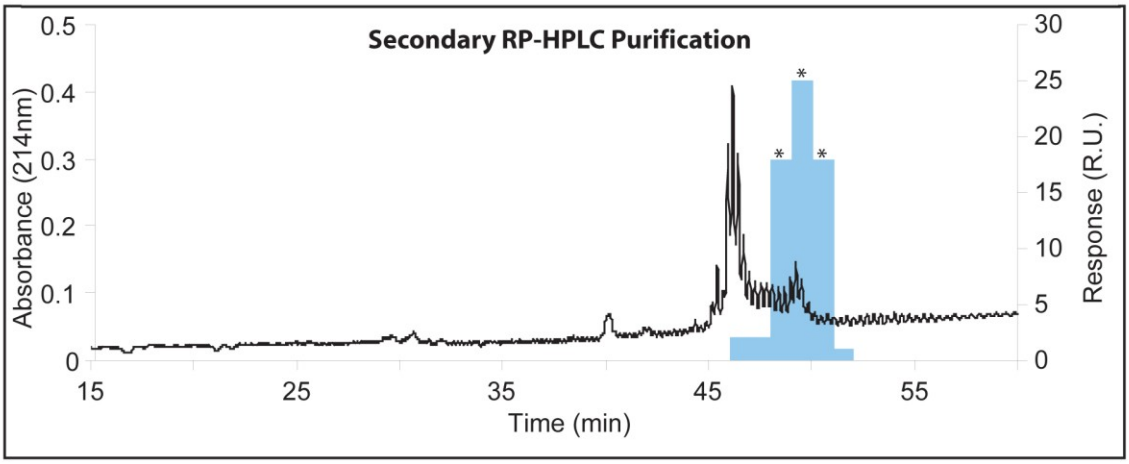


Figure 18. Validation of the Progranulin-A/V5/His ligand coupling and antibody kinetics. SPR analysis of the interaction between the site specific progranulin-A antibody and amine coupled progranulin-A/V5/His illustrating rapid association and slow dissociation kinetics. The interaction is characterized by biphasic association and dissociation kinetics.


Primary RP-HPLC Separation
Absorbance (214nm)
Response (R.U.)
Time (min)
2.5
2.0
1.5
1.0
0.5
0.0
100
80
60
40
20
0
15
25
35
45
55
* *
*


Secondary RP-HPLC Purification
Absorbance (214nm)
Response (R.U.)
Time (min)
0.5
0.4
0.3
0.2
0.1
0
30
25
20
15
10
5
0
15
25
35
45
55
*
* *

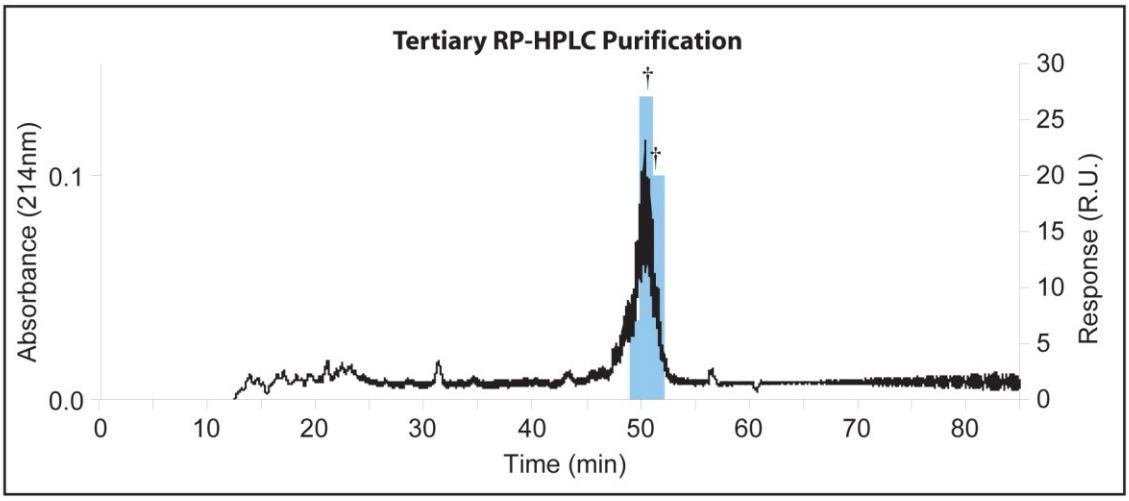
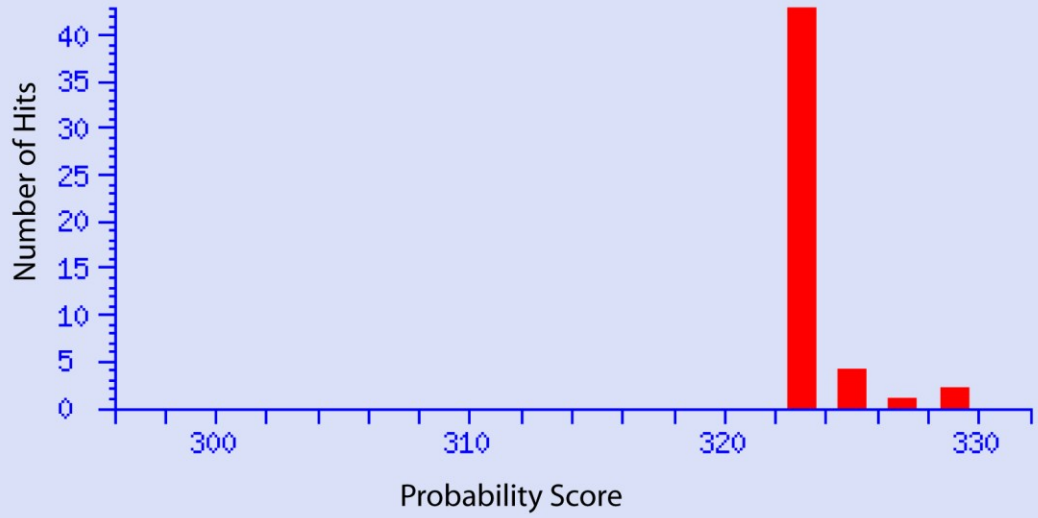

Tertiary RP-HPLC Purification
Absorbance (214nm)
Response (R.U.)
Time (min)
0.1
0.0
30
25
20
15
10
5
0
0
10
20
30
40
50
60
70
80
†
†

Figure 19. Purification of a progranulin-A/V5/His binding protein from zebrafish embryos. Overlay of RP-HPLC spectrophotometric chromatograms and surface plasmon resonance results response for each collected fraction. A reversed-phase extract of 8000 zebrafish embryos (29 hours post-fertilization) was subjected to sequential rounds of RP- HPLC. Aliquots from each column fraction were assayed for protein-protein interaction using SPR with recombinant progranulin-A/V5/His as bait. Specific progranulin-A binding is shown as response units. Fractions demonstrating a robust binding characteristic (asterisks) were subject to further purification. After three rounds of RP-HPLC and SPR screening active fractions were sent for protein identification by mass spectrometry (daggers) (See sections 2.20-2.23 for details).

A



B

Protein Summary Report

Accession Number	Score	Peptides Matched	Percent Coverage
BI710011	323	5	38%
BQ078431	323	5	38%
BM156776	270	4	36%
BI708498	266	5	37%
BM095856	211	2	17%
CD053083	76	2	9%
AI313718	72	1	12%

Figure 20. The indentification of a putative Progranulin-A binding protein. (A) Summarized Probability based mowse score generated by Mascot search engine for a single SPR positive fraction from the third round of purification. (B) Summarized protein identification list of all probability hits with a score greater than seventy. Percent coverage is expressed as the total coverage of the corresponding EST open reading frame. The first four proteins indentified (ESTs BI710011, BQ078431, BM156776, BI708498 and BM095896) correspond to the same protein, reffered to as progranulin binding potein (PBP). Similar results were observed in other SPR positive fractions, whereas SPR negative fractions did not identify this protein (data not shown). The ESTs CD053083 and AI373718 correspond to Keratin-type 5 and Keratin-type 4, respectively.

A

Deduced Coding Sequence of PBP

```
atgatggtgatgctgagaagtcttcttctgtttttcttcttgttctgcatgggcattgca
 M  M  V  M  L  R  S  L  L  L  F  F  F  L  F  C  M  G  I  A
gcagaaaggagatgcccacgtcaatggagaaggtctggatcgcgatgctttaggcttttc
 A  E  R  R  C  P  R  Q  W  R  R  S  G  S  R  C  F  R  L  F
tctacatcagtgaactgggccacagcagagaaaaactgtcaacgtcttggtgggaatctt
 S  T  S  V  N  W  A  T  A  E  K  N  C  Q  R  L  G  G  N  L
gcctctgtgctaaacgatgtggaaaatgattttctgctgagtctgatacccaactccaaa
 A  S  V  L  N  D  V  E  N  D  F  L  L  S  L  I  P  N  S  K
cgtttctttattggtggctataatgttgatgaacagaactggttttggtctgatggatct
 R  F  F  I  G  G  Y  N  V  D  E  Q  N  W  F  W  S  D  G  S
ccatttggttatacaaactggtgctcaggagagcctaacaacatgaatacggagcactgc
 P  F  G  Y  T  N  W  C  S  G  E  P  N  N  M  N  T  E  H  C
ttggagattaactggaccgcaaaccgttgttggaacaacttgccttgttcagttgaacta
 L  E  I  N  W  T  A  N  R  C  W  N  N  L  P  C  S  V  E  L
ggctatatctgtgccaagaacctaaatgattgctcttaa
 G  Y  I  C  A  K  N  L  N  D  C  S  -
```

B

Sequence Comparison of Zebrafish PBP and the Human REG Family

```
REG1a   ----MAQTSSYFMLISCLMFLSQSQGQEAQTELPQARISCPEGTNAYRSYCYYFNEDRET
REG1b   ----MAQTNSFFMLISSLMFLSLSQGQESQTELPNPRISCPEGTNAYRSYCYYFNEDPET
REG3a   MLPPMALPSVSWMLLSCLMLLSQVQGEEPQRELPSARIRCPKGSKAYGSHCYALFLSPKS
REG3g   MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKS
REG4    ----MASRSMRLLLLLSCLAKTGVLG------DIIMRPSCAPGWFYHKSNCYGYFRKLRN
PBP     ---------MMVMLRSLLLFFFLFCMG------IAAERRCPRQWRRSGSRCFRLFSTSVN

REG1a   WVDADLYCQNMNSGNLVSVLTQAEGAFVASLIKESGTDDFNVWIGLHDPKKNRR-----W
REG1b   WVDADLYCQNMNSGNLVSVLTQAEGAFVASLIKESSTDDSNVWIGLHDPKKNRR-----W
REG3a   WTDADLACQKRPSGNLVSVLSGAEGSFVSSLVKSIGNSYSYVWIGLHDPTQGTEPNGEGW
REG3g   WMDADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGW
REG4    WSDAELECQSYGNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPQKRQQ-----W
PBP     WATAEKNCQRLG-GNLASVLNDVENDFLLSLIPNSKR----FFIGGYNVDEQN------W

REG1a   HWSSGSLVSYKSWGIGAPSSVNPGYCVSLTSSTGFQKWKDVPCEDKFSFVCKFKN----
REG1b   HWSSGSLVSYKSWDTGSPSSANAGYCASLTSCSGFKKWKDESCEKKFSFVCKFKN----
REG3a   EWSSSDVMNYFAWERNPSTISSPGHCASLSRSTAFLRWKDYNCNVRLPYVCKFTD----
REG3g   EWSSTDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD----
REG4    QWIDGAMYLYRSWS--GKSMGGNKHCAEMSSNNNFLTWSSNECNKRQHFLCKYRP----
PBP     FWSDGSPFGYTNWCSGEPNNMNTEHCLEIN-WTANRCWNNLPCSVELGYICAKNLNDCS
```

Figure 21. Deduced coding sequence of the putative progranulin-A binding protein and alignment with its orthologous mammalian family members. (A) The nucleotide and corresponding amino acid sequence of the progranulin-A binding protein (PBP) based on mass spectrometry data mining by MASCOT and subsequent TBLASTN search results of zebrafish expressed sequence tag libraries. (B) Comparison of the sequences zebrafish PBP and members of the human REG family.

4. Discussion

4.1 Diversification of the progranulin gene family

4.1.1 The intragenic plasticity of progranulin genes – The suggested evolutionary pressure to maintain *progranulin* representation within organisms as a single gene (with the exception of the teleost lineage) is contrasted with the high variation in the number of tandem granulin repeats that are encoded within each gene (Figure 6). Perhaps the most striking examples of this plasticity is evident through the comparison of the chicken (*Gallus gallus*), humans, sea squirts (*Ciona intestinalis*) and the sea urchin (*Strongylocentrotus purpuratus*) *progranulin* genes which encode four, seven and one half, eleven and twenty-two granulin modules, respectively.

The variation of internally repeated domains is not limited to the *progranulins*. Several genes that consist of tandem repeated protein motifs demonstrate a relative tendency towards expansion and/or contraction across species including the Kruppel-related zinc finger proteins and the structural muscle protein titin (Kenny et al., 1999; Looman et al., 2002). Interestingly, organism complexity (in particular prokaryotes relative to multicellular eukaryotes) appears to increase in conjunction with an increase in the number of tandem repeats within several gene families (Bjorklund et al., 2006; Marcotte et al., 1999). It has been hypothesized that such intragenic domain expansion is a force for evolutionary change, similar to that of genome-wide duplication.

Previously, the concept of gene duplications and the subsequent theoretical release of molecular constraints for one orthologue was considered a reasonable explanation to understand the drive to retain both copies (Ohno et al., 1968). This statement is clearly applicable to the aforementioned example of the two zebrafish Nodal co-orthologues

(Section 1.2.3), although the evolution of novel function did not involve gross changes in domain repeat number. Interestingly, the origins of *progranulin-a* and *–b,* presumably products the 3R genome-wide duplication ~350mya, has given rise to two genes that no longer share a common number of granulin repeats (Figure 2 and 6). Moreover, the presence of *progranulin-1* and *-2* represent the extremity of gene contraction as both genes result in the secretion of single granulin domain peptides (also add in the original carp paper) (Cadieux et al., 2005).

The mechanism by which a gene may undergo the expansion or contraction of internal protein motifs in an inheritable manner requires an unequal crossover event during meiosis.

4.1.2 Mechanisms of intragenic expansion and contraction – The number of inherited diseases that are associated with small and large genomic alterations and/or rearrangements is ever increasing. The most simplistic examples of genetic disease inheritance involves single nucleotide polymorphisms, such as those found within the *progranulin* and leading to FTD-U (Baker et al., 2006; Cruts et al., 2006). In contrast, chromosomal trisomy, segmental deletions and trinucleotide repeat expansions have all been associated with various forms of mental retardation (Jacobs et al., 1959; Lejeune et al., 1959; Shaw-Smith et al., 2006). However, subtle and gross alterations within the human genome have also been shown to be beneficial in nature, including haplotype inversions that confer increased fertility and SNPs that provide protection against myocardial infarction (Girelli et al., 2000; Stefansson et al., 2005).

Due to their phenotypic consequences, perhaps the most well documented genomic alterations are those that involve the expansion of intragenic trinucleotide

repeats, which are the eitiological factor in Huntington's disease, spinocerebellar ataxia and Kennedy's disease (Everett and Wood, 2004). The expansion of trinucleotide repeat series is the result numerous factors and can occur within meiotic and mitotic cells, both of which can lead to disease development (Pearson, 2003; Pearson et al., 2005). Although recombination errors and double strand breaks can expand/contract TNRs the most common TNRs are those derived from DNA replication errors, wherein the machinery stalls and the formation of short hairpin structures leads to their excision or insertion (Pearson et al., 2002). However, DNA slippage is typically constrained to introducing small aberrations such as TNRs, rather than the genesis of multi-exon/intron blocks that is necessary for the expansion of protein domains such as that observed in *progranulin* genes.

Although not necessarily associated with disease, the most common instances of genomic alterations that yield an expansion or contraction of encoded domains involves DNA transposons and retrotransposons (Kazazian, 2004; Miller and Capy, 2006; Prak and Kazazian, 2000; Reich et al., 2002; Slotkin and Martienssen, 2007). Both DNA transposons and retrotransposons are referred to as mobile genetic elements that are capable of integrating into different regions of a cells genome. DNA transposons involve the mobilization of a segment of genomic DNA and its subsequent integration within a new region, proximal to its parental location (Boyd and Hartl, 1997; Vigdal et al., 2002). The genesis of the zebrafish *progranulin-1* and *-2* genes are likely products of a DNA transposition event with their highly conserved architecture, expression patterns and more conclusively their association with *ASgrn1/2* which contains an ancient *tc1/mariner* DNA transposon (Cadieux et al., 2005). Although capable of expanding upon *progranulin* gene number it is difficult to consider DNA transposons as mediators of intragenic

expansion/contraction as the transposon itself would become integrated within the gene similar to that of *ASgrn1/2*. Similarly, retrotransposons are capable of integrating RNA sequences within the genome, such that the novel insertion typically consists of an intronless sequence together with a characteristic polyadenine stretch (Dewannieux et al., 2003). Although the integrated retrotransposon product is theoretically capable of encoding a protein its expression (if any) is dependent on the presence of an active promoter region 5' to the integration site. Such retrotransposition was likely involved in the formation of the sea urchin *progranulin* which harbours a 5' exon that consists of twenty-five intronless granulin domains (without a polyadenine stretch), followed by a single intron and 3' exon (Figure 6). Based on this architecture it would appear that retrotransposition of the *progranulin* RNA occurred within the parental *progranulin* gene and EST database mining accessions (CD320922, CX682418, CD304373, CX681063) suggests that this pseudogene is transcriptionally active, although conserved biological function has not been investigated. The sea urchin architecture is unique within the multitude of *progranulin* gene architectures where no other gene harbours intronless segments, therefore for this reason alone retrotranspositions can be discarded as a potential mechanism of general intragenic expansion/contraction.

There remains two potential mechanisms that can possibly lead to the generalized expansion and contraction of granulin domains; aberrant replication via DNA slippage or unequal crossover. The former has been previously mentioned and is well known for the expansion/contraction of intragenic triplet repeats, such as those associated with Huntington chorea and fragile X syndrome, however DNA slippage is typically involved with microsatellites rather than large protein domains (Jin et al., 2004; Li et al., 2004). Although DNA slippage during replication has been demonstrated to allow for the

multiple expansions of a ~100 nucleotide sequence in yeast, the observed multiplication of conserved granulin domains would require much longer segmental loops (Verstrepen et al., 2005). Although DNA slippage cannot be ruled out other mechanisms are also possible including unequal crossover between chromosome pairs or sister chromatids that will generate both an expansion and contraction (Tom Strachan, 2003). Indeed, unequal crossover occurs between homologous and closely related paralogous genes (Jelesko et al., 2004; Kadyk and Hartwell, 1992; Onoda et al., 2004). Furthermore, the reciprocal exchange of genomic between chromosomes can manifest as unique phenotypes due to over under-representation of specific genes or chromosome segments (Sinnott et al., 1990).

4.1.3 The molecular relevance of intragenic expansion and contraction – The number and type of domains within any given protein can be extremely diverse. For particular protein domains the molecular significance of domain multiplication or deletion can be quickly assessed, at least hypothetically, for those domains whose biological activities are well characterized such as kinase, phosphatase, shc adapters (SH2/3) or farnesylation sites (Eungdamrong and Iyengar, 2004a; Eungdamrong and Iyengar, 2004b; Pawson, 2007; Remenyi et al., 2006). The same hypothetical evaluation of the biological significance of the expansion/contraction of tandem granulin domains is experimentally possible when assessed within the contexts of downstream intracellular signalling pathways. Such experimentation may be able to define a minimum and maximum granulin reiteration number. The significance of such information is not purely academic in nature; rather the definition of a minimum repeat bioactivity for progranulin may have clinical applications. This stems from the fact that NMD is associated with the

haploinsufficiency of several genetic diseases, including FTD-U, cystic fibrosis and Duchenne muscular dystrophy (Baker et al., 2006; Clancy et al., 2001; Cruts et al., 2006; Roberts et al., 1994). Numerous inhibitors of NMD are currently in clinical trials for the treatment of genetic disease who's aetiology is dependent upon this pathway and for which the truncated protein is likely to harbour sufficient bioactivity as to relieve or cure the disease. Therefore, it is particularly interesting to determine if severe truncations of progranulin such as c.373C>T (Q125X) or IVS8+1G>A (V279GfsX4) that would encode one and three granulin motifs respectively, could be sufficient or equivalent bioactivity compare the full-length precursor. Several specific and large-scale studies have attempted to assign biological significance to alterations in the number of tandem protein repetitions and may provide general directions of experimental interest.

The ankyrin protein domain consists of 33 amino acids and is found within numerous genes and in various tandem repetitions, ranging from four in $p16^{INK4a}$ to seven in Notch (Breeden and Nasmyth, 1987; Tripp and Barrick, 2004). Individual ankyrin repeats form two anti-parallel α-helices that compact to form a repetitive array within multi-ankyrin repreat proteins (REF). This "*beaded string*" architecture would suggest that ankyrin repeats of *drosophila* Notch lack a globular structure and that the proper folding of each α-helical doublet is intrinsically programmed without the need of inter motif stabilization (Zweifel et al., 2003). However, the deletion of distal ankyrin motifs (N- or C-terminal) of *drosophila* Notch decreased the stability of adjacent repeats (Zweifel and Barrick, 2001; Zweifel et al., 2003). The deletion of internal ankyrin repeats within the *drosophila* Notch produced more deleterious effects in protein stability, suggesting that cooperative stabilization between adjacent repeats has become selective. In contrast, the duplication of ankyrin domains directly adjacent to the parental repeat had

no affect on stability (Zweifel and Barrick, 2001; Zweifel et al., 2003). The cooperative folding of the ankyrin repeats of *drosophila* Notch suggests that expansion is tolerable, whereas deletion can be deleterious with regards to protein half-life.

Similar to the ankyrin repeat architecture of *drosophila* Notch, progranulins are likely to adopt a "*beaded string*" architecture. Granulin motifs are *prima facie* constrained to adopt a highly constrained structure that is established by the six internal cystine-bridges and therefore it is easy to presuppose that the folding of each motif is largely predetermined and cooperative folding is reserved for protein-protein interactions. Additionally, the plasticity and propensity for progranulins to undergo internal multiplications/deletions argues against cooperative folding.

