Identifying Novel Candidate Regulators of Kidney Morphogenesis using Developmental Metabolomics as a Tool

A Thesis Submitted
in Partial Fulfillment of the Requirements
for the
Five Year Dual Degree BS-MS Programme, IISER Pune

By

Apeksha Tare
Indian Institute of Science Education & Research (IISER) Pune

Under Supervision of
Dr Amitabha Bandyopadhyay
Department of Biological Sciences and Bioengineering
Indian Institute of Technology Kanpur
KANPUR – 208016

April 2012

Certificate

Certified that the work contained in the thesis entitled: *Identifying Novel Candidate Regulators of Kidney Morphogenesis using Developmental Metabolomics as a Tool* by *Ms Apeksha Tare*, 5th year, Five year dual degree BS-MS programme, Indian Institute of Science Education & Research (IISER), Pune, has been carried out under my supervision during August 1, 2011 – March 31, 2012.

Dr Amitabha Bandyopadhyay,
Department of Biological Sciences and Bioengineering
Indian Institute of Technology Kanpur
KANPUR – 208016

April 2012

Declaration

I hereby declare that the matter embodied in the thesis entitled: *Identifying Novel Candidate Regulators of Kidney Morphogenesis using Developmental Metabolomics as a Tool* are the results of the investigations carried out by me at the Department of Biological Sciences and Bioengineering, Indian Institute of Technology Kanpur, under the supervision of Dr Amitabha Bandyopadhyay and the same has not been submitted elsewhere for any other degree.

Apeksha Tare
April 2012

Acknowledgements

I express my profound thanks and gratitude to Dr Amitabha Bandyopadhyay for giving me an opportunity to work with him, introducing me to newer aspects in the field of developmental biology and for guiding me all along. A special thanks to Dr Jonaki Sen for allowing me to use the facilities of her lab and for providing useful insights.

I thank Dr L S Shashidhara and Dr Aurnab Ghose for encouraging me to carry out my research work at IIT Kanpur.

I extend my sincere thanks to Dr Mayurika Lahiri, in whose lab I have been trained for the past two years. She has been a constant source of encouragement and guidance.

I am extremely grateful to Priti, Ayan and Paritosh for their invaluable advice, help and support. I also extend a heartfelt thanks to Pratik, Brijesh, Prem, Shweta, Sandeep, Aditi, Monika, Shivani, Nashrah, Saborni, Suvimal, Divya, Ritika, Yachna, Barkha, Samarjeet, Shourya, and Asmita for all their help and cooperation and for creating such a friendly work environment. It has been a great learning experience for me and I thank all these people for making it an enjoyable one too. I also thank Naresh bhaiyya for taking care of the day-to-day autoclaving of necessary lab equipment and reagents.

Last but not the least, I am extremely indebted to my parents and two special friends, Ashutosh and Ayesha, for standing by me and lifting up my spirits at times when I felt so low and helpless. I also sincerely thank Rakesh bhaiyya for all his last minute help.

Apeksha Tare

Contents

List of Figures

List of Tables

Abstract

It was speculated that the dynamic metabolite profile of a tissue during organogenesis must be reflected by spatio-temporally well-regulated expression of metabolic enzymes. This was well elucidated by the results of a genome-wide expression screen of mRNAs encoding 1716 metabolic enzymes through high-throughput WM-ISH, across four different developmental stages of chicken (*Gallus gallus*), out of which 416 exhibited tissue-specific expression patterns. The work presented in this thesis includes a part of the expression screening, subsequent analysis of kidney-specific expression data and testing of an emerging hypothesis.

From the kidney-specific expression data set, a co-expression cluster of 106 metabolic genes was identified in HH26 kidney tubules which did not map on to any common biochemical pathway. However, several putative common upstream regulators for these genes were identified through microarray data mining analysis, one of them being HNF4α, a nuclear hormone receptor known to be a master regulator of liver-specific genes, and implicated to have a role in kidney development. Its expression in HH26 kidney tubules was verified through WM-ISH. To validate its regulatory role in the developing kidney, it was planned to monitor the expression changes of its putative downstream genes following HNF4α loss-of-function studies *in vivo*, for which purpose miRNA expression constructs and a retroviral dominant negative construct have been developed. It is strongly believed that the upstream regulators of such a large cohort of kidney-specific metabolic genes identified via this approach are likely to be novel determinants of renal morphogenesis, which is the primary motivation behind this study.