Despite the relevance of these points there is evidence to indicate that cooperative folding mechanistics may exist for tandem granulin arrays. First, although some alternative splice variants have been identified, they consist of truncations rather than products of exon-skipping products, in which the sequential architecture of the domains might be conserved (Figure 1). Second, familial FTD-U has been associated with missense mutations including Pro248Leu (c.743C>T), Ser258Asn (c.773G>A) and Arg432Cys (c.1294C>T) (van der Zee et al., 2007). Presumably, these mutations do not result in NMD-mediated haploinsufficiency and may affect protein folding. The Arg432Cys mutation could clearly result in misfolding through inappropriate cystine-bridging, whereas the significance of Pro248Leu is hypothesized as pathogenic based on coding conservation and *in silico* molecular modeling analysis (van der Zee et al., 2007). Recently, Haass and colleagues demonstrated that over expression of the Pro248Leu and Arg432Cys mutants in HeLa cells does not affect translation efficiency (Shankaran et al., 2008). Rather, these mutations were inefficiently secreted through the ER/Golgi system

suggesting that the mutation lead to targeted degradation. In contrast, the Ala9Asp mutatant was not expressed, suggesting that such misssorting to the cytoplasm is under surveillance by NMD. Based on these results some non-synonomous mutation of progranulin lead to haploinsufficiency through NMD or post-translational degradation. In addition, if granulin domains folded independently of their up and downstream counterparts, we would expect such mutations to harbour only localized deleterious folding effects, rather than perturbing the biological activity of the entire precursor. This suggests that the folding of the progranulin precursor may be a concerted event or the improper folding of a single motif is sufficient to alert the degradative mechanism. It will be important to ascertain the significance of all non-synonomous mutations *in vitro* as not all missense mutations have pathogenic associations. For instance, the missense polymorphisms Cys105Arg (c313T>C) and Gly515Ala (c.1544G>C) have been identified in control individuals (Gass et al., 2006).

Given the variation of the *progranulin* gene from one motif to multitudes of repeats, it is arguable that expansion/contraction is well tolerated. This concept of intragenic tolerance appears true in terms of evolutionary diversity. Previously, the concept of treating haploinsufficiency disease with NMD antagonists was referred to as a possible therapeutic option for FTD-U patients with *progranulin* mutations. Presuming that such NMD antagonist prove effective in chronic diseases, the underlying mutation remains, therefore the novel protein product is inevitably a truncated version of the original precursor. Therefore, what is the phenotypic manifestation, if any, of such truncation products?

4.1.4 The phenotypic relevance of intragenic expansion and contraction – Phenotypic consequences that may result from altering the number of internal repetitive protein domains within a gene is contingent upon devising a system that can assign a qualitative or quantitative significance to each construct. The yeast, *Saccharomyces cerevisiae* has been recently proven to be a worthy system to investigate such questions for two reasons (Verstrepen et al., 2005). First, it has been observed that domain repeat proteins harbour a high degree of repeat number variation between yeast strains, unlike their non-repeat counterparts, indicating that there is an active mechanism of inaccurate genomic replication which can be used to generate genetic variants from a single parental construct. Second, 75% of *Saccharomyces cerevisiae* genes that harbour tandem domain repeats are cell-surface proteins suggesting that alterations in domain repeat number will likely affect cell adhesion. Fink and colleagues took advantage of these *S. cerevisiae* properties to assess the genetic variants and associated phenotypic changes derived from the introduction of the cell-surface adhesin gene *FLO1* that contains 18 repeats of a ~100 nucleotide sequence (Verstrepen et al., 2005). Analysis of *FLO1* transformants illustrated that both expansion and contraction of domain repeat number were generated, however truncations were predominant. Furthermore, the adhesive properties of each *S. cerevisiae* strain increased in a linear proportion to the increase in *FLO1* encoded domain repeats (Verstrepen et al., 2005).

Clearly, utilizing a unicellular eukaryotic system to investigate phenotypic correlations with gene variants is relatively simple in comparison to initiating a similar investigation within a multicellular model such as C. elegans, the fruitfly or the zebrafish. Rather than initiate such an investigation, Fondon and Garner attempted to infer phenotypic correlations by cataloguing alterations in the domain repeat architectures of

developmentally important genes in 24 dog breeds (including two Hox clusters) in comparison to craniofacial and limb morphology (Fondon and Garner, 2004; Fondon and Garner, 2007). The connection between genotype and phenotype was most clearly described through the identification of a deletion encompassing the PQ_n repeat domain of Aristaless-like 4 (Alx-$4^{\Delta 51}$) found exclusively in the Great Pyrenees breed (Fondon and Garner, 2004). The homozygous phenotypic consequence of Alx-$4^{\Delta 51}$ allele manifests as hindlimb polydactyly, similar to the murine Alx-4 null phenotype (Qu et al., 1998). Of particular interest, the difference between Gln_n and Ala_n homopolymer repeats within the *Runx-2* gene correlates with overall variation in craniofacial morphology across dog breeds (Fondon and Garner, 2007).

Obviously, it would be interesting to analyze whether the rapid speciation of dog breeds has manifested as species-specific alterations within *progranulin* gene architecture and whether the progranulin genotype can be correlated to a phenotypic predisposition. As an alternative to direct investigation, the expansion of eukaryotic genome sequencing projects by the European Bioinformatic Institute and the Wellcome Trust Sanger Institute are readily available for analysis through the National Center for Biotechnology Information (NCBI) and the Ensembl genome browser (Hubbard et al., 2007; National Center for Biotechnology Information (U.S.), 2002).

4.1.5 Recent alterations in mammalian progranulin genes – The variation of granulin domain number between the zebrafish co-orthologues *progranulin-a* and *–b* results from relatively recent evolutionary events following their appearance ~350mya. This is in keeping with the assumption that molecular variation is an acceptable occurrence provided that at least one orthologue retains the general biological activity of the common

ancestral gene. One would predict that, due to the representation of *progranulin* as a single gene, mammalian species would not allow similar variation to occur. Indeed, of those species whose genome sequencing is sufficiently advanced, the majority of mammals including the common rabbit (*Oryctolagus cuniculus*), mouse (*mus musculus*), chimpanzee (*Pan troglodytes*) and the marsupial possum (*Monodelphis domestica*) all harbour a *progranulin* gene with near-identical architecture (Figure 7). Surprisingly, the dog (*canis familiaris*) and the macaque monkey (*macaca mulatta*) do not share the common mammalian architecture. The dog harbours a *progranulin* gene that encodes for six and one half granulin domains where the tenth coding exon is absent based on genomic sequencing (Figure 7). The macaque harbours a *progranulin* gene that contains a truncated ninth coding exon, yielding an internal C-terminal half domain. At present, these *in silico* architectures have yet to be confirmed from available EST libraries.

Although the identification of mammalian species with architectural variations within the *progranulin* gene is worthwhile, this information should not be considered as an end in and of itself. Rather, variations may represent species-specific functionalizations or the redundancy of the bioactive moiety within progranulin that may be inherent in a protein composed of tandem repeat domains. The underlying question is, what is the biologically relevant segment or segments and does the variation in the number of coding domains affect its biological activity? To answer such structure/function questions one typically relies on a quantititave or qualitative assay similar to investigation of the *FLO1* gene in yeast and the genotype/phenotype inference in dogs by Fondon and Gerner (Fondon and Garner, 2004; Verstrepen et al., 2005). However, the validation of identified changes in gene architecture relied on *prima facie* knowledge of the molecular consequences of repeat contraction/exanpsion. Simply put,

an understanding of the general signalling mechanism associated with progranulin would prove the most straightforward system to screen progranulin constructs of variable repeat number. The elucidation of the progranulin signalling cascade has clearly defined downstream targets, but the identification of direct protein-protein interactions (i.e. binding proteins and/or cell surface receptors) that mediate these events.

4.2 Progranulin; Diverse affects of gene abrogation in various species

4.2.1 FTD-U animal models; Making sense of dementia through missense – Animal models of human diseases are tools with which the research community continuously strives to develop and validate. The ongoing crux of any animal model for whatever disease (or developmental pathway) is the degree to which human phenotypes and cellular abnormalities are faithfully recreated. These concerns become particularly pronounced for neurodegenerative diseases that develop over several years or decades, far surpassing the normal life cycle of small animal models. Despite this significant caveat a number of mouse models of human neurodegenerative disorders have been developed including Alzheimer's disease, amyotrophic lateral sclerosis, Parkinson's disease, FTD and Huntington's disease (Denk and Wade-Martins, 2007; Hickey and Chesselet, 2003; Julien and Kriz, 2006; Levine et al., 2004; McGowan et al., 2006; Melrose et al., 2006).

The first step in modeling *progranulin* associated FTD-U in an animal model would be the generation of a heterozygous genotype with knockout technology (Beglopoulos and Shen, 2004). Presuming that the development of FTD-U is strictly associated with *progranulin* haploinsufficiency, the mouse *progranulin*$^{+/-}$ genotype should represent the equivalent outcome but without the involvement of NMD (as defined earlier, Section 1.1.1). However, the *progranulin* haploinsufficiency does not manifest

clinically until 59+/-7 years of age and the penetrance of some mutations that activate NMD are incomplete (~90%) (Cruts et al., 2006; Gass et al., 2006). Based on this information it is unclear whether *progranulin*$^{+/-}$ mice would recapitulate the disease prior to natural age-associated neurodegeneration. This presumed lack of penetrance for the mouse *progranulin*$^{+/-}$ genotype would likely result from the lengthy incubation period of the disease in humans as well as the variability of the age of onset (Gass et al., 2006). Similar penetrance problems have been reported for mouse models of FTD that were generated through microtubule associated protein tau (MAPT) knockout (Dawson et al., 2001; Harada et al., 1994).

Presuming the issue of penetrance to be the primary hurdle in generating a mouse model of *progranulin* associated FTD-U, a system of cumulative genetic dysfunction may provide solutions. This concept is reflected in the human superoxide dismutase G93A (SOD^{G93A}) over-expressing mice that are the mainstream *in vivo* model for the investigation of familial amyotrophic lateral sclerosis (ALS) (Gurney et al., 1994). Currently two transgenic strains of SOD^{G93A} mice are used in ALS research that differ in copy number (8 versus 25 copies) and life-span (236 versus 129 days) (Alexander et al., 2004; Gurney et al., 1994). To date it remains uncertain as to whether SOD^{G93A} can be considered a loss or gain-of function allele, but it is clear that in familial ALS and transgenic mouse models the SOD protein is found within intracellular inclusions suggesting that part of the SOD^{G93A} aetiology may be associated with protein misfolding and aggregation.

Not all *progranulin* mutations are believed to cause haploinsufficiency through NMD. Specifically, it is not known whether those *progranulin* missense mutations that cause non-synonomous changes within the coding sequence of *progranulin* result in the

activation of the NMD (or similar) pathway leading to haploinsufficiency or, if the affected allele is translated and an inert product is generated through protein misfolding. As expected for those *progranulin* mutations that cause NMD activation and haploinsufficiency, neuronal inclusion do not contain progranulin protein (Baker et al., 2006). Interestingly, those *progranulin* missense mutations that are associated with FTD-U have not been assessed for progranulin aggregates. If missense mutations within *progranulin* indeed cause misfolding or aggregation then transgenic over-expression of multiple copies of the missense allele in mice may recapitulate some aspects of progranulin-associated FTD-U. This concept is similar to that of the SOD^{G93A} over-expressing mice previously described.

Although a system that can be manipulated with respect to overall penetrance and age of onset through the increase in disease-related allele copy number would be desirable, there remains the possibility of the confounding effects of introducing molecular or clinical manifestations outside that of FTD-U.

4.2.2 The validity and penetrance of progranulin mutations in neurodegenerative disorders - Despite significant advances in the understanding of progranulin action, studies aimed at elucidating the physiological functions of progranulin *in vivo*, in particular during embryogenesis and development, have been lacking. Recent studies have reported that the addition of exogenous progranulin or a progranulin function-blocking antibody results in accelerated or arrested growth of murine pre-implantation embryos, respectively (Diaz-Cueto et al., 2000). Additionally, use of the progranulin function-blocking antibody inhibited blastocyst adhesion and outgrowth during peri-

implantation *in vitro* (Qin et al., 2005b). Based on these results it is likely that a murine progranulin knockout will lead to embryonic lethality prior to or during implantation.

The recent identification of mutations within human *progranulin* that lead to FTD-U has prompted the re-evaluation of familial dementia patients for similar novel *progranulin* mutations. At present, those *progranulin* mutations that have been clearly demonstrated to result in haploinsufficiency, through activation of the NMD pathway, are commonly referred to as autosomal dominant. We can presume that a single allele is sufficient to meet the needs for *progranulin* expression during development and the majority of adulthood. Although these *progranulin* mutations are rightly characterized as autosomal dominant for several families, this generality may be misleading since some mutations are not completely penetrant (Cruts et al., 2006; Gass et al., 2006). Van Broeckhoven and colleagues identified the IVS0+5G>C mutation located within the first intron that is presumed to result in retention of this intron during maturation and activation of a nuclear degradation pathway (Vinciguerra and Stutz, 2004). Haploinsufficiency related to the IVS0+5G>C *progranulin* mutation was confirmed for those with symptomatic FTD-U. Curiously, two deceased individuals who were obligate carriers of the diseased allele died within the broad onset range without presenting clinical FTD-U. Despite attempts to identified genetic modifiers that may affect disease penetrance, such as *APOE* genotype and H1/H2 haplotype, no co-segregation between these markers and *progranulin* mutations has been identified thus far. Based on the questionable penetrance of some mutations, the use of the "autosomal dominant" descriptor should be used with care. In addition, *progranulin* mutations have been described in published material and within this text as the etiological factor that underlies FTD-U. It remains to be determined whether *progranulin* haploinsufficiency is sufficient

to evoke neurodegeneration or whether this genotype is an enhancer of other genotypic manifestations and/or environmental insults.

It would appear that *progranulin* haploinsufficiency alone is insufficient to manifest as FTD-U based on the highly variable penetrance of familial FTD-U patients that have been genotyped for *progranulin* mutations (Bronner et al., 2007; Bruni et al., 2007). This may reflect the presence of compensatory mechanisms, such as the increased expression of full-length progranulin from the unaffected allele. Interestingly, for those extended families carrying mutant *progranulin* alleles, there has been no report of the identification of individuals homozygous for a deleterious mutation (Bronner et al., 2007; Bruni et al., 2007). If nothing else, this fact highlights the presumption that complete loss of progranulin function is presumed to be early embryonic lethal.

4.2.3 Recent progress in progranulin biology – The preliminary phenotypic consequences of progranulin-B abrogation are noteworthy and further investigation would likely revolve around the analysis of overall cell cycle progression, apoptosis, spatiotemporal patterning of the three germ layers and early dorso-ventral molecular distinctions. In addition, a second independent morpholino engineered to inactivate *progranulin-b* that produces the same phenotype is typically necessary to further validate the results as well as mRNA "rescue" experiments. Recently, Haass and colleagues described the inactivation of zebrafish progranulin-B using a MO targeting the exon6/intron7 splice site (Shankaran et al., 2008). The use of this *progranulin-b* splice-site MO had no affect on zebrafish development. Although Hass and colleagues demonstrated that the MO in question affected *progranulin-b* transcription, the resultant RT-PCR products argue for the introduction of an alternative splice-site rather than

abrogation. Furthermore, there is no Western blot analysis to demonstrate that the protein product is affected.

Recently, Nishihara and collegues published phenotyptic data regarding the mouse *progranulin* knockout animal (Kayasuga et al., 2007). The affects of murine progranulin knowckout are subtle and associated with behavioural alterations, specifically copulation and aggression. There was no affect on the Mendelian heritance or litter size during breeding of the progranulin$^{+/-}$ or *progranulin*$^{-/-}$. Given the mean age of FTLD-U onset in humans is 59+/-7 years of age and the penetrance of some mutations that activate NMD are incomplete (~90%) it should not be surprising that such a clinical manifestion is not evident in an animal model whose life-span is ~2 years (Cruts et al., 2006; Gass et al., 2006). However, the lack of overt defects in the *progranulin*$^{-/-}$ background is surprising. Presumably, a human with such a genetic background would have a clinical manifestation that is much more severe that the FTLD-U associated with *progranulin* heterozygosity. Furthermore, these results are in stark contrast to the developmental roles attributed to the zebrafish *progranulin-a* and *–b* that have been documented herein. This reiterates the need to complete the phenotypic validation of our zebrafish results using mRNA "rescue". It should be highlighted that, although the zebrafish is a worthwhile comparative model of vertebrate biology, there remains the possibility that speciation has placed further burden on the zebrafish *progranulin* family during development. Several examples of such differences between zebrafish and other vertebrate models including the zebrafish insulin-like growth factor binding protein-2 (IGFBP-2) and Hey-2 (Gessler et al., 2002; Rutenberg et al., 2006; Wood et al., 2005; Wood et al., 2000). Inactivation of zebrafish IGFBP-2 results in developmental delay, reduced body growth and cardiovascular abnormalities whereas the murine knockout of the orthologous gene has

no gross clinical defects apart from a modest decrease in spleen and liver weight. In contrast, inactivation of zebrafish Hey-2 (the *gridlock* mutant) causes expansion of the venous endothelial population at the expense of arterial counterparts. The knockout of mouse Hey-2 results in post-natal cardiac hypertrophy, a direct reversal of the phenotype predicted by the zebrafish.

4.3 Morpholino inhibition of zebrafish Progranulin-A

4.3.1 Preamble – Unlike *progranulin-b*, its orthologous counterpart *progranulin-a* is expressed at low levels throughout gastrulation and segmentation with initial tissue-specific expression beginning during late somitogenesis (~18hpf) within the developing eyes and central nervous system (Cadieux et al., 2005). Commencing at 24hpf and maintaining a relatively consistent expression pattern throughout the rest of development, progranulin-a is expressed within several tissues including the pharyngeal endoderm and presumptive hematopoietic stem cells as well leukocytic lineages. In addition, progranulin-a is expressed within a large proportion of the developing epithelial lining of the visceral digestive system. Apart from the initial variation during gastrulation and somitogenesis (6-24hpf) between *progranulin-a* and *–b*, the most distinguishing divergence in the expression pattern between these two co-orthologous is the prevalence of *progranulin-b* as the major family member expressed within neural tissue.

4.3.2 Morpholino efficacy in progranulin-a knockdown – In this particular case one of three *progranulin-a* targeted MOs proved effective in translation inhibition in vivo and resulted in observable defects (Fig.9 and data not shown). The associated 5 base-pair mismatch proved less effective in translation inhibition and this decreased efficacy

was reflected in our analysis of the statistical data. As mismatch MOs are traditionally used as purely negative controls, unlike the way the mismatch MO in this study has been used, the modulation of the expression of a similar endogenous transcript, namely the *progranulin-b* orthologue, was used to validate MO efficacy. Neither the *progranulin-a* MO nor its mismatch affected alterations in progranulin-B levels (Figure 9A), indicating that the phenotypic results obtained do, in fact, reflect phenotypic consequences of progranulin-A abrogation.

Rarely is the use of a single MO-derived phenotypic consequence considered sufficient to reliably interpret the results as specifically associated with the targeted transcript. For this reason a series of three non-overlapping MOs targeting the *progranulin-a* 5'UTR were assessed for their ability to manifest a common phenotype and determine their efficacy for translation inhibition. As previously described only one progranulin-a targeted MO successfully produced a phenotype and actively inhibited translation (Figure 9). As an alternative to translation inhibition, MOs can also be used to force inaccurate processing of pre-mRNA, resulting in the improper excision of the targeted exon or the inclusion of the associated intron (Draper et al., 2001; Kramer-Zucker et al., 2005; Thummel et al., 2006). Targeting of *progranulin*-a or any other repeat domain protein with splice-site directed MOs is inherently an imperfect strategy since the loss of one exon may allow for translation and may yield a biologically viable truncated protein. In addition, the phase of the exon/intron boundary to be targeted must be compared to the foremost downstream boundary since the excision of the targeted intron may allow for the continued in-frame translation of the remaining transcript. Lastly, evolutionary arguments suggest that the progranulin genes undergo a large degree of internal plasticity, indicating that only a few tandem domains may be required for

biological function. Therefore, it is important to target splice-sites that will introduce a premature stop codon while limiting the number of granulin domains potentially expressed, ideally none. To this end a splice-MO (Sp-MO) targeting the fourth exon/intron boundary of *progranulin-a* was used (Figure 10A). The microinjection of Sp-MO at 10ng/embryo faithfully recreated the overall visible morphological defects including craniofacial dysmorphogenesis, gut dysplasia and pericardial edema (Figure 10B). However, assessment of the presumed alteration in *progranulin*-a transcript size was not possible, since the morphants do not appear to express progranulin-a (Figure 10C). The lack of progranulin-a expression in Sp-MO morphants is most likely due to the inclusion of intron four within the mature transcript thereby introducing an early termination codon, leading to activation of the zebrafish NMD pathway and subsequent degradation of the transcript. Similar results have been observed for splice MOs targeting the zebrafish lymphoid enhancer binding factor-1 (Lef1) transcription factor and erythroid specific protein 4.1R, which are involved in hypothalamic development and red blood cell integrity, respectively (Lee et al., 2006; Shafizadeh et al., 2002). Prior to the obtaining of these findings, the use of splice-site targeting MOs to introduce early termination codons had been considered as a potential method of truncating the *progranulin-a* transcript at various positions with the intention of identifying which granulin domains or domain combinations harboured *progranulin-a* bioactivity. Based on the results of the Sp-MO it is likely that any early termination sequence regardless of position, will likely induce the degradation of the progranulin-a transcript. Despite the presumed inability to produce multiple truncated progranulin-A proteins, the activation of NMD could be advantageous for other experimental avenues, particularly one that involves the inactivation of NMD.

4.3.3 Zebrafish as an animal model of NMD inhibitor screening – Although the zebrafish is most commonly employed as a model of vertebrate development, it has also gained popularity in *in vivo* high-throughput small molecule screens to discover novel biologically active compounds (Murphey et al., 2006; Murphey and Zon, 2006; Stern and Zon, 2003). The advantages of using the zebrafish for such screens are similar to the reasons this species has gained popularity in developmental studies namely, *ex utero* development, optical clarity, small size and large breeding capacity conducive to 96- and 384-well assays and the ability to treat whole organisms with waterborne molecules. Furthermore, screening *in vivo* allows for the assessment of tissue-specificity and toxicity of the molecule in question. Coupled with the growing number of transgenic zebrafish models a rapid and automated reporter system can be envisioned. For example, Tsang and colleagues utilized two transgenic enhanced green fluorescent protein (EGFP) reporter systems to assess alterations in the activation of fibroblast growth factor (FGF) or the vascular endothelial growth factor (VEGF) signalling pathways in a small molecule screen of tyrosine kinase inhibitors (Molina et al., 2007). The zebrafish mutant *crash&burn* harbours mutations within the cell cycle gene *bmyb* and results in cell cycle arrest and genome instability (Shepard et al., 2005). Using terminal-deoxytransferase-mediated nick end labelling (TUNEL) to assess apoptosis, Stern and Zon screened ~16,000 molecules within a 16 week period and identified a single compound capable of rescuing the *crash&burn* phenotype (Stern et al., 2005).

Zebrafish small molecule screens of medium to high-throughput are obviously dependent on obtaining large cohorts, a demand which is not altogether suited to screening morphants which require individual microinjection. To date small molecule

screens in the zebrafish are reliant on readily available mutants, such as *crash&burn* or transgenic animals that employ the over expression of specific gene constructs. With respect to NMD, two cell culture reporter systems have been developed using luciferase and EGFP chimeric constructs with known NMD targets (Boelz et al., 2006; Paillusson et al., 2005). There is little doubt that a similar construct could be used to produce a transgenic zebrafish harbouring the *progranulin-a* transcript with an early termination codon (or a human progranulin equivalent) and a 3'-EGFP sequence. An important caveat to such experimental design is the targeting of the NMD pathway during development, since inactivation of this mechanism may have deleterious effects. Indeed, knockout of the NMD regulator gene Upf1 is embryonic lethal in mice (Medghalchi et al., 2001). In addition, NMD has been shown to affect the immune response and offer protection against lipid cytotoxicity (Frischmeyer and Dietz, 1999; Li and Wilkinson, 1998; Sun et al., 2001). Despite these potential problems several clinical trials have been undertaken to assess the efficacy of the known NMD inhibitor gentamicin for the treatment of genetic diseases that involve aberrant early termination codons including cystic fibrosis and Duchenne muscular dystrophy (Politano et al., 2003; Wagner et al., 2001; Wilschanski et al., 2000; Wilschanski et al., 2003). Although the efficacy of gentamicin treatment was found to be highly variable with respect to disease treatment and translation of the affected gene product, these early findings may suggest that gentamicin derivatives may provide a starting point for developing and screening for more efficacious compounds.

3.3.4 Progranulin processing and the bioactive core – In addition to the expansion in gene number, the zebrafish utilizes other mechanisms to diversify gene expression and/or

function. This includes the expression of antisense transcripts for *progranulin-a*, *progranulin-1 and –2* as well as an unusual trans-splicing mechanism to generate a hybrid form of *progranulin1/2*. The existence of splice variants for *progranulin-a* has recently come to our attention. Exon skipping and alternate exon/intron boundaries appear to exist, since the genomic sequence for *progranulin-a* appears to encode more granulin repeats than are evident in isolated cDNA sequences (data not shown). Splice variance may explain the presentation of progranulin-A as two ~55 kDa proteins roughly half the theoretical size of the full-length precursor. A similar apparent discrepancy is evident for progranulin-B (Figure 9). Alternatively, progranulin-A may undergo post-translational processing similar to the extracellular proteolysis of mammalian progranulin by neutrophil-derived elastase (Zhu et al., 2002). In zebrafish the expression of progranulin-A and –B occurs prior to and does not overlap that of elastase-a or –b, indicating an alternative protease may be responsible for progranulin processing (Mudumana et al., 2004; Wan et al., 2006). The exact sequence of these progranulin-A and –B immunoreactive species remains to be determined since efforts to manipulate endogenous progranulin species by immunoprecipitation form zebrafish extracts have been unsuccessful thus far.

4.4 Progranulin and hematopoiesis

4.4.1 Progranulin-A knockdown and hematopoietic alterations – The expression of *progranulin-a* within the ICM occurs relatively late (24 hpf) compared to other hematopoietic markers such as embryonic *globins*, *gata-1* and *scl*, precluding its involvement in hemangioblast formation (~11 hpf) or initial HSC specification (11−18 hpf) (Amatruda and Zon, 1999). As the embryo matures, *progranulin-a* expression

increases progressively within the dorsal aorta, posterior ICM, and peripheral neutrophils and macrophages (Cadieux et al., 2005). This pattern of *progranulin-a* expression within definitive HSCs and peripheral leukocytes continues throughout all developmental stages analyzed (24−120 hpf). Preliminary results obtained with the progranulin-A knockdown morphant highlight the functional requirement of this growth factor during HSC development.

As expected the primitive HSC progenitors (which form prior to 24hpf) are unaffected in 83% (n=19/23) of progranulin-A morphants (Figure 11A). However, the definitive HSC population within the ICM appears absent in 83% (n=35/41, Figure 11B), as do myelo-erythroid progenitors in 67% (n=12/18, Figure 11C) of progranulin-A morphants. In conjunction with the absence of HSCs, downstream markers of macrophages, neutrophils and erythrocytes are substantially decreased (Figure 11D,F). Based on these results, we suggest that progranulin-A is required for the maintenance of definitive HSCs (Figure 11G). Whether this is the result of alterations in a common multipotent progenitor, a distinct progranulin-A requirement in each lineage, or both is currently unknown. It is worthwhile investigating the expression patterns of other hematopoietic genes such as *c-myb*, *runx1* and *Pu.1* to further characterize the progranulin-A morphant phenotype.

4.4.2 The functions of mammalian progranulin during hematopoiesis and the innate immune response – Progranulin undergoes intracellular processing in macrophages and neutrophils to yield individual granulin domains (Bateman et al., 1990; Zhu et al., 2002). During an inflammatory response these innate immune cells release granulin peptides that cause decreased epithelial growth and promote further neutrophil recruitment (Zhu et al.,

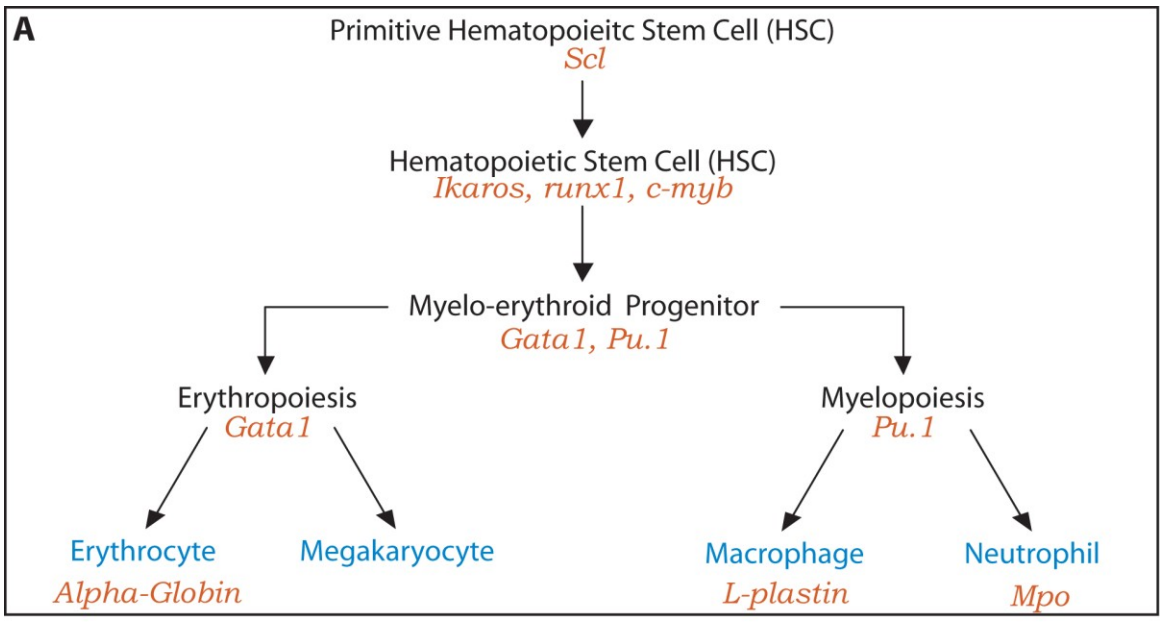
2002). In contrast, endothelial cells and fibroblasts secrete full-length progranulin following experiment wound. Progranulin derived from these cells enhances re-epithelialization, angiogenesis and suppresses neutrophil degranulation (He et al., 2003; Zhu et al., 2002). Furthermore, recombinant progranulin applied to murine transcutaneous puncture wounds has been shown to enhance healing *in vivo* (He et al., 2003).

Apart from their involvement in the innate immune response, progranulin and its constituent granulin peptides have been implicated in hematopoietic differentiation in both normal and pathological settings. For instance, granulins induce the proliferation of goldfish myeloid progenitors (Hanington et al., 2006). Although progranulin is highly expressed in mammalian hematopoietic stem cells (HSCs), its functional relevance is not understood (Golan-Mashiach et al., 2005). In multiple myeloma (a clonal B-cell expansion) progranulin acts as an autocrine growth factor via mitogen-activated protein kinase (MAPK) and phosphatidylinositol 3'-kinase (PI3K) pathways (Wang et al., 2003). Furthermore, progranulin over-expression in multiple myeloma cells confers resistance to dexamethasone, the conventional drug used to treat this disease (Wang et al., 2006). In general it is not clear whether progranulin is necessary for overall hematopoietic cell proliferation or if its function is more subtle with respect to the maturation of the various lineages.

4.4.3 Discussion of hematopoietic defects; expanding on preliminary data – Defining the function progranulin-A during zebrafish hematopoiesis can be extended to investigating the hierarchical position of progranulin-A in comparison to other genes with defined hematopoietic functions. For instance, MOs targeting the stem cell leukemia

(Scl) transcription factor result in almost complete loss of the erythroid and myeloid lineages (Figure 22) [64]. Downstream regulators include *Pu.1* and *Gata-1* which are necessary for myeloid and erythroid differentiation, respectively [65]. It would be worthwhile to investigate the expression pattern of *progranulin-a* in the context of *Scl*, *Gata-1* and *Pu.1* morphant backgrounds. Importantly, the efficacy and phenotypic consequence of each such morphant (Scl, Pu.1 and Gata-1 morphants) have been previously documented. Therefore, the affected downstream lineages within Scl, Pu.1 and Gata-1 morphants can be assessed following *progranulin-a* inactivation using MOs and over-expression by mRNA microinjection extent to which progranulin-A over-expression can rescue downstream cell populations.

As previously mentioned the regulatory pathways governing hematopoiesis are generally conserved between mammals and the zebrafish (Section 3.3.1). With this in mind it is possible to envision a series of *in vitro* experiments using commercially available human HSCs to determine progranulin function in this context (Figure 23). This experimental design would rely upon the use of expression patterns of genes orthologous to those in the zebrafish as reporters. First, assessment of the expression


A
Primitive Hematopoieitc Stem Cell (HSC)
Scl
Hematopoietic Stem Cell (HSC)
Ikaros, runx1, c-myb
Myelo-erythroid Progenitor
Gata1, Pu.1
Erythropoiesis
Gata1
Myelopoiesis
Pu.1
Erythrocyte
Alpha-Globin
Megakaryocyte
Macrophage
L-plastin
Neutrophil
Mpo

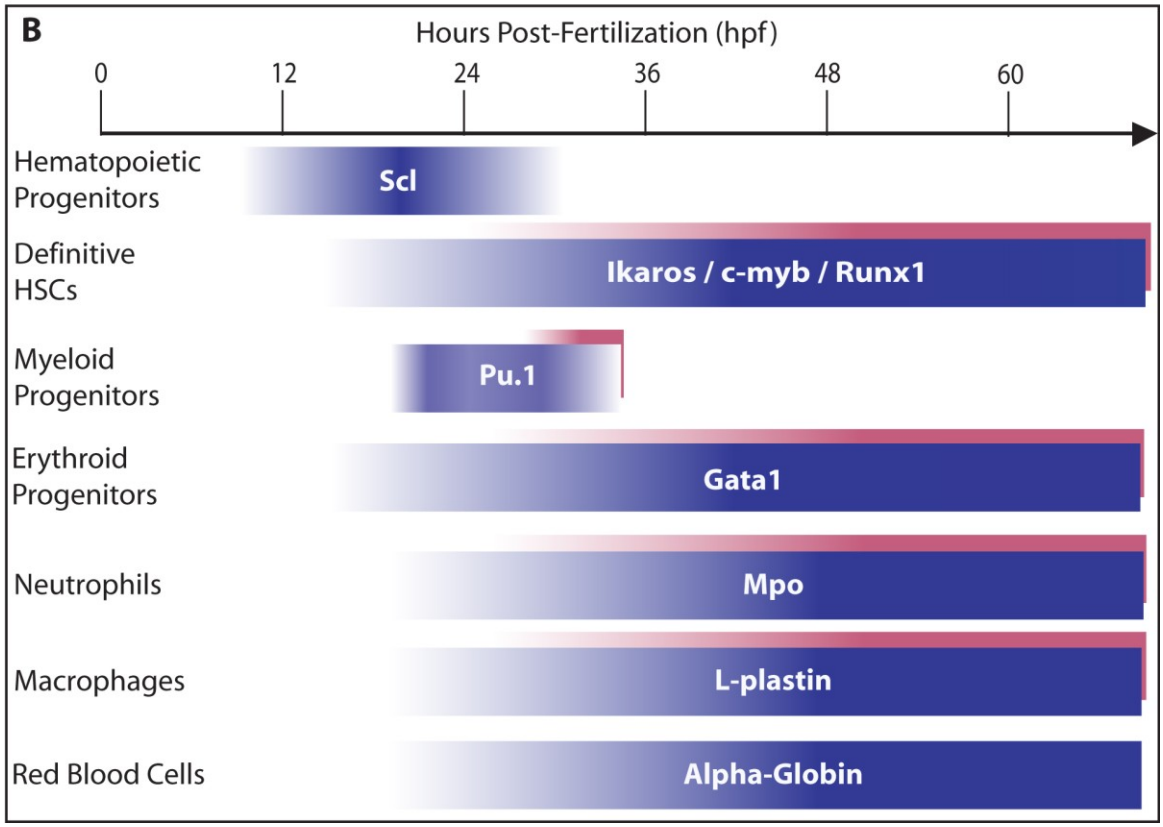
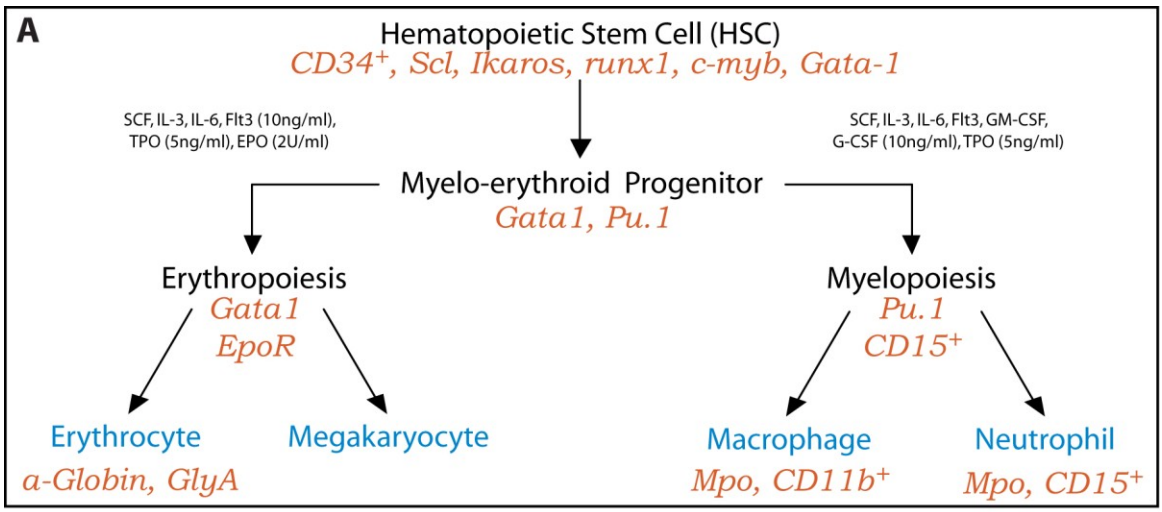

B
Hours Post-Fertilization (hpf)
0
12
24
36
48
60
Hematopoietic Progenitors
Scl
Definitive HSCs
Ikaros / c-myb / Runx1
Myeloid Progenitors
Pu.1
Erythroid Progenitors
Gata1
Neutrophils
Mpo
Macrophages
L-plastin
Red Blood Cells
Alpha-Globin

Figure 22. Proposed experimental design for the continued investigation of zebrafish progranulin-A morphant hematopoietic defects. (A) Genetic infrastructure of definitive hematopoiesis in the zebrafish. Myeloid and erythroid differentiation is depicted with various expression markers for each lineage in red (Bennett et al., 2001a; Brownlie et al., 2003; Kalev-Zylinska et al., 2002; Liao et al., 1998; Rhodes et al., 2005; Thompson et al., 1998; Willett et al., 2001). (B) Temporal gene expression pattern during zebrafish definitive hematopoiesis. Expression patterns related to the intermediate cell mass (ICM) and circulating derivatives are illustrated (purple), together with the known temporal and regional pattern of *progranulin-a* expression (red). Progranulin-a is found within HSC and all differentiated cell types except red blood cells (Bennett et al., 2001a; Brownlie et al., 2003; Kalev-Zylinska et al., 2002; Liao et al., 1998; Rhodes et al., 2005; Thompson et al., 1998; Willett et al., 2001).


A
Hematopoietic Stem Cell (HSC)
CD34+, Scl, Ikaros, runx1, c-myb, Gata-1
SCF, IL-3, IL-6, Flt3 (10ng/ml), TPO (5ng/ml), EPO (2U/ml)
SCF, IL-3, IL-6, Flt3, GM-CSF, G-CSF (10ng/ml), TPO (5ng/ml)
Myelo-erythroid Progenitor
Gata1, Pu.1
Erythropoiesis
Gata1
EpoR
Myelopoiesis
Pu.1
CD15+
Erythrocyte
a-Globin, GlyA
Megakaryocyte
Macrophage
Mpo, CD11b+
Neutrophil
Mpo, CD15+

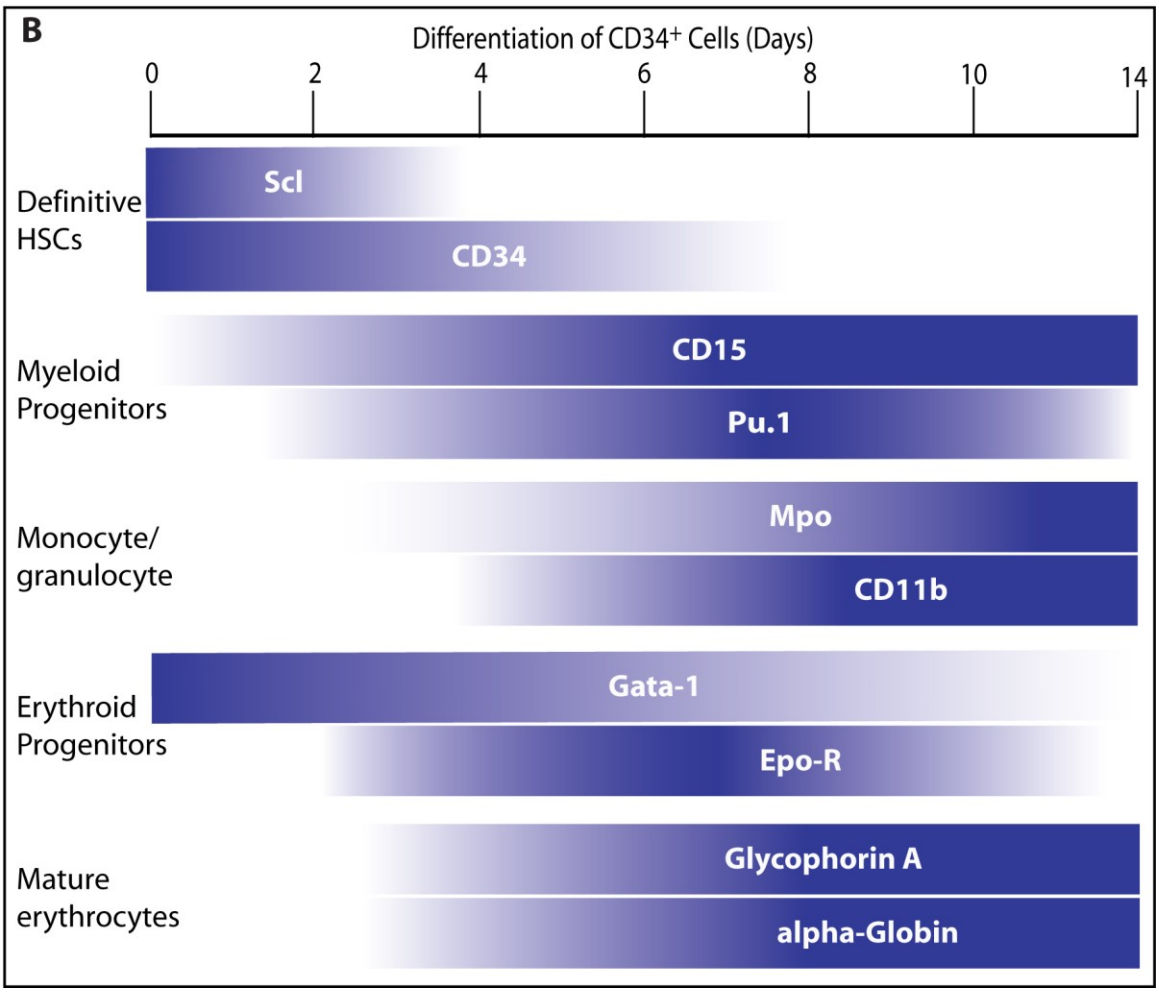

B
Differentiation of CD34+ Cells (Days)
0
2
4
6
8
10
14
Definitive HSCs
Scl
CD34
Myeloid Progenitors
CD15
Pu.1
Monocyte/ granulocyte
Mpo
CD11b
Erythroid Progenitors
Gata-1
Epo-R
Mature erythrocytes
Glycophorin A
alpha-Globin

Figure 23. Proposed experimental design for the investigation human progranulin function during hematopoietic differentiation. (A) Genetic infrastructure of the *in vitro* differentiation human $CD34^+$ HSCs. Myeloid and erythroid differentiation is described following application of a combination of specific cytokines with various expression markers for each lineage in red (Balducci et al., 2003; Cheng et al., 1996; Edvardsson et al., 2004). (B) Putative temporal gene expression pattern during *in vitro* differentiation human $CD34^+$ HSCs based on published data (Balducci et al., 2003; Cheng et al., 1996; Edvardsson et al., 2004).

pattern of progranulin in comparison to known expression markers during the *in vitro* differentiation of HSCs along the myeloid and erythroid lineages would be necessary to determine the context(s) in which progranulin may play a role (Balducci et al., 2003; Cheng et al., 1996; Edvardsson et al., 2004). The more challenging experimental analysis will involve the alteration of progranulin bioavailability during differentiation, through application of exogenous recombinant protein and the use of RNA interference protocols. With regards to the latter, a few studies have been reported where HSC transfection has been undertaken the most efficient being the use of lentiviral transfection (Reynaud et al., 2005; Scherr et al., 2006; Verhoeyen et al., 2005).

4.5 The many faces of progranulin

4.5.1 Discussion of craniofacial morphogenesis – The ontogeny of *progranulin-a* expression originally highlighted this growth factor as the sole *progranulin* family member to be expressed within the posterior intermediate cell mass (ICM) as well as peripheral leukocytes (Cadieux et al., 2005). In addition *progranulin-a* is expressed within the pharyngeal endoderm between 24 and 48 hpf. Therefore, alterations in craniofacial morphogenesis were not unexpected in progranulin-A morphants since this tissue has become increasingly implicated in viscerocranial arch development (Crump et al., 2004; Piotrowski and Nusslein-Volhard, 2000). Cranial neural crest (NC) populations 1, 2 and 3 originate from the anteroposterior polarization of hindbrain within the midbrain-hindbrain boundary, rhombomere 1 (R1) and R2 domains, the preotic R4 and the postotic R6-7, respectively (Kontges and Lumsden, 1996; Schilling and Kimmel, 1994). In addition, formation of cranial NC cell streams 1-3 and initial ventrolateral migration is dependent on hindbrain patterning mechanisms without the involvement of

endodermal cues (David et al., 2002). In contrast, the segmentation of of the NC cells to form the five posterior ectomesenchymal arches (EMAs) is intimately associated with the coincidental segmentation of the endoderm to form pouches (Piotrowski and Nusslein-Volhard, 2000). In keeping with the onset of *progranulin-a* expression within the pharyngeal endoderm at ~24 hpf (Cadieux et al., 2005), the specification of cranial NC streams 1-3 from the hindbrain are unaffected (Figure 13D,J). In addition EMA-3 undergoes segmentation to yield one additional arch (arch 4), however continued segmentation does not occur (Figure 13E,F,K,L). In tandem with the discontinued posterior EMA segmentation the pharyngeal endoderm likewise fails to develop posterior pouches (Figure 14D,E). Interestingly, the pharyngeal endoderm in progranulin-A morphants remains localized as two bilateral streams without indication of lateral morphogenesis to form posterior pharyngeal pouches (Fig. 14J), whereas the posterior neural crest are no longer visible (Fig. 14L). The apparent absence of posterior EMAs suggested that these segments may have undergone apoptosis. However, TUNEL based detection of DNA fragmentation at 34, 48 and 72 hpf did not localize such a mechanism within the craniofacial arches thus the fate of the posterior EMAs is unknown (Fig. 14F, data not shown). Based on these observations it was concluded that the progranulin-A is primarily necessary for endodermal pouch formation and by extension, causes the disruption of cranial NC neural segmentation and maintenance.