1 Introduction

1.1 Need for a systems biology approach

Much of the work in molecular developmental biology till date has focused on identifying and characterizing signaling molecules, transcription factors, growth factors and their receptors involved in regulating key developmental processes. These are regarded to be the most superior players in the hierarchy of a developmental signaling cascade. The most downstream effector molecules in these signaling cascades are largely unexplored. Although this kind of a reductionist approach has proved to be extremely useful in dissecting the role of individual components of a complex biological system, it is not sufficient for a thorough mechanistic understanding of intricate processes like differentiation, organogenesis and morphogenesis. In order to be able to describe the complex properties of such multifactorial biological systems which involve the interaction of complex regulatory networks at multiple levels, there is need to adopt a more holistic approach which emphasizes the whole system and not only the components in isolation.

1.2 Significance of the metabolome

There is no denying the fact that in multicellular organisms, spatio-temporal regulation of gene expression is extremely critical during development (Gehring & Ikeo, 1999; Takahashi *et al.*, 2007) and that the first and foremost level at which differential gene expression is regulated is at the transcriptional level. And therefore, elucidating transcription factor networks is crucial for elucidating developmental pathways. However, the ultimate biochemical manifestation of any genetic or environmental perturbation, the most reliable and sensitive indicator of the biological phenotype is the metabolome (Dunn *et al.*, 2011). Metabolites are the most downstream in the biochemical flow of information (Fig.1.1) and are the building blocks of all other biochemical entities including proteins (amino acids), DNA and RNA (nucleotides), and extracellular matrices (lipids, carbohydrates and lipid- or carbohydrate-modified proteins). It is because of the wide range of properties of the metabolites that they are able to confer distinct organs and tissues (or cell types) their unique physico-chemical properties which are commensurate with their

function. The metabolome is a rapid indicator of any kind of perturbation in the system because it is extremely dynamic (Dunn *et al.*, 2011). This implies that during organogenesis, the metabolite profile of any tissue must be constantly changing and would be a fairly accurate indicator of the developmental stage of tissue specification. However, the highly diverse range of physico-chemical properties and concentrations of metabolites make it extremely challenging to carry out comprehensive investigations at the level of the metabolome by making it necessary to employ multiple analytical platforms (Dunn *et al.*, 2011). Furthermore, our limited knowledge of the range of metabolites produced in a given tissue makes it even more difficult to study metabolomes in a dynamic system like a developing animal. To the contrary, single analytical platforms are sufficient for detection at the protein or transcript levels.