Proper morphogenesis of the pharyngeal endoderm has been attributed to FGF (David et al., 2002; Nissen et al., 2003; Walshe and Mason, 2003) and retinoic acid (Kopinke et al., 2006; Niederreither et al., 2003; Wendling et al., 2001) signaling as well as the activity of the T-box transcription factor *tbx1* (Arnold et al., 2006; Piotrowski et al., 2003). The down-regulation of FGF signaling through ablation of both *fgf3* and *fgf8*

demonstrates that these growth factors are required for the formation of all endodermal pouches and that proper hindbrain segmentation is required for first and second pouch segmentation, whereas posterior pouch segmentation occurs independently of the hindbrain (Crump et al., 2004). Like the progranulin-A morphants the zebrafish *tbx1* mutant *van gogh* (*vgo*) harbours defects encompassing pouches 3-7 and associated EMAs (Piotrowski et al., 2003; Piotrowski and Nusslein-Volhard, 2000). The same posterior arch defects are produced by temporal antagonism of retinoic acid (RA) signalling between 16 and 36 hpf (Kopinke et al., 2006). This suggests that RA, *tbx1* and *progranulin-a* are regulators of posterior pharyngeal endoderm morphogenesis.

There is little published evidence regarding the function and regulation of mammalian progranulin in contexts other than in tumourigenesis, making associations with these results difficult. However, several groups have identified progranulin as a target of RA-dependent expression in primary cultures during mesothelial and hematopoietic stem cell differentiation (Ong et al., 2006; Sun et al., 2004). Further investigation is necessary to determine if such an association exist during zebrafish development. Alternatively, the expression of progranulin-a may be regulated by FGFs or *tbx1* and the availability of the associated zebrafish mutants for each would provide an ideal system to provide a conclusion. Of course the reverse may also be true, where the expression of FGFs and/or *tbx1* is dependent upon *progranulin-a*. Perhaps the most curious result is the lack apoptosis within the region of the posterior EMA (Figure 14). It is possible that our TUNEL analysis failed to identify these cells since only two time-points were assessed. Alternatively (and more tantalizing), is that the loss of progranulin-A may have resulted in differentiation of the posterior EMA mass. Unfortunately, assessing this possibility without prior knowledge of a possible marker of the novel

differentiated state would be difficult. Alternatively, EMA cells could be labelled prior to ventral migration using photo-activated fluorescent dye, such as DMNB-caged fluorescein (David et al., 2002).

Although there is no corroborative evidence for a role for progranulins in regulating craniofacial development in mammals, there is suggestive evidence that progranulins may be involved the facial development of various cichlid fish (Kijimoto et al., 2005). Okada and colleagues have focussed their studies on the explosive speciation of cichlid species in Lake Victoria, in the (Eastern African) which is believed to have generated ~500 cichlid species over a 14,000 year period including numerous species that are dependent on different food sources with a craniofacial morphology indicative of feeding habits (Seehausen, 2006; Turner et al., 2001). The cichlid species *Ptyochromis sp*. which has smaller lower jaw architecture, expresses a granulin-like orthologue ~16-fold greater than *Haplochromis parvidens* species which has the longest low jaw architecture of Lake Victoria cichlids (Kijimoto et al., 2005).

4.6 Progranulin and zebrafish digestive system

4.6.1 Discussion of visceral gut development – Apart from those defects associated with the pharyngeal endoderm, derivatives of the remaining posterior endoderm are also affected in progranulin-A morphants. It should be noted that despite the visibly gross morphological defects in progranulin-A morphants at 120 hpf (Figure 9B), the specification and initial development including left-right asymmetric patterning of the digestive system and its associated organs is unaffected (Figures 15 and 16). Specifically, the expression patterns of *pdx-1* and *isl-1* at 32 and 48 hpf, respectively remain largely unchanged in progranulin-A morphants (Figure 15E,F). Despite the presence of a *foxa1*

positive liver bud at 48 hpf, the expression domain appears smaller and has not infiltrated the leftward mesenchyme to the same degree as seen in wild-type animals (Figure 15). The intestine appears similarly decreased in size although it has been correctly oriented. This gross reduction in the size of digestive and accessory organs is in keeping with the overall decrease in animal size. However, the loss of continued *foxa1* expression (Figure 15H) and the absence of hepatic and intestinal markers of differentiation (Figure 15L,M,N) suggests further development of these organs is perturbed in progranulin-A morphants. The visceral gut phenotype is not unexpected given the expression of *progranulin-a* within the epithelial lining of the stomach and intestine, although detection of the transcript within these tissues is only evident at ~120 hpf (Cadieux et al., 2005).

In contrast to this decreased size and absence of differentiation of the liver and intestine, all markers of endocrine pancreatic development and overall dimensions remain unaltered in progranulin-A morphants (Figures 15F and 16E). Intriguingly, the exocrine pancreas is not present based on *trypsin* and carboxypeptidase-A expression (Figure 16F,G,H). To uncover potential patterns of *progranulin-a* expression that were not observed in the original ontogeny study we repeated this exercise with emphasis on stages encompassing 28 to 48 hpf. Interestingly, we identified *progranulin-a* expression within the presumptive endocrine pancreas at 36 (Figure. 16I), however this pattern was not observed at 30 or 40 hpf suggesting a defined temporal requirement for this growth factor (data not shown).

It is not particularly surprising to observe defects in zebrafish pancreas development that specifically affects the exocrine lineage independently of its endocrine counterpart (Mayer and Fishman, 2003; Pack et al., 1996). This is in part due to the spatial and temporal division of endocrine and exocrine pancreas development, wherein

the formation of two independent pancreatic anlagen correspond to specific lineages (Field et al., 2003; Lin et al., 2004a; Wendik et al., 2004). Furthermore, the zebrafish has maintained conserved roles for several genes whose functions pertain to both pancreatic (*pdx-1)* or exclusively exocrine (*p48*/*ptf1a)* anlagen (Lin et al., 2004a; Milewski et al., 1998). However, in the zebrafish no gene has been associated with endocrine pancreas development without affecting its digestive counterpart.

At this time it is not clear whether the expression of *progranulin-a* within the endocrine pancreas affects its surrounding endodermal counterparts, however the first molecular manifestation of the exocrine pancreas coincides with the initiation of *progranulin-a* expression within the developing endocrine lineage (Lin et al., 2004a; Wendik et al., 2004; Zecchin et al., 2004). Given that the development of the exocrine pancreas is perturbed in our morphants we hypothesize that progranulin-A may be involved in paracrine signalling between these two cell populations. Further investigation is required to elucidate the status and migration of progenitors of the exocrine pancreas.

Murine *progranulin* is similarly expressed within the epithelial lining of the digestive tract during development and in adulthood, however this growth factor was not observed to be expressed within the liver or pancreas (Daniel et al., 2003; Daniel et al., 2000). With regards to the absence of *progranulin* expression within the pancreas it is clear that the zebrafish restricts *progranulin-a* expression to a short window of pancreatic development. Therefore the study of *progranulin* expression during murine development could have easily missed this restricted temporal pattern.

A recent study involving the Mozambique tilapia species *Oreochromis mossambicus* identified a *progranulin* gene as a downstream target of growth hormone signalling within the liver (Chen et al., 2007b). In addition, there is mention of MO

knockdown of the identified *progranulin* protein causing a substantial decrease in liver growth. Unfortunately, this result remains as unpublished data.

4.7 Discussion of observed progranulin-A morphant CNS defects – Although the focus of the current body of research has concentrated on the endoderm and its derivatives, the recent identification of mutations within human *progranulin* that result in haploinsufficiency that manifests as FTD-U prompted a strict evaluation of the neuronal expression components of several of our endodermal markers. Most notable was the absence of *dlx2* expression within the presumed hypothalamus (Figure 13E,F) and the loss of Zn8 localization within cranial sensory ganglia and commisurral neurons of the hindbrain (Figure 14E,F). The significance of *progranulin-a* gene expression within the developing zebrafish hypothalamus is unknown. However, it should be noted that *progranulin* expression within the neonatal rat basomedial hypothalamus is involved in sex-dependent differentiation of the brain (Suzuki et al., 2000; Suzuki et al., 1998). With respect to the cranial sensory ganglia, the absence of Zn8 immunolocalization was not correlated with tandem alterations in the expression of *isl-1* expression (data not shown), therefore the defect may pertain to neuronal fasciculation rather than formation (Fashena and Westerfield, 1999; Higashijima et al., 2000).

3.7 Summary of progranulin-A morphant phenotyping

3.7.1 With a cherry on top; final validation of morphant phenotypes – As previously discussed the use of two independent non-overlapping MOs against the target transcript is necessary for validation of the resultant phenotypes. However, the use of progranulin-a mRNA to rescue the morphant phenotype is considered a more stringent validation. It is

unlikely that these results will be accepted for publication without an attempt at mRNA rescue. However, there is reason to believe that this experiment may prove particularly difficult since mRNA rescue is typically used with a reporter system at <28 hpf. As the molecular disparities in progranulin-A morphants only commence at this stage it may be difficult to interpret a rescue experiment.

3.7.2 The Devil's advocate of developmental delay – The explanation of the mechanisms underlying the majority of phenotypes uncovered in progranulin-A morphants suffer from the argument that they can be construed as resulting from developmental delay, which is generally considered as a product of non-specific effects of MO microinjection (Heasman, 2002). Developmental delay can be addressed in two ways. Primarily, a successful mRNA rescue of the phenotype in question remains the benchmark validation method. In addition, it is worthwhile to assess those expression patterns and associated tissues using multiple markers, including those distinctive for differentiated and progenitor populations at several time points to determine the tissue in question becomes apparent after a short period. For example, in progranulin-A morphants the presence of liver and exocrine pancreas progenitors are unaffected at the earliest stages of tissue-specific marker expression (Figure 16). However, it would be informative to assess the expression patterns of these markers when differentiation expression patterns are absent (*trypsin* and *transferrin*) (Figure 17). Similarly, one could argue that assessing the expression of differentiation markers later in development would be beneficial however, there remains the possibility that the injected MO has been diluted sufficiently to allow for the differentiation of any remaining progenitor pool. Furthermore, With regards to the progranulin-A morphant defects in posterior EMA

development (Figure 13) it is important to note that cranial NC cells are also required for development of the thymus, which appears to be intact (Figure 11).

4.8.3 Future zebrafish directions; mutants or morphants – With regards to hematopoiesis, a description of both zebrafish and possible *in vitro* analysis of human HSCs has been described in detail in previous sections (Figures 22 and 23). In addition, a series of follow up experiments has been outlined for each section, pertaining to each phenotype. The immediate concern is to complete the validation of the progranulin-A morphant phenotypes by mRNA rescue experiments. However, further investigation and dissection of the molecular pathways involved and affected by progranulin-A ablation need not rely solely on the use of MOs for gene inactivation. Znomics (Portland, OR) has developed a commercially available insertional mutant database that can be readily mined for insertions within a gene of interest (Inc., http://zenemark.znomics.com/index.html). Interestingly, there appears to be a single viral insertional mutant for the *progranulin-a* gene within intron four (Reference ZM_00236472) the same intron that is targeted by the progranulin-a Sp-MO (Figure 10). Further validation of the insertion site is required to confirm the exa ct location of the viral insertion. However, viral integration within an intronic sequence does not necessarily mean that *progranulin-a* transcription will be affected. Despite this caveat, a known mutant background can be of exceptional value for not only confirming the results of the morphants studies described within this text, but also for the continued investigation of *progranulin-a* function without the need for microinjection of individual embryos.

4.9 Progranulin protein-protein interactions

4.9.1 Preamble – As described previously (section 1.5) several putative progranulin binding proteins have been identified, the majority of these have been defined using the yeast two-hybrid (Y2H) system. The best characterized and validated progranulin protein-protein interactions include those involving SLPI, COMP and perlecan (Gonzalez et al., 2003; Xu et al., 2007; Zhu et al., 2002). Despite positive Y2H results, corroborative co-immunoprecipitation data and demonstration of synergism in cell growth assays that argue for the interaction between progranulin and SLPI, COMP and perlecan, there remains no clear assessment as to whether these proteins augment the progranulin-mediated activation of known intracellular signalling pathways (Xu et al., 1998; Zanocco-Marani et al., 1999). Although this lack of characterization of signalling mechanism does not bring into question the validity of the published data, without this information the validation of the physiological significance of the progranulin-binding remains incomplete the lack of signalling mechanisms does not bring into question the validity of the published material it is intriguing as to why this type of assay or their results were not included or attempted.

Based on the documented protein-protein interactions involving progranulin it would appear that this molecule does not have a defined and restricted interaction network. Indeed, with the number of protein-protein interactions that are presumed to involve progranulin it is tempting to invoke a "scaffold protein" descriptor to this growth factor. For instance, the ubiquitous transcription factor CREB-binding protein (CBP) interacts with at least 312 other nuclear proteins (Kasper et al., 2006). Interestingly, proteins which are comprised of tandem repetitive domains have often been found to be scaffold proteins that act as signalling "hubs", from which numerous downstream

pathways are modified (Pawson and Nash, 2003). The trend for repeat domain proteins to act as major interaction hubs originates from several large-scale "interactome" experiments in the yeast (Ekman et al., 2006; Han et al., 2004).

Although progranulin protein-protein interactions are clearly not specific in regards to one protein family, they may be selective for proteins that contain particular domains. Interestingly, the progranulin binding proteins COMP, perlecan and DLK1 all contain epidermal growth factor (EGF) domains within the region of progranulin binding (Baladron et al., 2002; Gonzalez et al., 2003; Xu et al., 2007; Zhu et al., 2002). It would be premature to suggest that progranulin is a generalized binding protein for EGF-containing proteins in view of the circumstances in which there interacting partners were identified. Specifically, one can theoretically conceive that an increased rate of interaction between progranulin and these similarly cysteine rich proteins may be solely a reflection of aberrant intermolecular cystine-bridging since the yeast lack the perquisite redox microenvironment for proper disulphide bridging. If the Y2H system were in fact prone to such intermolecular cystine-bridging events, there is the expectation for an inordinately high rate of progranulin self-association. However, such a result has not been described thus far. Overall, the Y2H system is not compatible with screening for secreted cystine-rich proteins and it is necessary to use a different system of protein-protein interaction screening for progranulin.

For the identification of protein-protein interactions of secreted factors the tandem affinity purification tag (TAP-tag) and protein microarrays represent the two current avenues of experimental investigation (Burckstummer et al., 2006; Gavin et al., 2002; Rigaut et al., 1999). Unfortunately, protein microarray screens remain prohibitively expensive. Furthermore, this type of screen is limited in the number of baits incorporated

into the array and, if interactions are found, little information is revealed regarding binding affinities. The TAP-tag approach requires two-rounds of affinity purification, therefore a large degree of experimental validation is necessary to ensure ligand and target protein capture are compatible. The use of surface plasmon resonance (SPR) as a system to monitor protein-protein interactions addresses several of these limitations.

The SPR technique was first introduced in 1990 as a platform for high throughput *in vitro* kinetic assays for small molecule and protein interactions (Fagerstam et al., 1990; Pollard-Knight et al., 1990). SPR-based biosensors sense the mass concentration of molecules close to a surface in terms of a refractive index change with a readout measured in "response units". By covalently linking a ligand of interest to the surface one can pass over various cellular/tissue extracts to monitor for the presence of binding partners. When molecules bind to the ligand, the local concentration changes and an SPR response is measured through the generation of a sensorgram. The response is directly proportional to the mass of molecules that bind to the surface. Initial validation of the covalent linkage of recombinant progranulin-A and its ability to interact with proteins within the solute were assessed using various dilutions of a rabbit anti-zebrafish progranulin-A (Figure 18).

4.9.2 The methodology of LC-SPR-MS – We have described a simple method of offline proteomic fractionation which is compatible with SPR kinetic analysis and final MS identification of target binding proteins. The exact methods used, including the acidic extraction, RP-HPLC fractionation using a C_{18}-column and lyophilisation prior to SPR analysis are extremely harsh manipulations that are not conducive to protein solubility nor the maintenance of tertiary structure for large proteins. These facts were not unknown

during the experimental design. Indeed, the harsh conditions are more suited to the purification of small cysteine-rich proteins and peptides, such as growth factors and other proteins containing single EGF domains. The structure of PBP does conform to these presumptions as it is a small (~20kDa) protein that is presumed to conform to the typical cystine-bridging motif found in its orthologous antifreeze and REG counterparts (Figure 21). The caveat of this analysis is that we did not test the versatility of the LC-SPR interface using multiple chromatographic methods, due to the sensitivity of SPR to variations in salt and detergent concentrations (Canziani et al., 1999; Fagerstam et al., 1990). Indeed, it is the sensitivity of SPR to sample mixtures that has limited its use as a tool for protein-protein detection. However, we have demonstrated that LC-SPR can be used for protein-protein interaction screening if sufficient care is taken to allow compatibility.

The combination of SPR-MS has been limited to experiments such as ours, wherein a sample is screened for a positive interaction and characterized independently. Efforts to place these to powerful techniques in tandem by eluting captured ligand from the SPR chip surface and characterizing by MS, are ongoing and may provide more efficient through-put in protein-protein interaction screening although the initial kinetics or association must still be optimized for both target protein(s) and equipment limitations (Nedelkov, 2007; Yu et al., 2006).

4.9.3 Validation of PBP binding – Previously, the data regarding Y2H screens for identification of progranulin binding proteins was assessed for accurate and meaningful validation through co-immunoprecipitation and potential modulation of known progranulin intracellular signalling pathways (Ewart et al., 1998; Sonnichsen et al., 1995).

These assays represent the experimental analysis required to confirm the association between progranulin-A/V5/His and PBP. Furthermore, the nature of the ligand used and putative binding protein require additional experimental validation. Specifically, the epitope(s) associated with the protein-protein interaction must be confirmed to occur between progranulin and PBP, rather than within the C-terminal epitope tag. As well, the structure of PBP is similar to that of fish antifreeze proteins (data not shown) which are circulating humoral proteins in that prohibit water crystallization (Knight et al., 1984). The effective inhibition of ice formation by antifreeze proteins is dependent on monosaccharide (typically mannose) and divalent ion (typically Ca^{+2}) binding (Ewart et al., 1998; Sonnichsen et al., 1995). As the recombinant progranulin-A/V5/His was produced in insect cells that typically introduce simple glycosylation products with terminal mannose moieties, it will be necessary to add oligosaccharides as part of the solvent during co-immunoprecipitation analysis (Jarvis and Finn, 1996). Provided that the interaction between progranulin-A and PBP proves genuine, it is conceivable to begin investigating the function of this gene during zebrafish development within the contexts of those organ systems whose growth and maintenance are dependent upon progranulin-A.

4.9.4 Progranulins and the C-type lectin superfamily – Zebrafish PBP is a cysteine-rich polypeptide, a characteristic shared with almost all known growth factor binding proteins, receptors and co-receptors. The relatively low molecular weight of PBP (~20,000 Da) in comparison to progranulin-A itself (~100,000 Da) suggests that the latter may recruit PBP to form a bioactive complex. The characterization of the C-type lectin superfamily, including the members structurally similar to PBP, has not been undertaken

in the zebrafish at this time. However, a growing body of data exists for the human REG family within multiple contexts.

The human PBP orthologues are considered to be secreted growth factors that signal through multiple pathways including the MAPK and PI3K cascades (Malka et al., 2000). They have been characterized as either trophic or prognostic factors in many cancers including pancreatic hepatic and gastric carcinoma (Cavard et al., 2006; Dhar et al., 2004; Kimura et al., 1992; Oue et al., 2005; Takehara et al., 2006). The REG family has also been associated with the inflammatory response and diabetes and (Zhang et al., 2003). Of the four murine REG family members *Reg1* and *Reg3γ* have been knocked out and have been shown to be involved in the formation and maturation of the fetal intestinal villi and the regulation of the inflammatory response during pancreatitis, respectively (Gironella et al., 2007; Ose et al., 2007). Human Reg1 and rat Reg1α have been shown to function in association with a putative transmembrane receptor (EXTL3), an α1, 4-N-acetylglucosaminyltransferase involved in heparin sulfate biosynthesis (Kobayashi et al., 2000). Although EXTL3 has yet to be directly associated with any signal transduction pathway, this may represent a potential progranulin receptor or co-receptor.

The association between progranulin and PBP may be a common trend towards C-type lectin association. The C-type lectin superfamily of proteins represent a superfamily that includes a number of orphan transmembrane proteins including several NK cell receptors (Drickamer and Fadden, 2002).

4.10 General Conclusions – The *progranulin* gene family represents and intriguing example of both stringent structural conservation as well as rapid and ongoing genetic

variation. Indeed, the identifcation of such a large number of evolutionary pressures incuding genome-wide duplication, DNA and RNA transposons as well as the discrete segmental deletions observed in betweenvarious mammals, illustrates the impressive evolutionary pressures that shape our genomes. It may be argued that the reason why the core granulin motif has not evolved at a similar pace is due to the restrictive nature of the motif itself with its twelve cysteine bridges. However, it should be noted that the granulin motif is unlike other similar cysteine-rich domains (e.g. EGFs), in that there has been no (perhaps viable?) incorporation of this domain within other proteins of the chordate genome. The single exemption being the Cysteine protease of *Arabidopsis thalania*. Ultimately, a dissection of the progranulin protein for bioactive domains and specific residues that are intrinsic to proper folding will provide a strong biochemical sieve to assess the consequences of such a diverse array of gene repeat structures throughout evolution. With the identification of human *progranulin* mutations as an underlying cause of FTLD-U and dissection of specific mutants, this road has already become well travelled and will continue to be a rhealm of investigation for the entire progranulin community.

The results described in this study represent experimental analyses that have evolved in parallel with the elucidation of both conserved and unique mechanisms of zebrafish development. Indeed, the number of organ systems that appear perturbed due to progranulin-A knockdown are sufficient to provide several decades worth of experimental investigation that will inevitably migrate into mammalian models, both *in vitro* and *in vivo*. With this in mind the zebrafish can be considered in some respects a stepping stone for the eventual migration of all experimental analysis into mammalian systems. However, the inherent nature of the zebrafish model including large breeding

capacity, optical clear and so forth, is likely to ensure that the model will continue to be used a worthwhile comparative model. Previous sections hae given a strong outline for the continued investigation of progranulin function in hematopoiesis and our understanding of this growth factor within this context is further supported by a large body of evidence that indicates many neurodegenerative disorders are associated with abnormal microglial activation. With respect to the craniofacial and visceral gut development, the intial steps haven been taken and this work will always be considered a descriptive investigation of progranulin function within these contexts. The next plateau that must be strived towards is a mechanistic understanding of progranulin function within these developmental events. Placing progranulin within the context of other known growth factors, transcription factor and mutant backgrounds will provide a wealth of information.

Recently, the zebrafish has left its developmental infancy and formed inroads as an adult disease model of cancer (Chen et al., 2007a; Faucherre et al., 2007; Goessling et al., 2007). Transgenic zebrafish models of several neurodegenerative disease have also been developed, although the deleterious effects of transgene over expression manifest during development (Lemmens et al., 2007; Lumsden et al., 2007; Tomasiewicz et al., 2002). It will be particular interest to develop adult zebrafish models of neurodegenerative disorders, as teleosts are capable of regeneration and therefore, may offer new insight the mechanisms as well as possible therapeutic targets (Akimenko et al., 2003; Curado and Stainier, 2006).

The use of morpholinos as a method to uncover gene function will continue to play a large role in zebrafish development, however the amount of validation and the possibility of affecting both specific and non-specific phenotypes will remain a concern.

Several investigators have begun the task of applying gene knockout technology to the zebrafish although it remains to be seen if this technology will impart sufficient efficiency to justify its use (Wu et al., 2006). As previously mentioned, there would appear to a commercially available insertional mutation of the zebrafish *progranulin-a* gene. Although the viral insertion is intronic and may not result in inactivation, the value of a *bona fide* mutant background to continue the investigation of gene function during development (and beyond) is sufficient to justify the perquisite costs.

The development of a LC-SPR-MS methodology to investigate progranulin protein-protein interactions is a particularly novel advent which highlights the potential of such technology as well as the difficulties in combining classical protein fractionation with SPR biosensors. While the experimental design developed in this study is amenable to small scale, single ligand protein-protein interaction screens the true potential LC-SPR-MS is to wed these technologies as a single continuous unit. As biotechnology appears to reinvent itself at an extraordinary rate it is inevitable that on-line SPR surface capture and downstream MS characterization will be realized. Such a high-throughput system will undoubtedly produce vast databases of interaction networks. However, validation of these results will require classical immunoprecipitation or similar experimental analysis. Therein lies the crux of such high-throughput analysis and this is reminiscent of the larger scale real-time PCR validation of microarray data. It may be argued with some certainty that if one choose to endeavour on such a vast experimental agenda a preliminary concept of specific gene families or pathways of interest is of strong importance. Paradoxically, if a gene family or pathway is of interest this may question the utility of an "omic" study and a defined analysis of these few targets may be worthwhile. With this in mind the identification of PBP should be weighed, not solely as

a single binding protein, but as a C-type lectin that may imply a generalized interaction network between progranulins and this superfamily.

3.10 Contributions to Knowledge

The following experimental observations represent original contributions to science:

1. Abrogation of the translation of progranulin-a in the zebrafish developmental model.

2. Identification and characterization of progranulin-A morphant defects in hematopoietic stem cells and downstream lineages.

3. Identification and molecular dissection of progranulin-A morphant defects in craniofacial morphogenesis and the requirement for this growth factor in the morphogenesis of the pharyngeal endoderm and posterior neural crest.

4. Identification and molecular dissection of progranulin-A morphant defects in visceral gut and accessory organ growth and differentiation with respect to the liver and exocrine pancreas.

5. Development and application of a novel system of protein purification and characterization during protein-protein interaction screening, LC-SPR-MS.

6. Identification of a putative zebrafish progranulin-A binding protein using LC-SPR-MS.

5.1 References

(1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282**, 2012-8.

(2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69-87.

Abi-Rached, L., Gilles, A., Shiina, T., Pontarotti, P. and Inoko, H. (2002). Evidence of en bloc duplication in vertebrate genomes. *Nat Genet* **31**, 100-5.

Ahlgren, U., Jonsson, J. and Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development* **122**, 1409-16.

Akimenko, M. A., Ekker, M., Wegner, J., Lin, W. and Westerfield, M. (1994). Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head. *J Neurosci* **14**, 3475-86.

Akimenko, M. A., Mari-Beffa, M., Becerra, J. and Geraudie, J. (2003). Old questions, new tools, and some answers to the mystery of fin regeneration. *Dev Dyn* **226**, 190-201.

Albini, A. and Sporn, M. B. (2007). The tumour microenvironment as a target for chemoprevention. *Nat Rev Cancer* **7**, 139-47.

Alexander, G. M., Erwin, K. L., Byers, N., Deitch, J. S., Augelli, B. J., Blankenhorn, E. P. and Heiman-Patterson, T. D. (2004). Effect of transgene copy number on survival in the G93A SOD1 transgenic mouse model of ALS. *Brain Res Mol Brain Res* **130**, 7-15.

Amatruda, J. F. and Zon, L. I. (1999). Dissecting hematopoiesis and disease using the zebrafish. *Dev Biol* **216**, 1-15.

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L. et al. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science* **282**, 1711-4.

Amores, A., Suzuki, T., Yan, Y. L., Pomeroy, J., Singer, A., Amemiya, C. and Postlethwait, J. H. (2004). Developmental roles of pufferfish Hox clusters and genome evolution in ray-fin fish. *Genome Res* **14**, 1-10.

Amsterdam, A., Burgess, S., Golling, G., Chen, W., Sun, Z., Townsend, K., Farrington, S., Haldi, M. and Hopkins, N. (1999). A large-scale insertional mutagenesis screen in zebrafish. *Genes Dev* **13**, 2713-24.

Arnold, J. S., Werling, U., Braunstein, E. M., Liao, J., Nowotschin, S., Edelmann, W., Hebert, J. M. and Morrow, B. E. (2006). Inactivation of Tbx1 in the pharyngeal endoderm results in 22q11DS malformations. *Development* **133**, 977-87.

Aronheim, A., Zandi, E., Hennemann, H., Elledge, S. J. and Karin, M. (1997). Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. *Mol Cell Biol* **17**, 3094-102.

Asaka, S., Fujimoto, T., Akaishi, J., Ogawa, K. and Onda, M. (2006). Genetic prognostic index influences patient outcome for node-positive breast cancer. *Surg Today* **36**, 793-801.

Baba, T., Nemoto, H., Watanabe, K., Arai, Y. and Gerton, G. L. (1993). Exon/intron organization of the gene encoding the mouse epithelin/granulin precursor (acrogranin). *FEBS Lett* **322**, 89-94.

Bailey, W. J., Kim, J., Wagner, G. P. and Ruddle, F. H. (1997). Phylogenetic reconstruction of vertebrate Hox cluster duplications. *Mol Biol Evol* **14**, 843-53.

Baker, M., Mackenzie, I. R., Pickering-Brown, S. M., Gass, J., Rademakers, R., Lindholm, C., Snowden, J., Adamson, J., Sadovnick, A. D., Rollinson, S. et al. (2006). Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature* **442**, 916-9.

Baladron, V., Ruiz-Hidalgo, M. J., Bonvini, E., Gubina, E., Notario, V. and Laborda, J. (2002). The EGF-like homeotic protein dlk affects cell growth and interacts with growth-modulating molecules in the yeast two-hybrid system. *Biochem Biophys Res Commun* **291**, 193-204.

Balducci, E., Azzarello, G., Valenti, M. T., Capuzzo, G. M., Pappagallo, G. L., Pilotti, I., Ausoni, S., Bari, M., Rosetti, F., Sartori, D. et al. (2003). The impact of progenitor enrichment, serum, and cytokines on the ex vivo expansion of mobilized peripheral blood stem cells: a controlled trial. *Stem Cells* **21**, 33-40.

Bateman, A., Belcourt, D., Bennett, H., Lazure, C. and Solomon, S. (1990). Granulins, a novel class of peptide from leukocytes. *Biochem Biophys Res Commun* **173**, 1161-8.

Bateman, A. and Bennett, H. P. (1998). Granulins: the structure and function of an emerging family of growth factors. *J Endocrinol* **158**, 145-51.

Beglopoulos, V. and Shen, J. (2004). Gene-targeting technologies for the study of neurological disorders. *Neuromolecular Med* **6**, 13-30.

Behrens, M. I., Mukherjee, O., Tu, P. H., Liscic, R. M., Grinberg, L. T., Carter, D., Paulsmeyer, K., Taylor-Reinwald, L., Gitcho, M., Norton, J. B. et al. (2007). Neuropathologic heterogeneity in HDDD1: a familial frontotemporal lobar degeneration with ubiquitin-positive inclusions and progranulin mutation. *Alzheimer Dis Assoc Disord* **21**, 1-7.

Belcourt, D. R., Lazure, C. and Bennett, H. P. (1993). Isolation and primary structure of the three major forms of granulin-like peptides from hematopoietic tissues of a teleost fish (Cyprinus carpio). *J Biol Chem* **268**, 9230-7.

Bennett, C. M., Kanki, J. P., Rhodes, J., Liu, T. X., Paw, B. H., Kieran, M. W., Langenau, D. M., Delahaye-Brown, A., Zon, L. I., Fleming, M. D. et al. (2001a). Myelopoiesis in the zebrafish, Danio rerio. *Blood* **98**, 643-51.

Bennett, C. M., Kanki, J. P., Rhodes, J., Liu, T. X., Paw, B. H., Kieran, M. W., Langenau, D. M., Delahaye-Brown, A., Zon, L. I. and Fleming, M. D. e. a. (2001b). Myelopoiesis in the zebrafish, Danio rerio. *Blood* **98**, 643-651.

Bennett, H. P., Browne, C. A. and Solomon, S. (1981). Purification of the two major forms of rat pituitary corticotropin using only reversed-phase liquid chromatography. *Biochemistry* **20**, 4530-8.

Benussi, L., Binetti, G., Sina, E., Gigola, L., Bettecken, T., Meitinger, T. and Ghidoni, R. (2006). A novel deletion in progranulin gene is associated with FTDP-17 and CBS. *Neurobiol Aging*.

Bhandari, V. and Bateman, A. (1992). Structure and chromosomal location of the human granulin gene. *Biochem Biophys Res Commun* **188**, 57-63.

Bhandari, V., Giaid, A. and Bateman, A. (1993). The complementary deoxyribonucleic acid sequence, tissue distribution, and cellular localization of the rat granulin precursor. *Endocrinology* **133**, 2682-9.

Bhandari, V., Palfree, R. G. and Bateman, A. (1992). Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains. *Proc Natl Acad Sci U S A* **89**, 1715-9.

Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B. and Driever, W. (2001). Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. *Dev Biol* **230**, 189-203.

Bjorklund, A. K., Ekman, D. and Elofsson, A. (2006). Expansion of protein domain repeats. *PLoS Comput Biol* **2**, e114.

Boelz, S., Neu-Yilik, G., Gehring, N. H., Hentze, M. W. and Kulozik, A. E. (2006). A chemiluminescence-based reporter system to monitor nonsense-mediated mRNA decay. *Biochem Biophys Res Commun* **349**, 186-91.

Boeve, B. F., Baker, M., Dickson, D. W., Parisi, J. E., Giannini, C., Josephs, K. A., Hutton, M., Pickering-Brown, S. M., Rademakers, R., Tang-Wai, D. et al. (2006). Frontotemporal dementia and parkinsonism associated with the IVS1+1G->A mutation in progranulin: a clinicopathologic study. *Brain* **129**, 3103-14.

Boyd, E. F. and Hartl, D. L. (1997). Nonrandom location of IS1 elements in the genomes of natural isolates of Escherichia coli. *Mol Biol Evol* **14**, 725-32.

Breeden, L. and Nasmyth, K. (1987). Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of Drosophila. *Nature* **329**, 651-4.

Broder, Y. C., Katz, S. and Aronheim, A. (1998). The ras recruitment system, a novel approach to the study of protein-protein interactions. *Curr Biol* **8**, 1121-4.

Bronner, I. F., Rizzu, P., Seelaar, H., van Mil, S. E., Anar, B., Azmani, A., Kaat, L. D., Rosso, S., Heutink, P. and van Swieten, J. C. (2007). Progranulin mutations in Dutch familial frontotemporal lobar degeneration. *Eur J Hum Genet* **15**, 369-74.

Brownlie, A., Donovan, A., Pratt, S. J., Paw, B. H., Oates, A. C., Brugnara, C., Witkowska, H. E., Sassa, S. and Zon, L. I. (1998). Positional cloning of the zebrafish sauternes gene: a model for congenital sideroblastic anaemia. *Nat Genet* **20**, 244-50.

Brownlie, A., Hersey, C., Oates, A. C., Paw, B. H., Falick, A. M., Witkowska, H. E., Flint, J., Higgs, D., Jessen, J., Bahary, N. et al. (2003). Characterization of embryonic globin genes of the zebrafish. *Dev Biol* **255**, 48-61.

Bruni, A. C., Momeni, P., Bernardi, L., Tomaino, C., Frangipane, F., Elder, J., Kawarai, T., Sato, C., Pradella, S., Wakutani, Y. et al. (2007). Heterogeneity within a large kindred with frontotemporal dementia: a novel progranulin mutation. *Neurology* **69**, 140-7.

Burckstummer, T., Bennett, K. L., Preradovic, A., Schutze, G., Hantschel, O., Superti-Furga, G. and Bauch, A. (2006). An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat Methods* **3**, 1013-9.

Cadieux, B., Chitramuthu, B. P., Baranowski, D. and Bennett, H. P. (2005). The zebrafish progranulin gene family and antisense transcripts. *BMC Genomics* **6**, 156.

Canziani, G., Zhang, W., Cines, D., Rux, A., Willis, S., Cohen, G., Eisenberg, R. and Chaiken, I. (1999). Exploring biomolecular recognition using optical biosensors. *Methods* **19**, 253-69.

Carruthers, S. and Stemple, D. L. (2006). Genetic and genomic prospects for Xenopus tropicalis research. *Semin Cell Dev Biol* **17**, 146-53.

Cavard, C., Terris, B., Grimber, G., Christa, L., Audard, V., Radenen-Bussiere, B., Simon, M. T., Renard, C. A., Buendia, M. A. and Perret, C. (2006). Overexpression of regenerating islet-derived 1 alpha and 3 alpha genes in human primary liver tumors with beta-catenin mutations. *Oncogene* **25**, 599-608.

Chen, J., Jette, C., Kanki, J. P., Aster, J. C., Look, A. T. and Griffin, J. D. (2007a). NOTCH1-induced T-cell leukemia in transgenic zebrafish. *Leukemia* **21**, 462-71.

Chen, M. H., Li, Y. H., Chang, Y., Hu, S. Y., Gong, H. Y., Lin, G. H., Chen, T. T. and Wu, J. L. (2007b). Co-induction of hepatic IGF-I and progranulin mRNA by growth hormone in tilapia, Oreochromis mossambiccus. *Gen Comp Endocrinol* **150**, 212-8.

Chen, Y. and Schier, A. F. (2001). The zebrafish Nodal signal Squint functions as a morphogen. *Nature* **411**, 607-10.

Cheng, T., Shen, H., Giokas, D., Gere, J., Tenen, D. G. and Scadden, D. T. (1996). Temporal mapping of gene expression levels during the differentiation of individual primary hematopoietic cells. *Proc Natl Acad Sci U S A* **93**, 13158-63.

Cheung, S. T., Wong, S. Y., Leung, K. L., Chen, X., So, S., Ng, I. O. and Fan, S. T. (2004). Granulin-epithelin precursor overexpression promotes growth and invasion of hepatocellular carcinoma. *Clin Cancer Res* **10**, 7629-36.

Chisholm, R. L., Gaudet, P., Just, E. M., Pilcher, K. E., Fey, P., Merchant, S. N. and Kibbe, W. A. (2006). dictyBase, the model organism database for Dictyostelium discoideum. *Nucleic Acids Res* **34**, D423-7.

Clancy, J. P., Bebok, Z., Ruiz, F., King, C., Jones, J., Walker, L., Greer, H., Hong, J., Wing, L., Macaluso, M. et al. (2001). Evidence that systemic gentamicin suppresses premature stop mutations in patients with cystic fibrosis. *Am J Respir Crit Care Med* **163**, 1683-92.

Creuzet, S., Couly, G. and Le Douarin, N. M. (2005). Patterning the neural crest derivatives during development of the vertebrate head: insights from avian studies. *J Anat* **207**, 447-59.

Crump, J. G., Maves, L., Lawson, N. D., Weinstein, B. M. and Kimmel, C. B. (2004). An essential role for Fgfs in endodermal pouch formation influences later craniofacial skeletal patterning. *Development* **131**, 5703-16.

Cruts, M., Gijselinck, I., van der Zee, J., Engelborghs, S., Wils, H., Pirici, D., Rademakers, R., Vandenberghe, R., Dermaut, B., Martin, J. J. et al. (2006). Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature* **442**, 920-4.

Culouscou, J. M., Carlton, G. W. and Shoyab, M. (1993). Biochemical analysis of the epithelin receptor. *J Biol Chem* **268**, 10458-62.

Curado, S. and Stainier, D. Y. (2006). The HeArt of regeneration. *Cell* **127**, 462-4.

Curwen, V., Eyras, E., Andrews, T. D., Clarke, L., Mongin, E., Searle, S. M. and Clamp, M. (2004). The Ensembl automatic gene annotation system. *Genome Res* **14**, 942-50.

Dang, C. V., Barrett, J., Villa-Garcia, M., Resar, L. M., Kato, G. J. and Fearon, E. R. (1991). Intracellular leucine zipper interactions suggest c-Myc hetero-oligomerization. *Mol Cell Biol* **11**, 954-62.

Daniel, R., Daniels, E., He, Z. and Bateman, A. (2003). Progranulin (acrogranin/PC cell-derived growth factor/granulin-epithelin precursor) is expressed in the placenta, epidermis, microvasculature, and brain during murine development. *Dev Dyn* **227**, 593-9.

Daniel, R., He, Z., Carmichael, K. P., Halper, J. and Bateman, A. (2000). Cellular localization of gene expression for progranulin. *J Histochem Cytochem* **48**, 999-1009.

David, N. B., Saint-Etienne, L., Tsang, M., Schilling, T. F. and Rosa, F. M. (2002). Requirement for endoderm and FGF3 in ventral head skeleton formation. *Development* **129**, 4457-68.

Davion, S., Johnson, N., Weintraub, S., Mesulam, M. M., Engberg, A., Mishra, M., Baker, M., Adamson, J., Hutton, M., Rademakers, R. et al. (2007). Clinicopathologic correlation in PGRN mutations. *Neurology*.

Dawson, H. N., Ferreira, A., Eyster, M. V., Ghoshal, N., Binder, L. I. and Vitek, M. P. (2001). Inhibition of neuronal maturation in primary hippocampal neurons from tau deficient mice. *J Cell Sci* **114**, 1179-87.

De Block, M. and Debrouwer, D. (1993). RNA-RNA in situ hybridization using digoxigenin-labeled probes: the use of high-molecular-weight polyvinyl alcohol in the alkaline phosphatase indoxyl-nitroblue tetrazolium reaction. *Anal Biochem* **215**, 86-9.

Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D. M. et al. (2002). The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. *Science* **298**, 2157-67.

Denk, F. and Wade-Martins, R. (2007). Knock-out and transgenic mouse models of tauopathies. *Neurobiol Aging*.

Dewannieux, M., Esnault, C. and Heidmann, T. (2003). LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* **35**, 41-8.

Dhar, D. K., Udagawa, J., Ishihara, S., Otani, H., Kinoshita, Y., Takasawa, S., Okamoto, H., Kubota, H., Fujii, T., Tachibana, M. et al. (2004). Expression of regenerating gene I in gastric adenocarcinomas: correlation with tumor differentiation status and patient survival. *Cancer* **100**, 1130-6.

Di Gregorio, A., Spagnuolo, A., Ristoratore, F., Pischetola, M., Aniello, F., Branno, M., Cariello, L. and Di Lauro, R. (1995). Cloning of ascidian homeobox genes provides evidence for a primordial chordate cluster. *Gene* **156**, 253-7.

Diaz-Cueto, L., Stein, P., Jacobs, A., Schultz, R. M. and Gerton, G. L. (2000). Modulation of mouse preimplantation embryo development by acrogranin (epithelin/granulin precursor). *Dev Biol* **217**, 406-18.

Donald, C. D., Laddu, A., Chandham, P., Lim, S. D., Cohen, C., Amin, M., Gerton, G. L., Marshall, F. F. and Petros, J. A. (2001). Expression of progranulin and the epithelin/granulin precursor acrogranin correlates with neoplastic state in renal epithelium. *Anticancer Res* **21**, 3739-42.

Donovan, A., Brownlie, A., Zhou, Y., Shepard, J., Pratt, S. J., Moynihan, J., Paw, B. H., Drejer, A., Barut, B., Zapata, A. et al. (2000). Positional cloning of zebrafish ferroportin1 identifies a conserved vertebrate iron exporter. *Nature* **403**, 776-81.

Draper, B. W., Morcos, P. A. and Kimmel, C. B. (2001). Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis* **30**, 154-6.

Drickamer, K. and Fadden, A. J. (2002). Genomic analysis of C-type lectins. *Biochem Soc Symp*, 59-72.

Eckfeldt, C. E., Mendenhall, E. M., Flynn, C. M., Wang, T. F., Pickart, M. A., Grindle, S. M., Ekker, S. C. and Verfaillie, C. M. (2005). Functional analysis of human hematopoietic stem cell gene expression using zebrafish. *PLoS Biol* **3**, e254.

Edvardsson, L., Dykes, J., Olsson, M. L. and Olofsson, T. (2004). Clonogenicity, gene expression and phenotype during neutrophil versus erythroid differentiation of cytokine-stimulated CD34+ human marrow cells in vitro. *Br J Haematol* **127**, 451-63.

Ehrhard, K. N., Jacoby, J. J., Fu, X. Y., Jahn, R. and Dohlman, H. G. (2000). Use of G-protein fusions to monitor integral membrane protein-protein interactions in yeast. *Nat Biotechnol* **18**, 1075-9.

Eichinger, L., Pachebat, J. A., Glockner, G., Rajandream, M. A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q. et al. (2005). The genome of the social amoeba Dictyostelium discoideum. *Nature* **435**, 43-57.

Eisen, J. S. (1996). Zebrafish make a big splash. *Cell* **87**, 969-77.

Eisen, J. S. and Weston, J. A. (1993). Development of the neural crest in the zebrafish. *Dev Biol* **159**, 50-9.

Ekker, S. C. (2000). Morphants: a new systematic vertebrate functional genomics approach. *Yeast* **17**, 302-306.

Ekman, D., Light, S., Bjorklund, A. K. and Elofsson, A. (2006). What properties characterize the hub proteins of the protein-protein interaction network of Saccharomyces cerevisiae? *Genome Biol* **7**, R45.

Etcheverry, G. J. (1995). [Nobel prize of physiology or medicine 1955: Edward B. Lewis, Christiane Nusslein-Volhard, Eric Wieschaus. The flies and the keys of the embryonic development]. *Medicina (B Aires)* **55**, 715-7.