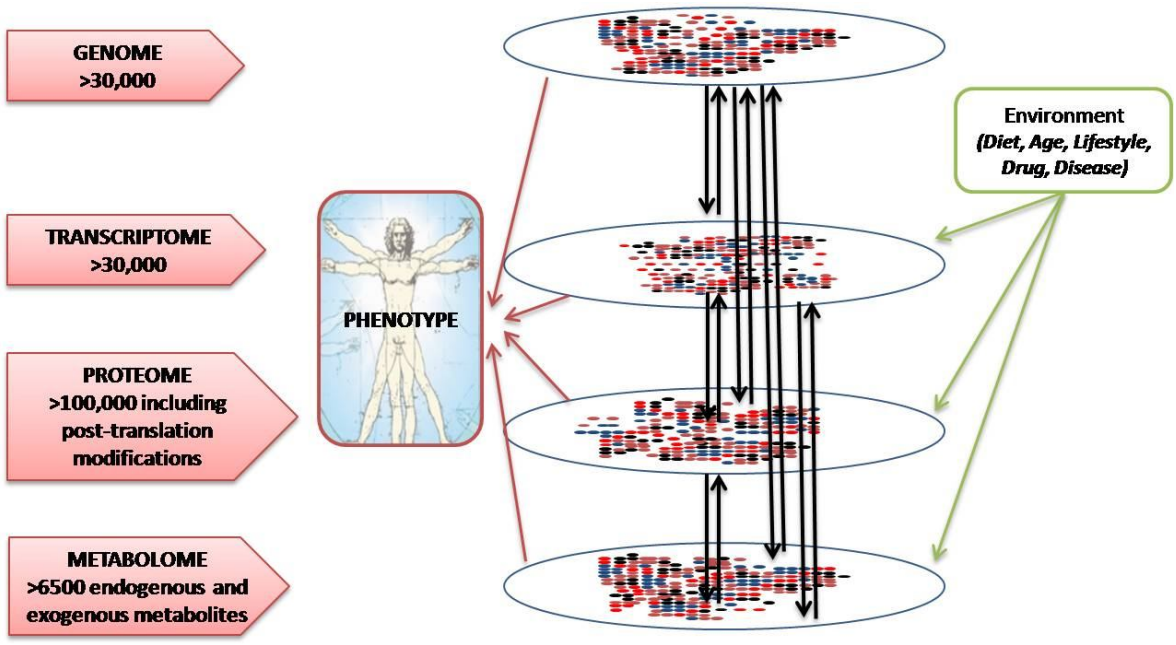


Figure 1.1: The complex interactions of genome, transcriptome, proteome and metabolome in biological systems (Adapted from Dunn *et al.*, 2011).

1.3 Metabolic enzymes in the context of development

The synthesis of metabolites is not template-driven, they are either substrates, intermediates or end-products of enzyme-catalyzed biochemical reactions. Therefore, it is speculated that the expression of metabolic enzymes must be tightly spatio-temporally regulated during development and that these might constitute a significant part of the most downstream target genes in developmental signaling

cascades, whose expression is most likely regulated by upstream signaling molecules and transcription factors. There exist several lines of evidence which support these speculations, some of which are summarized below.

The formation and maintenance of extracellular matrix (ECM), which not only provides structural support and orientation to tissues (Bosman & Stamenkovic, 2003) but also influences processes such as cell migration, proliferation, adhesion and differentiation (Gelse *et al.*, 2003; Kalluri, 2003) which are very crucial during organogenesis, also requires metabolic enzymes like those involved in collagen synthesis (Huxley-Jones *et al.*, 2007).

Another very important class of metabolic enzymes, the cytochrome P450 (CYP) gene family which is involved in metabolic oxidation of a variety of substrates, has been shown to be highly regulated in terms of the particular forms which are expressed at particular developmental time-points (Miller *et al.*, 1996). Dynamic expression profiles of enzymes involved in heme metabolism have also been reported during *Xenopus* development (Shi *et al.*, 2008). Manganese superoxide dismutase, an essential antioxidant enzyme involved in organogenesis, has been shown to have a dynamically changing expression profile during mouse organogenesis (Minyon *et al.*, 2011). Another recent study has unveiled a distinct pattern of expression of adenylate kinase and AMPK genes which facilitates the differentiation of embryonic stem cells leading to a cardiogenic fate (Dzeja *et al.*, 2011). In a study aimed at identifying downstream targets of RXRα, a nuclear hormone receptor, using RXRα null mutant mouse embryos, it was discovered that more than 50% of the putative downstream target genes identified were metabolic genes (Lozano *et al.*, 1998).

1.4 A general hypothesis leading to specific hypotheses

Even though many such independent studies indicating the significance of metabolic enzymes and their regulation in the context of development have been carried out in isolation, detailed metabolic circuits which orchestrate cellular differentiation and organogenesis are largely undefined. As a small yet significant step towards this goal and keeping in mind the importance of a holistic approach, we started with testing a rather general hypothesis that “metabolic enzymes must be expressed in a

spatio-temporally regulated manner during development". The idea was to start with an initial inductive or hypothesis-generating study which would help us acquire data on a wide range of metabolic enzymes, the subsequent analysis of which would guide us appropriately in framing specific hypotheses and testing them through our traditional reductionist approaches. For this purpose, a genome-wide expression screen of mRNAs encoding metabolic enzymes through high-throughput whole-mount RNA *in situ* hybridization (WM-ISH), across four different developmental stages of chicken (*Gallus gallus*), was undertaken in the laboratory (Roy *et al.*, 2012). The work presented in this thesis includes a part of the expression screening, subsequent analysis of some part of the data generated and testing of one of the emerging specific hypotheses.

A total of 1716 metabolic enzymes have been screened, out of which, 416 have been found to exhibit tissue-specific expression domains at one or more embryonic stages, thus supporting our hypothesis. The results of this screen suggest a distinctive role of these metabolic enzymes in organogenesis and provide several new leads for further functional analysis. The kidney ranks as the third highest organ, after limb and somites, in terms of the number of metabolic enzymes being expressed in it, within the developmental time window that was scanned. Within kidney, we identified a co-expression group of 106 genes which are expressed specifically in the kidney tubules. The vertebrate kidney has served as a very useful model organ-system for investigating fundamental processes of organogenesis because its development involves most of these, including inductive tissue interactions, cell polarization, mesenchymal-to-epithelial transformation, and branching morphogenesis. In the light of these observations and for the purpose of this study, the kidney-specific gene expression data was used as a starting point for further analysis and experimental design.

Broadly speaking, the objective of this study was to combine bioinformatic analysis with molecular genetic tools to establish a link between the metabolic genes and their upstream regulators, in the context of the developing chicken kidney. We expect that some of the upstream regulators identified in this way would be novel regulators of kidney morphogenesis. We used a novel bioinformatic approach as well as conventional candidate gene approaches to discover potential regulators of

kidney tubule specific co-expression group. Chicken was used as a model system for this study.

1.5 Kidney development

The urogenital system, which includes the kidney, derives from intermediate mesoderm (IM), a region of mesoderm lying between paraxial (or somitic) mesoderm and lateral plate mesoderm, during development. Three different kidney structures are known to form sequentially from anterior to posterior, during kidney development, namely, the pronephros, the mesonephros and the metanephros, in that order. The pro- and mesonephros form first and are transient developmental intermediates in birds and mammals. The metanephros is the permanent kidney and is formed from the posterior region of the IM.

Avian nephrogenesis begins at around HH stage 9-10 with the formation of the nephric duct via mesenchymal-to-epithelial transition (Obara-Ishihara et al., 1999) within the IM. This is later called the Wolffian duct and is essential for all further kidney organogenesis. Many of the early molecular players involved in nephric duct morphogenesis are fairly well-characterized, Pax2, Lim1, Sim1, WT1, Bmp etc. being some of them (Dressler, 2006). The duct extends caudally during development and fuses with the cloaca at around HH stage 22. Metanephric kidney development begins with the aggregation of cells in the posterior region of the IM to form a cell population known as the metanephric mesenchyme (MM), which contains the precursors for diverse elements of the adult kidney and has a unique transcription factor profile. Eya1 is known to be a critical regulator of MM specification (Sajithlal et al., 2005). The kidney is a fairly complex organ, consisting of about 25 distinct cell types. However, kidney development can be simplistically summarized as the cooperative development of two distinct types of tubules. The first i.e. the ureteric bud appears as an outgrowth from the nephric duct into the MM and undergoes branching morphogenesis to form the entire collecting system of the kidney, both these processes being controlled by Ret/Gdnf pathway (Sainio et al., 1997). All subsequent development of the metanephric kidney is governed by signaling and transcription networks activated as a result of reciprocal tissue interactions between the invading ureteric bud and the MM, most of which are unknown. The second type of tubules, the nephrons, are also formed via mesenchymal-to-epithelial transition

induced by the invasion of the ureteric bud into the MM. These tubules undergo further patterning and segmentation to adopt their fully functional form. Little is known about the signalling networks regulating nephron patterning and segmentation.

Owing to the advent of genomics and proteomics combined with genetic manipulation techniques, kidney is emerging as a very well characterized model system of organogenesis. The cellular and molecular bases for many early events during kidney morphogenesis have become increasingly clear. However, much is yet to be discovered.

1.6 Using chicken as a model system

Till date, mouse has been the most commonly used vertebrate model organism for investigating signaling pathways regulating early kidney development because of the availability of techniques such as *in vitro* organ culture, gene knockout and expression analyses. However, the chicken (*Gallus gallus*) embryo also has served as an important model system in developmental biology owing to its accessibility throughout embryogenesis and amenability to experimental manipulation. Its usefulness as a model system has all the more increased due to the sequencing of the chicken genome and the availability of a variety of web resources. Gene expression analysis techniques such as RNA *in situ* hybridisation have also been well established in the chicken. Although the development of transgenic technology in chicken has been hampered by technical challenges, successful employment of molecular approaches such as *in ovo* electroporation and retrovirus-mediated gene transfer, including RNAi (Smith *et al.*, 2009), have made the chicken embryo an extremely useful system for functional testing of genes, allowing gain- and loss-of-function studies *in vivo* in a spatially and temporally controlled manner. The retroviral gene transfer is relatively more stable, widespread and long-lasting such that the resultant phenotypes can be assessed at any desired time after infection, throughout development.