Eungdamrong, N. J. and Iyengar, R. (2004a). Computational approaches for modeling regulatory cellular networks. *Trends Cell Biol* **14**, 661-9.

Eungdamrong, N. J. and Iyengar, R. (2004b). Modeling cell signaling networks. *Biol Cell* **96**, 355-62.

Everett, C. M. and Wood, N. W. (2004). Trinucleotide repeats and neurodegenerative disease. *Brain* **127**, 2385-405.

Ewart, K. V., Li, Z., Yang, D. S., Fletcher, G. L. and Hew, C. L. (1998). The ice-binding site of Atlantic herring antifreeze protein corresponds to the carbohydrate-binding site of C-type lectins. *Biochemistry* **37**, 4080-5.

Fagerstam, L. G., Frostell, A., Karlsson, R., Kullman, M., Larsson, A., Malmqvist, M. and Butt, H. (1990). Detection of antigen-antibody interactions by surface plasmon resonance. Application to epitope mapping. *J Mol Recognit* **3**, 208-14.

Fashena, D. and Westerfield, M. (1999). Secondary motoneuron axons localize DM-GRASP on their fasciculated segments. *J Comp Neurol* **406**, 415-24.

Faucherre, A., Taylor, G. S., Overvoorde, J., Dixon, J. E. and Hertog, J. D. (2007). Zebrafish pten genes have overlapping and non-redundant functions in tumorigenesis and embryonic development. *Oncogene*.

Feldman, B., Gates, M. A., Egan, E. S., Dougan, S. T., Rennebeck, G., Sirotkin, H. I., Schier, A. F. and Talbot, W. S. (1998). Zebrafish organizer development and germ-layer formation require nodal-related signals. *Nature* **395**, 181-5.

Field, H. A., Dong, P. D., Beis, D. and Stainier, D. Y. (2003). Formation of the digestive system in zebrafish. II. Pancreas morphogenesis. *Dev Biol* **261**, 197-208.

Fields, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-6.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-11.

Fondon, J. W., 3rd and Garner, H. R. (2004). Molecular origins of rapid and continuous morphological evolution. *Proc Natl Acad Sci U S A* **101**, 18058-63.

Fondon, J. W., 3rd and Garner, H. R. (2007). Detection of length-dependent effects of tandem repeat alleles by 3-D geometric decomposition of craniofacial variation. *Dev Genes Evol* **217**, 79-85.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531-45.

Frischmeyer, P. A. and Dietz, H. C. (1999). Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* **8**, 1893-900.

Gaiano, N., Amsterdam, A., Kawakami, K., Allende, M., Becker, T. and Hopkins, N. (1996). Insertional mutagenesis and rapid cloning of essential genes in zebrafish. *Nature* **383**, 829-32.

Galloway, J. L. and Zon, L. I. (2003). Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. *Curr Top Dev Biol* **53**, 139-58.

Garcia-Fernandez, J. and Holland, P. W. (1994). Archetypal organization of the amphioxus Hox gene cluster. *Nature* **370**, 563-6.

Gass, J., Cannon, A., Mackenzie, I. R., Boeve, B., Baker, M., Adamson, J., Crook, R., Melquist, S., Kuntz, K., Petersen, R. et al. (2006). Mutations in progranulin are a major cause of ubiquitin-positive frontotemporal lobar degeneration. *Hum Mol Genet* **15**, 2988-3001.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M. et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-7.

Gessler, M., Knobeloch, K. P., Helisch, A., Amann, K., Schumacher, N., Rohde, E., Fischer, A. and Leimeister, C. (2002). Mouse gridlock: no aortic coarctation or deficiency, but fatal cardiac defects in Hey2 -/- mice. *Curr Biol* **12**, 1601-4.

Gibbs, R. A. Rogers, J. Katze, M. G. Bumgarner, R. Weinstock, G. M. Mardis, E. R. Remington, K. A. Strausberg, R. L. Venter, J. C. Wilson, R. K. et al. (2007). Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**, 222-34.

Girelli, D., Russo, C., Ferraresi, P., Olivieri, O., Pinotti, M., Friso, S., Manzato, F., Mazzucco, A., Bernardi, F. and Corrocher, R. (2000). Polymorphisms in the factor VII gene and the risk of myocardial infarction in patients with coronary artery disease. *N Engl J Med* **343**, 774-80.

Gironella, M., Folch-Puy, E., LeGoffic, A., Garcia, S., Christa, L., Smith, A., Tebar, L., Hunt, S. P., Bayne, R., Smith, A. J. et al. (2007). Experimental acute pancreatitis in PAP/HIP knock-out mice. *Gut* **56**, 1091-7.

Goessling, W., North, T. E. and Zon, L. I. (2007). New waves of discovery: modeling cancer in zebrafish. *J Clin Oncol* **25**, 2473-9.

Golan-Mashiach, M., Dazard, J. E., Gerecht-Nir, S., Amariglio, N., Fisher, T., Jacob-Hirsch, J., Bielorai, B., Osenberg, S., Barad, O., Getz, G. et al. (2005). Design principle of gene expression used by human stem cells: implication for pluripotency. *Faseb J* **19**, 147-9.

Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi, M., Artzt, K., Farrington, S. et al. (2002). Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nat Genet* **31**, 135-40.

Gonzalez, E. M., Mongiat, M., Slater, S. J., Baffa, R. and Iozzo, R. V. (2003). A novel interaction between perlecan protein core and progranulin: potential effects on tumor growth. *J Biol Chem* **278**, 38113-6.

Gradwohl, G., Dierich, A., LeMeur, M. and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc Natl Acad Sci U S A* **97**, 1607-11.

Graham, A., Okabe, M. and Quinlan, R. (2005). The role of the endoderm in the development and evolution of the pharyngeal arches. *J Anat* **207**, 479-87.

Gu, X. and Huang, W. (2002). Testing the parsimony test of genome duplications: a counterexample. *Genome Res* **12**, 1-2.

Gurney, M. E., Pu, H., Chiu, A. Y., Dal Canto, M. C., Polchow, C. Y., Alexander, D. D., Caliendo, J., Hentati, A., Kwon, Y. W., Deng, H. X. et al. (1994). Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science* **264**, 1772-5.

Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., Kane, D. A., Odenthal, J., van Eeden, F. J., Jiang, Y. J., Heisenberg, C. P. et al. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development* **123**, 1-36.

Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J., Cusick, M. E., Roth, F. P. et al. (2004). Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**, 88-93.

Hanington, P. C., Barreda, D. R. and Belosevic, M. (2006). A novel hematopoietic granulin induces proliferation of goldfish (Carassius auratus L.) macrophages. *J Biol Chem*.

Harada, A., Oguchi, K., Okabe, S., Kuno, J., Terada, S., Ohshima, T., Sato-Yoshitake, R., Takei, Y., Noda, T. and Hirokawa, N. (1994). Altered microtubule organization in small-calibre axons of mice lacking tau protein. *Nature* **369**, 488-91.

He, Z., Ismail, A., Kriazhev, L., Sadvakassova, G. and Bateman, A. (2002). Progranulin (PC-cell-derived growth factor/acrogranin) regulates invasion and cell survival. *Cancer Res* **62**, 5590-6.

He, Z., Ong, C. H., Halper, J. and Bateman, A. (2003). Progranulin is a mediator of the wound response. *Nat Med* **9**, 225-9.

Heasman, J. (2002). Morpholino oligos: making sense of antisense? *Dev Biol* **243**, 209-14.

Hennig, W. (2004). The revolution of the biology of the genome. *Cell Res* **14**, 1-7.

Hickey, M. A. and Chesselet, M. F. (2003). The use of transgenic and knock-in mice to study Huntington's disease. *Cytogenet Genome Res* **100**, 276-86.

Higashijima, S., Hotta, Y. and Okamoto, H. (2000). Visualization of cranial motor neurons in live transgenic zebrafish expressing green fluorescent protein under the control of the islet-1 promoter/enhancer. *J Neurosci* **20**, 206-18.

Ho, R. K. (1992). Cell movements and cell fate during zebrafish gastrulation. *Dev Suppl*, 65-73.

Ho, R. K. and Kimmel, C. B. (1993). Commitment of cell fate in the early zebrafish embryo. *Science* **261**, 109-11.

Hoegg, S., Brinkmann, H., Taylor, J. S. and Meyer, A. (2004). Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *J Mol Evol* **59**, 190-203.

Hoque, M., Tian, B., Mathews, M. B. and Pe'ery, T. (2005). Granulin and granulin repeats interact with the Tat.P-TEFb complex and inhibit Tat transactivation. *J Biol Chem* **280**, 13648-57.

Hoque, M., Young, T. M., Lee, C. G., Serrero, G., Mathews, M. B. and Pe'ery, T. (2003). The growth factor granulin interacts with cyclin T1 and modulates P-TEFb-dependent transcription. *Mol Cell Biol* **23**, 1688-702.

Horne-Badovinac, S., Lin, D., Waldron, S., Schwarz, M., Mbamalu, G., Pawson, T., Jan, Y., Stainier, D. Y. and Abdelilah-Seyfried, S. (2001). Positional cloning of heart and soul reveals multiple roles for PKC lambda in zebrafish organogenesis. *Curr Biol* **11**, 1492-502.

Horton, A. C., Mahadevan, N. R., Ruvinsky, I. and Gibson-Brown, J. J. (2003). Phylogenetic analyses alone are insufficient to determine whether genome duplication(s) occurred during early vertebrate evolution. *J Exp Zoolog B Mol Dev Evol* **299**, 41-53.

Hrabal, R., Chen, Z., James, S., Bennett, H. P. and Ni, F. (1996). The hairpin stack fold, a novel protein architecture for a new family of protein growth factors. *Nat Struct Biol* **3**, 747-52.

Hrabe de Angelis, M. H., Flaswinkel, H., Fuchs, H., Rathkolb, B., Soewarto, D., Marschall, S., Heffner, S., Pargent, W., Wuensch, K., Jung, M. et al. (2000). Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nat Genet* **25**, 444-7.

Huala, E., Dickerman, A. W., Garcia-Hernandez, M., Weems, D., Reiser, L., LaFond, F., Hanley, D., Kiphart, D., Zhuang, M., Huang, W. et al. (2001). The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res* **29**, 102-5.

Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T. et al. (2002). The Ensembl genome database project. *Nucleic Acids Res* **30**, 38-41.

Hubbard, T. J., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T. et al. (2007). Ensembl 2007. *Nucleic Acids Res* **35**, D610-7.

Huey, E. D., Grafman, J., Wassermann, E. M., Pietrini, P., Tierney, M. C., Ghetti, B., Spina, S., Baker, M., Hutton, M., Elder, J. W. et al. (2006). Characteristics of frontotemporal dementia patients with a Progranulin mutation. *Ann Neurol* **60**, 374-80.

Hughes, A. L. (1999). Phylogenies of developmentally important proteins do not support the hypothesis of two rounds of genome duplication early in vertebrate history. *J Mol Evol* **48**, 565-76.

Hughes, A. L. and Friedman, R. (2003). 2R or not 2R: testing hypotheses of genome duplication in early vertebrates. *J Struct Funct Genomics* **3**, 85-93.

Inc., Z. (http://zenemark.znomics.com/index.html). dbZeneMark, a Database of Retroviral Insertions in the Zebrafish Genome. In *Proc Natl Acad Sci U S A*, vol. 103 (ed., pp. 17284-9.

Itoh, N. and Ornitz, D. M. (2004). Evolution of the Fgf and Fgfr gene families. *Trends Genet* **20**, 563-9.

Jacobs, P. A., Baikie, A. G., Court Brown, W. M. and Strong, J. A. (1959). The somatic chromosomes in mongolism. *Lancet* **1**, 710.

Jaillon, O., Aury, J. M., Brunet, F., Petit, J. L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A. et al. (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **431**, 946-57.

Jarvis, D. L. and Finn, E. E. (1996). Modifying the insect cell N-glycosylation pathway with immediate early baculovirus expression vectors. *Nat Biotechnol* **14**, 1288-92.

Jelesko, J. G., Carter, K., Thompson, W., Kinoshita, Y. and Gruissem, W. (2004). Meiotic recombination between paralogous RBCSB genes on sister chromatids of Arabidopsis thaliana. *Genetics* **166**, 947-57.

Jin, P., Alisch, R. S. and Warren, S. T. (2004). RNA and microRNAs in fragile X mental retardation. *Nat Cell Biol* **6**, 1048-53.

Jing, X. H., Zhou, S. M., Wang, W. Q. and Chen, Y. (2006). Mechanisms underlying long- and short-range nodal signaling in Zebrafish. *Mech Dev* **123**, 388-94.

Jones, M. B., Michener, C. M., Blanchette, J. O., Kuznetsov, V. A., Raffeld, M., Serrero, G., Emmert-Buck, M. R., Petricoin, E. F., Krizman, D. B., Liotta, L. A. et al. (2003). The granulin-epithelin precursor/PC-cell-derived growth factor is a growth factor for epithelial ovarian cancer. *Clin Cancer Res* **9**, 44-51.

Josephs, K. A., Ahmed, Z., Katsuse, O., Parisi, J. F., Boeve, B. F., Knopman, D. S., Petersen, R. C., Davies, P., Duara, R., Graff-Radford, N. R. et al. (2007). Neuropathologic features of frontotemporal lobar degeneration with ubiquitin-positive inclusions with progranulin gene (PGRN) mutations. *J Neuropathol Exp Neurol* **66**, 142-51.

Jowsey, J., Riggs, B. L., Kelly, P. J. and Hoffman, D. L. (1978). Calcium and salmon calcitonin in treatment of osteoporosis. *J Clin Endocrinol Metab* **47**, 633-9.

Julien, J. P. and Kriz, J. (2006). Transgenic mouse models of amyotrophic lateral sclerosis. *Biochim Biophys Acta* **1762**, 1013-24.

Kadyk, L. C. and Hartwell, L. H. (1992). Sister chromatids are preferred over homologs as substrates for recombinational repair in Saccharomyces cerevisiae. *Genetics* **132**, 387-402.

Kalev-Zylinska, M. L., Horsfield, J. A., Flores, M. V., Postlethwait, J. H., Vitas, M. R., Baas, A. M., Crosier, P. S. and Crosier, K. E. (2002). Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development* **129**, 2015-30.

Kamrava, M., Simpkins, F., Alejandro, E., Michener, C., Meltzer, E. and Kohn, E. C. (2005). Lysophosphatidic acid and endothelin-induced proliferation of ovarian cancer cell lines is mitigated by neutralization of granulin-epithelin precursor (GEP), a prosurvival factor for ovarian cancer. *Oncogene* **24**, 7084-93.

Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* **95**, 5752-6.

Kasper, L. H., Fukuyama, T., Biesen, M. A., Boussouar, F., Tong, C., de Pauw, A., Murray, P. J., van Deursen, J. M. and Brindle, P. K. (2006). Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Mol Cell Biol* **26**, 789-809.

Kayasuga, Y., Chiba, S., Suzuki, M., Kikusui, T., Matsuwaki, T., Yamanouchi, K., Kotaki, H., Horai, R., Iwakura, Y. and Nishihara, M. (2007). Alteration of behavioural phenotype in mice by targeted disruption of the progranulin gene. *Behav Brain Res* **185**, 110-8.

Kazazian, H. H., Jr. (2004). Mobile elements: drivers of genome evolution. *Science* **303**, 1626-32.

Keng, V. W., Yagi, H., Ikawa, M., Nagano, T., Myint, Z., Yamada, K., Tanaka, T., Sato, A., Muramatsu, I., Okabe, M. et al. (2000). Homeobox gene Hex is essential for onset of mouse embryonic liver development and differentiation of the monocyte lineage. *Biochem Biophys Res Commun* **276**, 1155-61.

Kenny, P. A., Liston, E. M. and Higgins, D. G. (1999). Molecular evolution of immunoglobulin and fibronectin domains in titin and related muscle proteins. *Gene* **232**, 11-23.

Kijimoto, T., Watanabe, M., Fujimura, K., Nakazawa, M., Murakami, Y., Kuratani, S., Kohara, Y., Gojobori, T. and Okada, N. (2005). cimp1, a novel astacin family metalloproteinase gene from East African cichlids, is differentially expressed between species during growth. *Mol Biol Evol* **22**, 1649-60.

Kim, H. J., Schleiffarth, J. R., Jessurun, J., Sumanas, S., Petryk, A., Lin, S. and Ekker, S. C. (2005). Wnt5 signaling in vertebrate pancreas development. *BMC Biol* **3**, 23.

Kim, H. J., Sumanas, S., Palencia-Desai, S., Dong, Y., Chen, J. N. and Lin, S. (2006). Genetic analysis of early endocrine pancreas formation in zebrafish. *Mol Endocrinol* **20**, 194-203.

Kim, S. H. and Yi, S. V. (2006). Correlated asymmetry of sequence and functional divergence between duplicate proteins of Saccharomyces cerevisiae. *Mol Biol Evol* **23**, 1068-75.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310.

Kimura, N., Yonekura, H., Okamoto, H. and Nagura, H. (1992). Expression of human regenerating gene mRNA and its product in normal and neoplastic human pancreas. *Cancer* **70**, 1857-63.

Knight, C. A., DeVries, A. L. and Oolman, L. D. (1984). Fish antifreeze protein and the freezing and recrystallization of ice. *Nature* **308**, 295-6.

Knight, R. D., Javidan, Y., Zhang, T., Nelson, S. and Schilling, T. F. (2005). AP2-dependent signals from the ectoderm regulate craniofacial development in the zebrafish embryo. *Development* **132**, 3127-38.

Kobayashi, S., Akiyama, T., Nata, K., Abe, M., Tajima, M., Shervani, N. J., Unno, M., Matsuno, S., Sasaki, H., Takasawa, S. et al. (2000). Identification of a receptor for reg (regenerating gene) protein, a pancreatic beta-cell regeneration factor. *J Biol Chem* **275**, 10723-6.

Kong, W. J., Zhang, S. L., Chen, X., Zhang, S., Wang, Y. J., Zhang, D. and Sun, Y. (2007). PC cell-derived growth factor overexpression promotes proliferation and survival of laryngeal carcinoma. *Anticancer Drugs* **18**, 29-40.

Kontges, G. and Lumsden, A. (1996). Rhombencephalic neural crest segmentation is preserved throughout craniofacial ontogeny. *Development* **122**, 3229-42.

Kopinke, D., Sasine, J., Swift, J., Stephens, W. Z. and Piotrowski, T. (2006). Retinoic acid is required for endodermal pouch morphogenesis and not for pharyngeal endoderm specification. *Dev Dyn* **235**, 2695-2709.

Korzh, S., Emelyanov, A. and Korzh, V. (2001). Developmental analysis of ceruloplasmin gene and liver formation in zebrafish. *Mech Dev* **103**, 137-9.

Kramer-Zucker, A. G., Olale, F., Haycraft, C. J., Yoder, B. K., Schier, A. F. and Drummond, I. A. (2005). Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis. *Development* **132**, 1907-21.

Krapp, A., Knofler, M., Ledermann, B., Burki, K., Berney, C., Zoerkler, N., Hagenbuchle, O. and Wellauer, P. K. (1998). The bHLH protein PTF1-p48 is essential for the formation of the exocrine and the correct spatial organization of the endocrine pancreas. *Genes Dev* **12**, 3752-63.

Krauss, S., Concordet, J. P. and Ingham, P. W. (1993). A functionally conserved homolog of the Drosophila segment polarity gene hh is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* **75**, 1431-44.

Kulesa, P., Ellies, D. L. and Trainor, P. A. (2004). Comparative analysis of neural crest cell death, migration, and function during vertebrate embryogenesis. *Dev Dyn* **229**, 14-29.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.

Lam, W. L., Lee, T. S. and Gilbert, W. (1996). Active transposition in zebrafish. *Proc Natl Acad Sci U S A* **93**, 10870-5.

Lander, E. S. Linton, L. M. Birren, B. Nusbaum, C. Zody, M. C. Baldwin, J. Devon, K. Dewar, K. Doyle, M. FitzHugh, W. et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Langenau, D. M., Feng, H., Berghmans, S., Kanki, J. P., Kutok, J. L. and Look, A. T. (2005). Cre/lox-regulated transgenic zebrafish model with conditional myc-induced T cell acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A* **102**, 6068-73.

Langenau, D. M., Traver, D., Ferrando, A. A., Kutok, J. L., Aster, J. C., Kanki, J. P., Lin, S., Prochownik, E., Trede, N. S., Zon, L. I. et al. (2003). Myc-induced T cell leukemia in transgenic zebrafish. *Science* **299**, 887-90.

Lawson, N. D. and Weinstein, B. M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol* **248**, 307-18.

Le Ber, I., van der Zee, J., Hannequin, D., Gijselinck, I., Campion, D., Puel, M., Laquerriere, A., De Pooter, T., Camuzat, A., Van den Broeck, M. et al. (2007). Progranulin null mutations in both sporadic and familial frontotemporal dementia. *Hum Mutat*.

Lee, J. E., Wu, S. F., Goering, L. M. and Dorsky, R. I. (2006). Canonical Wnt signaling through Lef1 is required for hypothalamic neurogenesis. *Development* **133**, 4451-61.

Lee, K. H., Xu, Q. and Breitbart, R. E. (1996). A new tinman-related gene, nkx2.7, anticipates the expression of nkx2.5 and nkx2.3 in zebrafish heart and pharyngeal endoderm. *Dev Biol* **180**, 722-31.

Lejeune, J., Turpin, R. and Gautier, M. (1959). [Mongolism; a chromosomal disease (trisomy).]. *Bull Acad Natl Med* **143**, 256-65.

Lemmens, R., Van Hoecke, A., Hersmus, N., Geelen, V., D'Hollander, I., Thijs, V., Van Den Bosch, L., Carmeliet, P. and Robberecht, W. (2007). Overexpression of mutant superoxide dismutase 1 causes a motor axonopathy in the zebrafish. *Hum Mol Genet*.

Leslie, G. A. and Clem, L. W. (1969). Production of anti-hapten antibodies by several classes of lower vertebrates. *J Immunol* **103**, 613-7.

Leverenz, J. B., Yu, C. E., Montine, T. J., Steinbart, E., Bekris, L. M., Zabetian, C., Kwong, L. K., Lee, V. M., Schellenberg, G. D. and Bird, T. D. (2007). A novel progranulin mutation associated with variable clinical presentation and tau, TDP43 and alpha-synuclein pathology. *Brain* **130**, 1360-74.

Levine, M. S., Cepeda, C., Hickey, M. A., Fleming, S. M. and Chesselet, M. F. (2004). Genetic mouse models of Huntington's and Parkinson's diseases: illuminating but imperfect. *Trends Neurosci* **27**, 691-7.

Lewis, J. L., Bonner, J., Modrell, M., Ragland, J. W., Moon, R. T., Dorsky, R. I. and Raible, D. W. (2004). Reiterated Wnt signaling during zebrafish neural crest development. *Development* **131**, 1299-308.

Lewis, K. E. and Eisen, J. S. (2001). Hedgehog signaling is required for primary motoneuron induction in zebrafish. *Development* **128**, 3485-95.

Li, S. and Wilkinson, M. F. (1998). Nonsense surveillance in lymphocytes? *Immunity* **8**, 135-41.