2 Materials & Methods

2.1 Synthesis of Digoxigenin (DIG) - labeled probes

The EST for the gene of interest was PCR-amplified from the corresponding plasmid isolated from the ChEST clone, using T3-T7 primers. ChEST clones were obtained from Geneservice, UK and all ESTs were cloned in pBluescript II KS+ (Stratagene). Antisense RNA probe was transcribed from the linear PCR template with T3 RNA polymerase, in the presence of Digoxigenin-labeled nucleotide mix. Subsequently, RNA was precipitated with LiCl/EtOH and resuspended in TE [pH 8].

2.2 Whole-mount *in situ* Hybridization (WM-ISH)

WM-ISH was performed with few modifications of the protocol published by Riddle *et al.* in 1993. Chicken embryos were staged according to Hamburger-Hamilton (HH) staging (Hamburger & Hamilton, 1951).

HH18, HH22, HH26 and HH28 chicken embryos were harvested, fixed in 4% paraformaldehyde (PFA) in PBS overnight, washed once with PBT (PBS with 0.1% Tween 20), dehydrated through an increasing gradient of methanol/PBT (25%, 50%, 75%, 2 x 100% methanol), and stored at -20°C till further use. Unless mentioned otherwise, all washes were for 5 min each.

Embryos were rehydrated through a decreasing gradient of methanol/PBT followed by two PBT washes. Embryos were bleached with 6% hydrogen peroxide in PBT, washed thrice with PBT, permeabilized with HH stage-specific proteinase K treatment, washed twice with PBT, post-fixed with 4% PFA and 0.2% glutaraldehyde in PBT, washed twice with PBT and incubated with hybridization solution (50% formamide, 5X SSC [pH 4.5], 1% SDS, 50 µg/ml yeast tRNA, 50 µg/ml Heparin; 20X SSC: 3M NaCl, 0.3 M sodium citrate [pH 4.5]) for 1 hour at 70°C. DIG-labeled RNA probe was added in the hybridization solution and incubated overnight at 70°C.

After hybridization, embryos were washed thrice for 30 min each at 70°C with solution I (50% formamide, 5X SSC [pH 4.5], 1% SDS), thrice for 30 min each at 70°C with solution III (50% formamide, 5X SSC [pH 4.5]), and thrice at room

temperature with tris-buffered saline (TBS) containing 0.1% Tween 20 (TBST). To prevent nonspecific binding of antibody, embryos were blocked with TBST containing 10% heat-inactivated sheep serum for 1 hour at room temperature. Anti-DIG-Fab Fragment conjugated with Alkaline Phosphatase (Roche) (1:2000 in TBST containing 1% serum) was added and incubated overnight at 4°C.

Following antibody incubation, embryos were washed thrice with TBST at room temperature, five times for 1 hour each with TBST at room temperature, and overnight with TBST at 4°C.

Embryos were washed thrice for 10 min each with NTMT (100 mM NaCl, 100 mM Tris-HCL [pH 9.5], 50 mM $MgCl_2$, 0.1% Tween 20), followed by incubation with detection solution (125 mg/ml BCIP and 250 mg/ml NBT in NTMT) at room temperature, away from light. When the detection reaction was judged complete, embryos were washed with PBS, post-fixed with 4% PFA, washed thrice with PBS and stored in 1% PFA. Embryos were viewed with Leica DM S6D and photographed with Leica DFC290.

2.3 Section *in situ* Hybridization

HH26 and HH28 embryos were harvested, fixed with 4% PFA in PBS at 4°C overnight, washed twice with PBS and dehydrated through an increasing gradient of ethanol in PBS (25%, 50%, 75%, 2 x 100%). Embryos were incubated with xylene for 1-2 min, then with a 1:1 xylene-paraffin wax mix at 60°C twice for 15 min each, then with molten paraffin wax at 60°C for 1 hour, and then with fresh paraffin wax again at 60°C for 5-6 hours. The embryos (heads removed) were then embedded in paraffin wax blocks and stored at 4°C till further use. 9 µm thick sections were generated using Leica RM2255 microtome and collected on lysine-coated slides.

Before hybridization, slides were baked on a hot plate at 60°C for 1 hour, de-waxed in xylene twice at room temperature and rehydrated through a decreasing gradient of ethanol in PBS, followed by two washes with PBS. The rehydrated tissues were fixed with 4% PFA/PBS for 10 min, washed with PBT twice, treated with Proteinase K (1 µg/ml) for 10 min, washed twice with PBT, incubated with 4% PFA/PBS for 10 min, washed twice with PBT, incubated with acetic anhydride in 0.1M tri-ethanolamine,

and washed twice with PBT. Slides were air dried and pre-warmed hybridization solution containing DIG-labeled probe was added. Slides were covered with polypropylene cover slips, and incubated at 65°C overnight in a humidified chamber with paper towels soaked in 5X SSC/50% formamide.

Post hybridization, cover slips were removed by rinsing in 5X SSC [pH 7], slides were washed with 1X SSC [pH 7]/50% formamide at 65°C for 30 min, then with TNE (10 mM Tris-Cl [pH 7.5], 500 mM NaCl, 1 mM EDTA) for 10 min at 37°C, then with TNE containing RNase A (20 µ/ml, Roche) for 30 min at 37°C, then with TNE again for 10 min at 37°C, then with 2X SSC [pH 7] for 30 min, twice with 0.2 X SSC [pH 7] at 65°C for 20 min each.

Slides were then washed thrice with MABT (100 mM Maleic acid, 150 mM NaCl, 0.1% Tween 20 [pH 7.5]) at room temperature, then blocked with MABT containing 20% heat-inactivated goat serum for 1 hour at room temperature, followed by addition of Anti-DIG-AP antibody (1:2500 in MABT with 2% serum) and incubation at 4°C overnight.

The slides were then washed four times with MABT at room temperature, then with NTMT thrice, followed by incubation with detection solution (125 mg/ml BCIP and 250 mg/ml NBT in NTMT) at room temperature, away from light, till the reaction was judged complete. Slides were then washed with NTMT, then with PBS twice followed by post-fixing with 4% PFA/PBS, washing with PBS and water, twice each and finally mounting in gelvatol for storage. Stained sections were viewed in Zeiss Axiovision and photographed with Zeiss Axiocam HRc.

2.4 Design and cloning of RNAi hairpins against HNF4α

Three miRNAs were designed to target three different regions of the chick HNF4α mRNA (Fig. 3.7), using the online Invitrogen BLOCK-iT RNAi Designer tool (http://rnaidesigner.invitrogen.com/rnaiexpress/). For each miRNA, two single-stranded oligonucleotides were designed, one encoding the target pre-miRNA (top strand oligo) and the other complementary to it (bottom strand oligo), which were then annealed to generate the double-stranded oligo to be used for cloning.

The double-stranded oligos (Fig. 2.1) encoding the pre-miRNAs were cloned between the two BsaI sites of the miRNA expression vector, pRmiR (Fig. 3.6) to generate pRmiR-hairpin1, pRmiR-hairpin2, and pRmiR-hairpin3. The cloning site is flanked on either side by miR-155 flanking sequences to allow proper *in vivo* processing of the pre-miRNA. After transfection, the pre-miRNA forms the intramolecular stem-loop structure which is then processed by the endogenous Dicer enzyme to form a 22 nucleotide long mature miRNA. Also, since the EmGFP (Emerald Green Fluorescent Protein) gene present in the vector is expressed co-cistronically with the miRNA, its expression can be taken as a fairly accurate read-out of miRNA expression.