Li, Y. C., Korol, A. B., Fahima, T. and Nevo, E. (2004). Microsatellites within genes: structure, function, and evolution. *Mol Biol Evol* **21**, 991-1007.

Liao, E. C., Paw, B. H., Oates, A. C., Pratt, S. J., Postlethwait, J. H. and Zon, L. I. (1998). SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish. *Genes Dev* **12**, 621-6.

Liau, L. M., Lallone, R. L., Seitz, R. S., Buznikov, A., Gregg, J. P., Kornblum, H. I., Nelson, S. F. and Bronstein, J. M. (2000). Identification of a human glioma-associated growth factor gene, granulin, using differential immuno-absorption. *Cancer Res* **60**, 1353-60.

Lin, J. W., Biankin, A. V., Horb, M. E., Ghosh, B., Prasad, N. B., Yee, N. S., Pack, M. A. and Leach, S. D. (2004a). Differential requirement for ptf1a in endocrine and exocrine lineages of developing zebrafish pancreas. *Dev Biol* **274**, 491-503.

Lin, J. W., Biankin, A. V., Horb, M. E., Ghosh, B., Prasad, N. B., Yee, N. S., Pack, M. A. and Leach, S. D. (2004b). Differential requirement for ptf1a in endocrine and exocrine lineages of developing zebrafish pancreas. *Dev Biol* **270**, 474-86.

Lindblad-Toh, K. Wade, C. M. Mikkelsen, T. S. Karlsson, E. K. Jaffe, D. B. Kamal, M. Clamp, M. Chang, J. L. Kulbokas, E. J., 3rd Zody, M. C. et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803-19.

Liu, J. P., Baker, J., Perkins, A. S., Robertson, E. J. and Efstratiadis, A. (1993). Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). *Cell* **75**, 59-72.

Liu, W. Y., Wang, Y., Qin, Y., Wang, Y. P. and Zhu, Z. Y. (2007a). Site-Directed Gene Integration in Transgenic Zebrafish Mediated by Cre Recombinase Using a Combination of Mutant Lox Sites. *Mar Biotechnol (NY)*.

Liu, W. Y., Wang, Y., Sun, Y. H., Wang, Y., Wang, Y. P., Chen, S. P. and Zhu, Z. Y. (2005). Efficient RNA interference in zebrafish embryos using siRNA synthesized with SP6 RNA polymerase. *Dev Growth Differ* **47**, 323-31.

Liu, Y., Xi, L., Liao, G., Wang, W., Tian, X., Wang, B., Chen, G., Han, Z., Wu, M., Wang, S. et al. (2007b). Inhibition of PC cell-derived growth factor (PCDGF)/granulin-epithelin precursor (GEP) decreased cell proliferation and invasion through downregulation of cyclin D and CDK4 and inactivation of MMP-2. *BMC Cancer* **7**, 22.

Looman, C., Abrink, M., Mark, C. and Hellman, L. (2002). KRAB zinc finger proteins: an analysis of the molecular mechanisms governing their increase in numbers and complexity during evolution. *Mol Biol Evol* **19**, 2118-30.

Lowe, L. A., Yamada, S. and Kuehn, M. R. (2001). Genetic dissection of nodal function in patterning the mouse embryo. *Development* **128**, 1831-43.

Lu, R. and Serrero, G. (2001). Mediation of estrogen mitogenic effect in human breast cancer MCF-7 cells by PC-cell-derived growth factor (PCDGF/granulin precursor). *Proc Natl Acad Sci U S A* **98**, 142-7.

Lumsden, A. L., Henshall, T. L., Dayan, S., Lardelli, M. T. and Richards, R. I. (2007). Huntingtin-deficient zebrafish exhibit defects in iron utilization and development. *Hum Mol Genet*.

Lundin, L. G., Larhammar, D. and Hallbook, F. (2003). Numerous groups of chromosomal regional paralogies strongly indicate two genome doublings at the root of the vertebrates. *J Struct Funct Genomics* **3**, 53-63.

Malka, D., Vasseur, S., Bodeker, H., Ortiz, E. M., Dusetti, N. J., Verrando, P., Dagorn, J. C. and Iovanna, J. L. (2000). Tumor necrosis factor alpha triggers antiapoptotic mechanisms in rat pancreatic cells through pancreatitis-associated protein I activation. *Gastroenterology* **119**, 816-28.

Mangos, S., Vanderbeld, B., Krawetz, R., Sudol, K. and Kelly, G. M. (2001). Ran binding protein RanBP1 in zebrafish embryonic development. *Mol Reprod Dev* **59**, 235-48.

Mann, M. and Wilm, M. (1994). Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem* **66**, 4390-9.

Marcotte, E. M., Pellegrini, M., Yeates, T. O. and Eisenberg, D. (1999). A census of protein repeats. *J Mol Biol* **293**, 151-60.

Martin, J. Han, C. Gordon, L. A. Terry, A. Prabhakar, S. She, X. Xie, G. Hellsten, U. Chan, Y. M. Altherr, M. et al. (2004). The sequence and analysis of duplication-rich human chromosome 16. *Nature* **432**, 988-94.

Masellis, M., Momeni, P., Meschino, W., Heffner, R., Jr., Elder, J., Sato, C., Liang, Y., St George-Hyslop, P., Hardy, J., Bilbao, J. et al. (2006). Novel splicing mutation in the progranulin gene causing familial corticobasal syndrome. *Brain* **129**, 3115-23.

Matsumura, N., Mandai, M., Miyanishi, M., Fukuhara, K., Baba, T., Higuchi, T., Kariya, M., Takakura, K. and Fujii, S. (2006). Oncogenic property of acrogranin in human uterine leiomyosarcoma: direct evidence of genetic contribution in in vivo tumorigenesis. *Clin Cancer Res* **12**, 1402-11.

Mayer, A. N. and Fishman, M. C. (2003). Nil per os encodes a conserved RNA recognition motif protein required for morphogenesis and cytodifferentiation of digestive organs in zebrafish. *Development* **130**, 3917-28.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

McGowan, E., Eriksen, J. and Hutton, M. (2006). A decade of modeling Alzheimer's disease in transgenic mice. *Trends Genet* **22**, 281-9.

McLysaght, A., Hokamp, K. and Wolfe, K. H. (2002). Extensive genomic duplication during early chordate evolution. *Nat Genet* **31**, 200-4.

Medghalchi, S. M., Frischmeyer, P. A., Mendell, J. T., Kelly, A. G., Lawler, A. M. and Dietz, H. C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum Mol Genet* **10**, 99-105.

Melrose, H. L., Lincoln, S. J., Tyndall, G. M. and Farrer, M. J. (2006). Parkinson's disease: a rethink of rodent models. *Exp Brain Res* **173**, 196-204.

Mesulam, M., Johnson, N., Krefft, T. A., Gass, J. M., Cannon, A. D., Adamson, J. L., Bigio, E. H., Weintraub, S., Dickson, D. W., Hutton, M. L. et al. (2007). Progranulin Mutations in Primary Progressive Aphasia: The PPA1 and PPA3 Families. *Arch Neurol* **64**, 43-7.

Meyer, A. and Van de Peer, Y. (2005). From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays* **27**, 937-45.

Milewski, W. M., Duguay, S. J., Chan, S. J. and Steiner, D. F. (1998). Conservation of PDX-1 structure, function, and expression in zebrafish. *Endocrinology* **139**, 1440-9.

Miller, W. J. and Capy, P. (2006). Applying mobile genetic elements for genome analysis and evolution. *Mol Biotechnol* **33**, 161-74.

Miralles, F., Battelino, T., Czernichow, P. and Scharfmann, R. (1998). TGF-beta plays a key role in morphogenesis of the pancreatic islets of Langerhans by controlling the activity of the matrix metalloproteinase MMP-2. *J Cell Biol* **143**, 827-36.

Mohler, J. (1988). Requirements for hedgehod, a segmental polarity gene, in patterning larval and adult cuticle of Drosophila. *Genetics* **120**, 1061-72.

Molina, G. A., Watkins, S. C. and Tsang, M. (2007). Generation of FGF reporter transgenic zebrafish and their utility in chemical screens. *BMC Dev Biol* **7**, 62.

Monami, G., Gonzalez, E. M., Hellman, M., Gomella, L. G., Baffa, R., Iozzo, R. V. and Morrione, A. (2006). Proepithelin promotes migration and invasion of 5637 bladder cancer cells through the activation of ERK1/2 and the formation of a paxillin/FAK/ERK complex. *Cancer Res* **66**, 7103-10.

Moorman, S. J. (2001). Anatomical nomenclature for zebrafish. *Zebrafish Science Monitor*, 7-8.

Mudumana, S. P., Wan, H., Singh, M., Korzh, V. and Gong, Z. (2004). Expression analyses of zebrafish transferrin, ifabp, and elastaseB mRNAs as differentiation markers for the three major endodermal organs: liver, intestine, and exocrine pancreas. *Dev Dyn* **230**, 165-73.

Mukherjee, O., Pastor, P., Cairns, N. J., Chakraverty, S., Kauwe, J. S., Shears, S., Behrens, M. I., Budde, J., Hinrichs, A. L., Norton, J. et al. (2006). HDDD2 is a familial frontotemporal lobar degeneration with ubiquitin-positive, tau-negative inclusions caused by a missense mutation in the signal peptide of progranulin. *Ann Neurol* **60**, 314-22.

Mullins, M. C., Hammerschmidt, M., Haffter, P. and Nusslein-Volhard, C. (1994). Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. *Curr Biol* **4**, 189-202.

Murphey, R. D., Stern, H. M., Straub, C. T. and Zon, L. I. (2006). A chemical genetic screen for cell cycle inhibitors in zebrafish embryos. *Chem Biol Drug Des* **68**, 213-9.

Murphey, R. D. and Zon, L. I. (2006). Small molecule screening in the zebrafish. *Methods* **39**, 255-61.

Myszka, D. G. (1997). Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Curr Opin Biotechnol* **8**, 50-7.

Myszka, D. G. (2000). Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE. *Methods Enzymol* **323**, 325-40.

Nasevicius, A. and Ekker, S. C. (2000). Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* **26**, 216-20.

National Center for Biotechnology Information (U.S.). (2002). The NCBI handbook, (ed. [Bethesda, Md.?]: NCBI.

Nedelkov, D. (2007). Development of surface plasmon resonance mass spectrometry array platform. *Anal Chem* **79**, 5987-90.

Nelson, J. (1994). Fishes of the World, (ed. New York: Wiley.

Neuhauss, S. C., Solnica-Krezel, L., Schier, A. F., Zwartkruis, F., Stemple, D. L., Malicki, J., Abdelilah, S., Stainier, D. Y. and Driever, W. (1996). Mutations affecting craniofacial development in zebrafish. *Development* **123**, 357-67.

Niederreither, K., Vermot, J., Le Roux, I., Schuhbaur, B., Chambon, P. and Dolle, P. (2003). The regional pattern of retinoic acid synthesis by RALDH2 is essential for the development of posterior pharyngeal arches and the enteric nervous system. *Development* **130**, 2525-34.

Nissen, R. M., Yan, J., Amsterdam, A., Hopkins, N. and Burgess, S. M. (2003). Zebrafish foxi one modulates cellular responses to Fgf signaling required for the integrity of ear and jaw patterning. *Development* **130**, 2543-54.

Nolan, P. M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I. C., Vizor, L., Brooker, D., Whitehill, E. et al. (2000). A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat Genet* **25**, 440-3.

Oates, A. C., Bruce, A. E. and Ho, R. K. (2000). Too much interference: injection of double-stranded RNA has nonspecific effects in the zebrafish embryo. *Dev Biol* **224**, 20-8.

Ohno, S., Wolf, U. and Atkin, N. B. (1968). Evolution from fish to mammals by gene duplication. *Hereditas* **59**, 169-87.

Ong, C. H., He, Z., Kriazhev, L., Shan, X., Palfree, R. G. and Bateman, A. (2006). Regulation of progranulin expression in myeloid cells. *Am J Physiol Regul Integr Comp Physiol*.

Onoda, F., Seki, M., Wang, W. and Enomoto, T. (2004). The hyper unequal sister chromatid recombination in an sgs1 mutant of budding yeast requires MSH2. *DNA Repair (Amst)* **3**, 1355-62.

Orkin, S. H. and Zon, L. I. (1997). Genetics of erythropoiesis: induced mutations in mice and zebrafish. *Annu Rev Genet* **31**, 33-60.

Ose, T., Kadowaki, Y., Fukuhara, H., Kazumori, H., Ishihara, S., Udagawa, J., Otani, H., Takasawa, S., Okamoto, H. and Kinoshita, Y. (2007). Reg I-knockout mice reveal its role in regulation of cell growth that is required in generation and maintenance of the villous structure of small intestine. *Oncogene* **26**, 349-59.

Oue, N., Mitani, Y., Aung, P. P., Sakakura, C., Takeshima, Y., Kaneko, M., Noguchi, T., Nakayama, H. and Yasui, W. (2005). Expression and localization of Reg IV in human neoplastic and non-neoplastic tissues: Reg IV expression is associated with intestinal and neuroendocrine differentiation in gastric adenocarcinoma. *J Pathol* **207**, 185-98.

Pack, M., Solnica-Krezel, L., Malicki, J., Neuhauss, S. C., Schier, A. F., Stemple, D. L., Driever, W. and Fishman, M. C. (1996). Mutations affecting development of zebrafish digestive organs. *Development* **123**, 321-8.

Padgett, R. W., St Johnston, R. D. and Gelbart, W. M. (1987). A transcript from a Drosophila pattern gene predicts a protein homologous to the transforming growth factor-beta family. *Nature* **325**, 81-4.

Paillusson, A., Hirschi, N., Vallan, C., Azzalin, C. M. and Muhlemann, O. (2005). A GFP-based reporter system to monitor nonsense-mediated mRNA decay. *Nucleic Acids Res* **33**, e54.

Pan, C. X., Kinch, M. S., Kiener, P. A., Langermann, S., Serrero, G., Sun, L., Corvera, J., Sweeney, C. J., Li, L., Zhang, S. et al. (2004). PC cell-derived growth factor expression in prostatic intraepithelial neoplasia and prostatic adenocarcinoma. *Clin Cancer Res* **10**, 1333-7.

Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J., Herwig, R., Vingron, M. and Lehrach, H. (2003). New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. *Genome Res* **13**, 1056-66.

Pawson, T. (2007). Dynamic control of signaling by modular adaptor proteins. *Curr Opin Cell Biol* **19**, 112-6.

Pawson, T. and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445-52.

Pearson, C. E. (2003). Slipping while sleeping? Trinucleotide repeat expansions in germ cells. *Trends Mol Med* **9**, 490-5.

Pearson, C. E., Nichol Edamura, K. and Cleary, J. D. (2005). Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet* **6**, 729-42.

Pearson, C. E., Tam, M., Wang, Y. H., Montgomery, S. E., Dar, A. C., Cleary, J. D. and Nichol, K. (2002). Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: slipped-out repeats and slip-out junctions. *Nucleic Acids Res* **30**, 4534-47.

Perkins, D. N., Pappin, D. J., Creasy, D. M. and Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-67.

Peyrieras, N., Strahle, U. and Rosa, F. (1998). Conversion of zebrafish blastomeres to an endodermal fate by TGF-beta-related signaling. *Curr Biol* **8**, 783-6.

Piotrowski, T., Ahn, D. G., Schilling, T. F., Nair, S., Ruvinsky, I., Geisler, R., Rauch, G. J., Haffter, P., Zon, L. I., Zhou, Y. et al. (2003). The zebrafish van gogh mutation disrupts tbx1, which is involved in the DiGeorge deletion syndrome in humans. *Development* **130**, 5043-52.

Piotrowski, T. and Nusslein-Volhard, C. (2000). The endoderm plays an important role in patterning the segmented pharyngeal region in zebrafish (Danio rerio). *Dev Biol* **225**, 339-56.

Politano, L., Nigro, G., Nigro, V., Piluso, G., Papparella, S., Paciello, O. and Comi, L. I. (2003). Gentamicin administration in Duchenne patients with premature stop codon. Preliminary results. *Acta Myol* **22**, 15-21.

Pollard-Knight, D., Hawkins, E., Yeung, D., Pashby, D. P., Simpson, M., McDougall, A., Buckle, P. and Charles, S. A. (1990). Immunoassays and nucleic acid detection with a biosensor based on surface plasmon resonance. *Ann Biol Clin (Paris)* **48**, 642-6.

Postlethwait, J. H., Woods, I. G., Ngo-Hazelett, P., Yan, Y. L., Kelly, P. D., Chu, F., Huang, H., Hill-Force, A. and Talbot, W. S. (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res* **10**, 1890-902.

Postlethwait, J. H., Yan, Y. L., Gates, M. A., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E. S., Force, A., Gong, Z. et al. (1998). Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* **18**, 345-9.

Potter, J. D. (2007). Morphogens, morphostats, microarchitecture and malignancy. *Nat Rev Cancer* **7**, 464-74.

Prak, E. T. and Kazazian, H. H., Jr. (2000). Mobile elements and the human genome. *Nat Rev Genet* **1**, 134-44.

Qin, J., Diaz-Cueto, L., Schwarze, J. E., Takahashi, Y., Imai, M., Isuzugawa, K., Yamamoto, S., Chang, K. T., Gerton, G. L. and Imakawa, K. (2005a). Effects of progranulin on blastocyst hatching and subsequent adhesion and outgrowth in the mouse. *Biol Reprod* **73**, 434-42.

Qin, J., Diaz-Cueto, L., Schwarze, J. E., Takahashi, Y., Imai, M., Isuzugawa, K., Yamamoto, S., Chang, K. T., Gerton, G. L. and Imakawa, K. (2005b). Effects of Progranulin on Blastocyst Hatching and Subsequent Adhesion and Outgrowth in the Mouse. *Biol Reprod*.

Qu, S., Tucker, S. C., Ehrlich, J. S., Levorse, J. M., Flaherty, L. A., Wisdom, R. and Vogt, T. F. (1998). Mutations in mouse Aristaless-like4 cause Strong's luxoid polydactyly. *Development* **125**, 2711-21.

Rabbitts, T. H., Forster, A., Dunnick, W. and Bentley, D. L. (1980). The role of gene deletion in the immunoglobulin heavy chain switch. *Nature* **283**, 351-6.

Raible, D. W., Wood, A., Hodsdon, W., Henion, P. D., Weston, J. A. and Eisen, J. S. (1992). Segregation and early dispersal of neural crest cells in the embryonic zebrafish. *Dev Dyn* **195**, 29-42.

Reich, D. E., Schaffner, S. F., Daly, M. J., McVean, G., Mullikin, J. C., Higgins, J. M., Richter, D. J., Lander, E. S. and Altshuler, D. (2002). Human genome sequence variation and the influence of gene history, mutation and recombination. *Nat Genet* **32**, 135-42.

Reiter, J. F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N. and Stainier, D. Y. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev* **13**, 2983-95.

Reiter, J. F., Kikuchi, Y. and Stainier, D. Y. (2001). Multiple roles for Gata5 in zebrafish endoderm formation. *Development* **128**, 125-35.

Remenyi, A., Good, M. C. and Lim, W. A. (2006). Docking interactions in protein kinase and phosphatase networks. *Curr Opin Struct Biol* **16**, 676-85.

Reynaud, D., Ravet, E., Titeux, M., Mazurier, F., Renia, L., Dubart-Kupperschmitt, A., Romeo, P. H. and Pflumio, F. (2005). SCL/TAL1 expression level regulates human hematopoietic stem cell self-renewal and engraftment. *Blood* **106**, 2318-28.

Rhodes, J., Hagen, A., Hsu, K., Deng, M., Liu, T. X., Look, A. T. and Kanki, J. P. (2005). Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev Cell* **8**, 97-108.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-2.

Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D. and Nusse, R. (1987). The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* **50**, 649-57.

Rivera, E. J., Goldin, A., Fulmer, N., Tavares, R., Wands, J. R. and de la Monte, S. M. (2005). Insulin and insulin-like growth factor expression and function deteriorate with progression of Alzheimer's disease: link to brain reductions in acetylcholine. *J Alzheimers Dis* **8**, 247-68.

Roberts, R. G., Gardner, R. J. and Bobrow, M. (1994). Searching for the 1 in 2,400,000: a review of dystrophin gene point mutations. *Hum Mutat* **4**, 1-11.

Rossner, M. and Yamada, K. M. (2004). What's in a picture? The temptation of image manipulation. *J Cell Biol* **166**, 11-5.

Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N. et al. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173-8.

Rutenberg, J. B., Fischer, A., Jia, H., Gessler, M., Zhong, T. P. and Mercola, M. (2006). Developmental patterning of the cardiac atrioventricular canal by Notch and Hairy-related transcription factors. *Development* **133**, 4381-90.

S., O. (1970). Evolution by gene duplication: Springer publishing.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning - A laboratory manual.: Cold Spring Harbor Laboratory Press.

Samollow, P. B., Kammerer, C. M., Mahaney, S. M., Schneider, J. L., Westenberger, S. J., VandeBerg, J. L. and Robinson, E. S. (2004). First-generation linkage map of the gray, short-tailed opossum, Monodelphis domestica, reveals genome-wide reduction in female recombination rates. *Genetics* **166**, 307-29.

Scherr, M., Chaturvedi, A., Battmer, K., Dallmann, I., Schultheis, B., Ganser, A. and Eder, M. (2006). Enhanced sensitivity to inhibition of SHP2, STAT5, and Gab2 expression in chronic myeloid leukemia (CML). *Blood* **107**, 3279-87.

Schilling, T. F. and Kimmel, C. B. (1994). Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development* **120**, 483-94.

Schilling, T. F., Piotrowski, T., Grandel, H., Brand, M., Heisenberg, C. P., Jiang, Y. J., Beuchle, D., Hammerschmidt, M., Kane, D. A., Mullins, M. C. et al. (1996). Jaw and branchial arch mutants in zebrafish I: branchial arches. *Development* **123**, 329-44.

Schlombs, K., Wagner, T. and Scheel, J. (2003). Site-1 protease is required for cartilage development in zebrafish. *Proc Natl Acad Sci U S A* **100**, 14024-9.

Schmid, B., Furthauer, M., Connors, S. A., Trout, J., Thisse, B., Thisse, C. and Mullins, M. C. (2000). Equivalent genetic roles for bmp7/snailhouse and bmp2b/swirl in dorsoventral pattern formation. *Development* **127**, 957-67.

Schulte-Merker, S., Ho, R. K., Herrmann, B. G. and Nusslein-Volhard, C. (1992). The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development* **116**, 1021-32.

Schymick, J., Yang, Y., Andersen, P., Vonsattel, J., Greenway, M., Momeni, P., Elder, J., Chio, A., Restagno, G., Robberecht, W. et al. (2007). Progranulin mutations and ALS or ALS-FTD phenotypes. *J Neurol Neurosurg Psychiatry*.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proc Biol Sci* **273**, 1987-98.

Sell, C., Dumenil, G., Deveaud, C., Miura, M., Coppola, D., DeAngelis, T., Rubin, R., Efstratiadis, A. and Baserga, R. (1994). Effect of a null mutation of the insulin-like growth factor I receptor gene on growth and transformation of mouse embryo fibroblasts. *Mol Cell Biol* **14**, 3604-12.

Serrero, G. and Ioffe, O. B. (2003). Expression of PC-cell-derived growth factor in benign and malignant human breast epithelium. *Hum Pathol* **34**, 1148-54.

Shafizadeh, E., Paw, B. H., Foott, H., Liao, E. C., Barut, B. A., Cope, J. J., Zon, L. I. and Lin, S. (2002). Characterization of zebrafish merlot/chablis as non-mammalian vertebrate models for severe congenital anemia due to protein 4.1 deficiency. *Development* **129**, 4359-70.

Shankaran, S. S., Capell, A., Hruscha, A. T., Fellerer, K., Neumann, M., Schmid, B. and Haass, C. (2008). Missense mutations in the progranulin gene linked to frontotemporal lobar degeneration with ubiquitin-immunoreactive inclusions reduce progranulin production and secretion. *J Biol Chem* **283**, 1744-53.

Shaw-Smith, C., Pittman, A. M., Willatt, L., Martin, H., Rickman, L., Gribble, S., Curley, R., Cumming, S., Dunn, C., Kalaitzopoulos, D. et al. (2006). Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet* **38**, 1032-7.

Shepard, J. L., Amatruda, J. F., Stern, H. M., Subramanian, A., Finkelstein, D., Ziai, J., Finley, K. R., Pfaff, K. L., Hersey, C., Zhou, Y. et al. (2005). A zebrafish bmyb mutation causes genome instability and increased cancer susceptibility. *Proc Natl Acad Sci U S A* **102**, 13194-9.

Shin, D., Shin, C. H., Tucker, J., Ober, E. A., Rentzsch, F., Poss, K. D., Hammerschmidt, M., Mullins, M. C. and Stainier, D. Y. (2007). Bmp and Fgf signaling are essential for liver specification in zebrafish. *Development* **134**, 2041-50.

Shoham, N., Cohen, L., Gazit, A. and Yaniv, A. (2003). The Tat protein of the caprine arthritis encephalitis virus interacts with the Notch2 EGF-like repeats and the epithelin/granulin precursor. *Intervirology* **46**, 239-44.

Sinnott, P., Collier, S., Costigan, C., Dyer, P. A., Harris, R. and Strachan, T. (1990). Genesis by meiotic unequal crossover of a de novo deletion that contributes to steroid 21-hydroxylase deficiency. *Proc Natl Acad Sci U S A* **87**, 2107-11.

Slotkin, R. K. and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**, 272-85.

Snowden, J. S., Pickering-Brown, S. M., Mackenzie, I. R., Richardson, A. M., Varma, A., Neary, D. and Mann, D. M. (2006). Progranulin gene mutations associated with frontotemporal dementia and progressive non-fluent aphasia. *Brain* **129**, 3091-102.

Sodergren, E. Weinstock, G. M. Davidson, E. H. Cameron, R. A. Gibbs, R. A. Angerer, R. C. Angerer, L. M. Arnone, M. I. Burgess, D. R. Burke, R. D. et al. (2006). The genome of the sea urchin Strongylocentrotus purpuratus. *Science* **314**, 941-52.

Sonnichsen, F. D., Sykes, B. D. and Davies, P. L. (1995). Comparative modeling of the three-dimensional structure of type II antifreeze protein. *Protein Sci* **4**, 460-71.

Sood, R., English, M. A., Jones, M., Mullikin, J., Wang, D. M., Anderson, M., Wu, D., Chandrasekharappa, S. C., Yu, J., Zhang, J. et al. (2006). Methods for reverse genetic screening in zebrafish by resequencing and TILLING. *Methods* **39**, 220-7.

Spina, S., Murrell, J. R., Huey, E. D., Wassermann, E. M., Pietrini, P., Baraibar, M. A., Barbeito, A. G., Troncoso, J. C., Vidal, R., Ghetti, B. et al. (2007). Clinicopathologic features of frontotemporal dementia with Progranulin sequence variation. *Neurology*.

Stalker, J., Gibbins, B., Meidl, P., Smith, J., Spooner, W., Hotz, H. R. and Cox, A. V. (2004). The Ensembl Web site: mechanics of a genome browser. *Genome Res* **14**, 951-5.

Stamm, S., Riethoven, J. J., Le Texier, V., Gopalakrishnan, C., Kumanduri, V., Tang, Y., Barbosa-Morais, N. L. and Thanaraj, T. A. (2006). ASD: a bioinformatics resource on alternative splicing. *Nucleic Acids Res* **34**, D46-55.

Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., Baker, A., Jonasdottir, A., Ingason, A., Gudnadottir, V. G. et al. (2005). A common inversion under selection in Europeans. *Nat Genet* **37**, 129-37.

Stern, H. M., Murphey, R. D., Shepard, J. L., Amatruda, J. F., Straub, C. T., Pfaff, K. L., Weber, G., Tallarico, J. A., King, R. W. and Zon, L. I. (2005). Small molecules that delay S phase suppress a zebrafish bmyb mutant. *Nat Chem Biol* **1**, 366-70.

Stern, H. M. and Zon, L. I. (2003). Cancer genetics and drug discovery in the zebrafish. *Nat Rev Cancer* **3**, 533-9.

Storkebaum, E., Lambrechts, D. and Carmeliet, P. (2004). VEGF: once regarded as a specific angiogenic factor, now implicated in neuroprotection. *Bioessays* **26**, 943-54.

Sui, D. and Wilson, J. E. (2000). Interaction of insulin-like growth factor binding protein-4, Miz-1, leptin, lipocalin-type prostaglandin D synthase, and granulin precursor with the N-terminal half of type III hexokinase. *Arch Biochem Biophys* **382**, 262-74.

Summerton, J. (1999). Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochim Biophys Acta* **1489**, 141-58.

Summerton, J. and Weller, D. (1997). Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev* **7**, 187-95.

Sun, X., Gulyas, M. and Hjerpe, A. (2004). Mesothelial differentiation as reflected by differential gene expression. *Am J Respir Cell Mol Biol* **30**, 510-8.

Sun, X., Li, X., Moriarty, P. M., Henics, T., LaDuca, J. P. and Maquat, L. E. (2001). Nonsense-mediated decay of mRNA for the selenoprotein phospholipid hydroperoxide glutathione peroxidase is detectable in cultured cells but masked or inhibited in rat tissues. *Mol Biol Cell* **12**, 1009-17.

Suzuki, M., Bannai, M., Matsumuro, M., Furuhata, Y., Ikemura, R., Kuranaga, E., Kaneda, Y., Nishihara, M. and Takahashi, M. (2000). Suppression of copulatory behavior by intracerebroventricular infusion of antisense oligodeoxynucleotide of granulin in neonatal male rats. *Physiol Behav* **68**, 707-13.

Suzuki, M., Yoshida, S., Nishihara, M. and Takahashi, M. (1998). Identification of a sex steroid-inducible gene in the neonatal rat hypothalamus. *Neurosci Lett* **242**, 127-30.

Tabata, T. and Takei, Y. (2004). Morphogens, their identification and regulation. *Development* **131**, 703-12.

Takehara, A., Eguchi, H., Ohigashi, H., Ishikawa, O., Kasugai, T., Hosokawa, M., Katagiri, T., Nakamura, Y. and Nakagawa, H. (2006). Novel tumor marker REG4 detected in serum of patients with resectable pancreatic cancer and feasibility for antibody therapy targeting REG4. *Cancer Sci* **97**, 1191-7.

Talapatra, A., Rouse, R. and Hardiman, G. (2002). Protein microarrays: challenges and promises. *Pharmacogenomics* **3**, 527-36.

Tangkeangsirisin, W., Hayashi, J. and Serrero, G. (2004). PC cell-derived growth factor mediates tamoxifen resistance and promotes tumor growth of human breast cancer cells. *Cancer Res* **64**, 1737-43.

Tangkeangsirisin, W. and Serrero, G. (2004). PC cell-derived growth factor (PCDGF/GP88, progranulin) stimulates migration, invasiveness and VEGF expression in breast cancer cells. *Carcinogenesis* **25**, 1587-92.

Tavassoli, M. (1991). Embryonic and fetal hemopoiesis: an overview. *Blood Cells* **17**, 269-81; discussion 282-6.

Taylor, J. S., Braasch, I., Frickey, T., Meyer, A. and Van de Peer, Y. (2003). Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res* **13**, 382-90.

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X. Q. et al. (2001). Expression of the zebrafish genome during embryogenesis. In *ZFIN Direct Data Submission*, (ed.

Thisse, C. and Zon, L. I. (2002). Organogenesis--heart and blood formation from the zebrafish point of view. *Science* **295**, 457-62.

Thompson, M. A., Ransom, D. G., Pratt, S. J., MacLennan, H., Kieran, M. W., Detrich, H. W., 3rd, Vail, B., Huber, T. L., Paw, B., Brownlie, A. J. et al. (1998). The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Dev Biol* **197**, 248-69.

Thornburg, N. J., Kusano, S. and Raab-Traub, N. (2004). Identification of Epstein-Barr virus RK-BARF0-interacting proteins and characterization of expression pattern. *J Virol* **78**, 12848-56.

Thummel, R., Bai, S., Sarras, M. P., Jr., Song, P., McDermott, J., Brewer, J., Perry, M., Zhang, X., Hyde, D. R. and Godwin, A. R. (2006). Inhibition of zebrafish fin regeneration using in vivo electroporation of morpholinos against fgfr1 and msxb. *Dev Dyn* **235**, 336-46.

Thummel, R., Burket, C. T., Brewer, J. L., Sarras, M. P., Jr., Li, L., Perry, M., McDermott, J. P., Sauer, B., Hyde, D. R. and Godwin, A. R. (2005). Cre-mediated site-specific recombination in zebrafish embryos. *Dev Dyn* **233**, 1366-77.

Tirosh, I. and Barkai, N. (2007). Comparative analysis indicates regulatory neofunctionalization of yeast duplicates. *Genome Biol* **8**, R50.

Tom Strachan, A. R. (2003). Human Molecular Genetics - Third Edition: Garland Science.

Tomasiewicz, H. G., Flaherty, D. B., Soria, J. P. and Wood, J. G. (2002). Transgenic zebrafish model of neurodegeneration. *J Neurosci Res* **70**, 734-45.

Trainor, P. A. (2005). Specification of neural crest cell formation and migration in mouse embryos. *Semin Cell Dev Biol* **16**, 683-93.

Trainor, P. A. and Krumlauf, R. (2001). Hox genes, neural crest cells and branchial arch patterning. *Curr Opin Cell Biol* **13**, 698-705.

Trinh, D. P., Brown, K. M. and Jeang, K. T. (1999). Epithelin/granulin growth factors: extracellular cofactors for HIV-1 and HIV-2 Tat proteins. *Biochem Biophys Res Commun* **256**, 299-306.

Tripp, K. W. and Barrick, D. (2004). The tolerance of a modular protein to duplication and deletion of internal repeats. *J Mol Biol* **344**, 169-78.

Turner, G. F., Seehausen, O., Knight, M. E., Allender, C. J. and Robinson, R. L. (2001). How many species of cichlid fishes are there in African lakes? *Mol Ecol* **10**, 793-806.

van der Zee, J., Le Ber, I., Maurer-Stroh, S., Engelborghs, S., Gijselinck, I., Camuzat, A., Brouwers, N., Vandenberghe, R., Sleegers, K., Hannequin, D. et al. (2007). Mutations other than null mutations producing a pathogenic loss of progranulin in frontotemporal dementia. *Hum Mutat* **28**, 416.

Vasavada, H. A., Ganguly, S., Germino, F. J., Wang, Z. X. and Weissman, S. M. (1991). A contingent replication assay for the detection of protein-protein interactions in animal cells. *Proc Natl Acad Sci U S A* **88**, 10686-90.

Venkatesh, B., Kirkness, E. F., Loh, Y. H., Halpern, A. L., Lee, A. P., Johnson, J., Dandona, N., Viswanathan, L. D., Tay, A., Venter, J. C. et al. (2007). Survey Sequencing and Comparative Analysis of the Elephant Shark (Callorhinchus milii) Genome. *PLoS Biol* **5**, e101.

Venter, J. C. Adams, M. D. Myers, E. W. Li, P. W. Mural, R. J. Sutton, G. G. Smith, H. O. Yandell, M. Evans, C. A. Holt, R. A. et al. (2001). The sequence of the human genome. *Science* **291**, 1304-51.

Verhoeyen, E., Wiznerowicz, M., Olivier, D., Izac, B., Trono, D., Dubart-Kupperschmitt, A. and Cosset, F. L. (2005). Novel lentiviral vectors displaying "early-acting cytokines" selectively promote survival and transduction of NOD/SCID repopulating human hematopoietic stem cells. *Blood* **106**, 3386-95.

Verstrepen, K. J., Jansen, A., Lewitter, F. and Fink, G. R. (2005). Intragenic tandem repeats generate functional variability. *Nat Genet* **37**, 986-90.

Vienne, A., Rasmussen, J., Abi-Rached, L., Pontarotti, P. and Gilles, A. (2003). Systematic phylogenomic evidence of en bloc duplication of the ancestral 8p11.21-8p21.3-like region. *Mol Biol Evol* **20**, 1290-8.

Vigdal, T. J., Kaufman, C. D., Izsvak, Z., Voytas, D. F. and Ivics, Z. (2002). Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *J Mol Biol* **323**, 441-52.

Vinciguerra, P. and Stutz, F. (2004). mRNA export: an assembly line from genes to nuclear pores. *Curr Opin Cell Biol* **16**, 285-92.

Wagner, A. (2002). Asymmetric functional divergence of duplicate genes in yeast. *Mol Biol Evol* **19**, 1760-8.

Wagner, K. R., Hamed, S., Hadley, D. W., Gropman, A. L., Burstein, A. H., Escolar, D. M., Hoffman, E. P. and Fischbeck, K. H. (2001). Gentamicin treatment of Duchenne and Becker muscular dystrophy due to nonsense mutations. *Ann Neurol* **49**, 706-11.

Wallace, K. N. and Pack, M. (2003). Unique and conserved aspects of gut development in zebrafish. *Dev Biol* **255**, 12-29.

Wallace, K. N., Yusuff, S., Sonntag, J. M., Chin, A. J. and Pack, M. (2001). Zebrafish hhex regulates liver development and digestive organ chirality. *Genesis* **30**, 141-3.

Walshe, J. and Mason, I. (2003). Fgf signalling is required for formation of cartilage in the head. *Dev Biol* **264**, 522-36.

Wan, H., Korzh, S., Li, Z., Mudumana, S. P., Korzh, V., Jiang, Y. J., Lin, S. and Gong, Z. (2006). Analyses of pancreas development by generation of gfp transgenic zebrafish using an exocrine pancreas-specific elastaseA gene promoter. *Exp Cell Res* **312**, 1526-39.

Wang, W., Hayashi, J., Kim, W. E. and Serrero, G. (2003). PC cell-derived growth factor (granulin precursor) expression and action in human multiple myeloma. *Clin Cancer Res* **9**, 2221-8.

Wang, W., Hayashi, J. and Serrero, G. (2006). PC cell-derived growth factor confers resistance to dexamethasone and promotes tumorigenesis in human multiple myeloma. *Clin Cancer Res* **12**, 49-56.

Wang, X., Chu, L. T., He, J., Emelyanov, A., Korzh, V. and Gong, Z. (2001). A novel zebrafish bHLH gene, neurogenin3, is expressed in the hypothalamus. *Gene* **275**, 47-55.

Warga, R. M. and Nusslein-Volhard, C. (1999). Origin and development of the zebrafish endoderm. *Development* **126**, 827-38.

Waterston, R. H. Lindblad-Toh, K. Birney, E. Rogers, J. Abril, J. F. Agarwal, P. Agarwala, R. Ainscough, R. Alexandersson, M. An, P. et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-62.

Wendik, B., Maier, E. and Meyer, D. (2004). Zebrafish mnx genes in endocrine and exocrine pancreas formation. *Dev Biol* **268**, 372-83.

Wendling, O., Ghyselinck, N. B., Chambon, P. and Mark, M. (2001). Roles of retinoic acid receptors in early embryonic morphogenesis and hindbrain patterning. *Development* **128**, 2031-8.

Whitwell, J. L., Jack, C. R., Jr., Baker, M., Rademakers, R., Adamson, J., Boeve, B. F., Knopman, D. S., Parisi, J. F., Petersen, R. C., Dickson, D. W. et al. (2007). Voxel-

based morphometry in frontotemporal lobar degeneration with ubiquitin-positive inclusions with and without progranulin mutations. *Arch Neurol* **64**, 371-6.

Willett, C. E., Kawasaki, H., Amemiya, C. T., Lin, S. and Steiner, L. A. (2001). Ikaros expression as a marker for lymphoid progenitors during zebrafish development. *Dev Dyn* **222**, 694-8.

Wilschanski, M., Famini, C., Blau, H., Rivlin, J., Augarten, A., Avital, A., Kerem, B. and Kerem, E. (2000). A pilot study of the effect of gentamicin on nasal potential difference measurements in cystic fibrosis patients carrying stop mutations. *Am J Respir Crit Care Med* **161**, 860-5.

Wilschanski, M., Yahav, Y., Yaacov, Y., Blau, H., Bentur, L., Rivlin, J., Aviram, M., Bdolah-Abram, T., Bebok, Z., Shushi, L. et al. (2003). Gentamicin-induced correction of CFTR function in patients with cystic fibrosis and CFTR stop mutations. *N Engl J Med* **349**, 1433-41.

Wilson, M. E., Scheel, D. and German, M. S. (2003). Gene expression cascades in pancreatic development. *Mech Dev* **120**, 65-80.

Wisneski, L. A. (1990). Salmon calcitonin in the acute management of hypercalcemia. *Calcif Tissue Int* **46 Suppl**, S26-30.

Wolfe, K. H. and Li, W. H. (2003). Molecular evolution meets the genomics revolution. *Nat Genet* **33 Suppl**, 255-65.

Woltering, J. M. and Durston, A. J. (2006). The zebrafish hoxDb cluster has been reduced to a single microRNA. *Nat Genet* **38**, 601-2.

Wood, A. W., Schlueter, P. J. and Duan, C. (2005). Targeted knockdown of insulin-like growth factor binding protein-2 disrupts cardiovascular development in zebrafish embryos. *Mol Endocrinol* **19**, 1024-34.

Wood, T. L., Rogler, L. E., Czick, M. E., Schuller, A. G. and Pintar, J. E. (2000). Selective alterations in organ sizes in mice with a targeted disruption of the insulin-like growth factor binding protein-2 gene. *Mol Endocrinol* **14**, 1472-82.

Woods, I. G., Kelly, P. D., Chu, F., Ngo-Hazelett, P., Yan, Y. L., Huang, H., Postlethwait, J. H. and Talbot, W. S. (2000). A comparative map of the zebrafish genome. *Genome Res* **10**, 1903-14.

Wu, Y., Zhang, G., Xiong, Q., Luo, F., Cui, C., Hu, W., Yu, Y., Su, J., Xu, A. and Zhu, Z. (2006). Integration of double-fluorescence expression vectors into zebrafish genome for the selection of site-directed knockout/knockin. *Mar Biotechnol (NY)* **8**, 304-11.

Xu, K., Zhang, Y., Ilalov, K., Carlson, C. S., Feng, J. Q., Di Cesare, P. E. and Liu, C. J. (2007). Cartilage oligomeric matrix protein associates with granulin-epithelin precursor (GEP) and potentiates GEP-stimulated chondrocyte proliferation. *J Biol Chem* **282**, 11347-55.

Xu, S. Q., Tang, D., Chamberlain, S., Pronk, G., Masiarz, F. R., Kaur, S., Prisco, M., Zanocco-Marani, T. and Baserga, R. (1998). The granulin/epithelin precursor abrogates the requirement for the insulin-like growth factor 1 receptor for growth in vitro. *J Biol Chem* **273**, 20078-83.

Yan, Y. L., Willoughby, J., Liu, D., Crump, J. G., Wilson, C., Miller, C. T., Singer, A., Kimmel, C., Westerfield, M. and Postlethwait, J. H. (2005). A pair of Sox: distinct and overlapping functions of zebrafish sox9 co-orthologs in craniofacial and pectoral fin development. *Development* **132**, 1069-83.

Yang, J., Yu, M., Liu, B., Zhu, M. and Li, K. (2006). Porcine granulin gene (GRN): molecular cloning, polymorphism and chromosomal localization. *DNA Seq* **17**, 350-4.

Yee, N. S., Yusuff, S. and Pack, M. (2001). Zebrafish pdx1 morphant displays defects in pancreas development and digestive organ chirality, and potentially identifies a multipotent pancreas progenitor cell. *Genesis* **30**, 137-40.

Yelick, P. C. and Schilling, T. F. (2002). Molecular dissection of craniofacial development using zebrafish. *Crit Rev Oral Biol Med* **13**, 308-22.

Yoo, B. K., Triller, D. M. and Yoo, D. J. (2006). Exenatide: a new option for the treatment of type 2 diabetes. *Ann Pharmacother* **40**, 1777-84.

Yu, X., Xu, D. and Cheng, Q. (2006). Label-free detection methods for protein microarrays. *Proteomics* **6**, 5493-503.

Zamore, P. D., Tuschl, T., Sharp, P. A. and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25-33.

Zanocco-Marani, T., Bateman, A., Romano, G., Valentinis, B., He, Z. H. and Baserga, R. (1999). Biological activities and signaling pathways of the granulin/epithelin precursor. *Cancer Res* **59**, 5331-40.

Zaret, K. S. (2001). Hepatocyte differentiation: from the endoderm and beyond. *Curr Opin Genet Dev* **11**, 568-74.

Zecchin, E., Mavropoulos, A., Devos, N., Filippi, A., Tiso, N., Meyer, D., Peers, B., Bortolussi, M. and Argenton, F. (2004). Evolutionary conserved role of ptf1a in the specification of exocrine pancreatic fates. *Dev Biol* **268**, 174-84.

Zelensky, A. N. and Gready, J. E. (2004). C-type lectin-like domains in Fugu rubripes. *BMC Genomics* **5**, 51.

Zelensky, A. N. and Gready, J. E. (2005). The C-type lectin-like domain superfamily. *Febs J* **272**, 6179-217.

Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57-68.

Zhang, Y. W., Ding, L. S. and Lai, M. D. (2003). Reg gene family and human diseases. *World J Gastroenterol* **9**, 2635-41.

Zhou, J., Gao, G., Crabb, J. W. and Serrero, G. (1993). Purification of an autocrine growth factor homologous with mouse epithelin precursor from a highly tumorigenic cell line. *J Biol Chem* **268**, 10863-9.

Zhu, J., Nathan, C., Jin, W., Sim, D., Ashcroft, G. S., Wahl, S. M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Wright, C. D. et al. (2002). Conversion of proepithelin to epithelins: roles of SLPI and elastase in host defense and wound repair. *Cell* **111**, 867-78.

Zweifel, M. E. and Barrick, D. (2001). Studies of the ankyrin repeats of the Drosophila melanogaster Notch receptor. 1. Solution conformational and hydrodynamic properties. *Biochemistry* **40**, 14344-56.

Zweifel, M. E., Leahy, D. J., Hughson, F. M. and Barrick, D. (2003). Structure and stability of the ankyrin domain of the Drosophila Notch receptor. *Protein Sci* **12**, 2622-32.