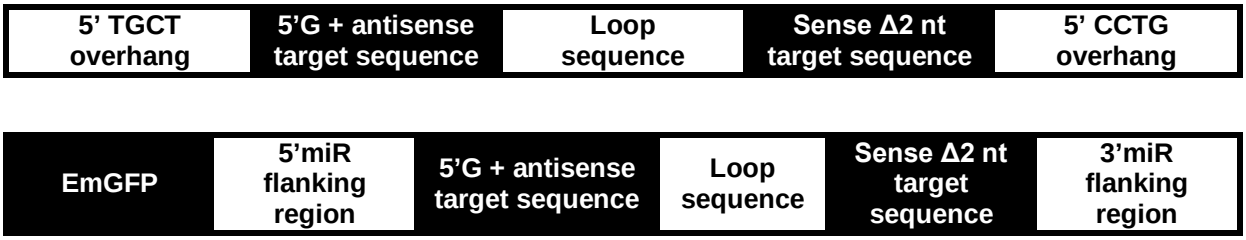


Figure 2.1: Structural features of the engineered pre-miRNA (a) before, and (b) after being cloned in the miRNA expression vector

The following top and bottom strand oligos were used for the three hairpins against HNF4α mRNA, the sequence features are shown in Fig. 2.2.

1. Hairpin 1 - targets 125 bp to 145 bp of ORF

Top strand:

5'- TGCTGTGTTGAGGTTGGCTGCTTCTGGTTTTGGCCACTGACTGACCAGAAGCACAACCTCAACA **-3'**

Bottom strand:

3'- CACAACTCCAACCGACGAAGACCAAAACCGGTGACTGACTGGTCTTCGTGTTGGAGTTGTGTCC **-5'**

2. Hairpin 2 – targets 944 bp to 964 bp of ORF

Top strand:

5'- TGCTGTGATGTAGTCCTCCAGGCTGAGTTTTGGCCACTGACTGACTCAGCCTGGGACTACATCA **-3'**

Bottom strand:

3'- CACTACATCAGGAGGTCCGACTCAAAACCGGTGACTGACTGAGTCGGACCCTGATGTAGTGTCC **-5'**

3. Hairpin 3 – targets 1739 bp-1759 bp of 3'UTR

Top strand:

5'- TGCTGATCACTTGTTCACAGGGTCTGGTTTTGGCCACTGACTGACCAGACCCTGAACAAGTGAT **-3'**

Bottom strand:

3'- CTAGTGAACAAGTGTCCCAGACCAAAACCGGTGACTGACTGGTCTGGGACTTGTTCACTAGTCC **-5'**

***Top strand**:*

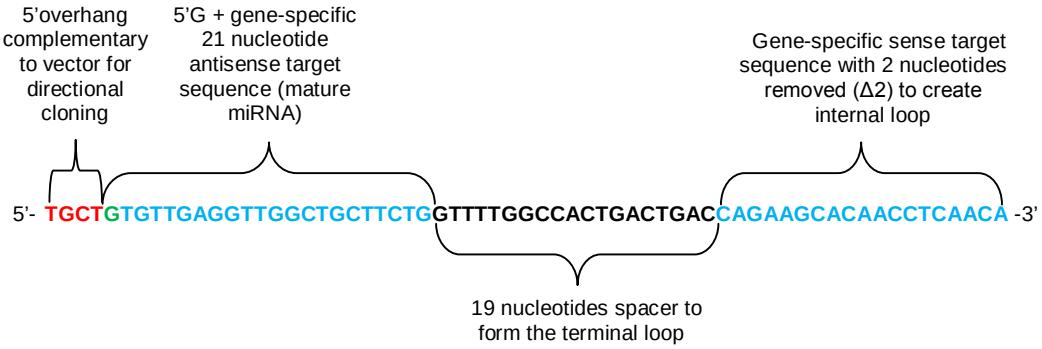


Figure 2.2: Sequence features of a top strand oligo encoding any miRNA, using hairpin 1 against HNF4α as an example.

2.5 Isolation and cloning of full length HNF4α and DNHNF4α cDNAs

2.5.1 Overall strategy

RNA was isolated from HH26 chicken embryos and cDNA was made through reverse transcription. The full length HNF4α CDS and dominant negative HNF4α (DNHNF4α) CDS were PCR-amplified from the cDNA with gene-specific primers, and subsequently cloned into TA vector (pCRII vector, Invitrogen) to generate pCRII-HNF4α and pCRII-DNHNF4α, respectively (Fig. 3.12). The pCRII-DNHNF4α was digested with BsaI and EcoRI, and the DNHNF4α insert was subcloned into the NcoI/EcoRI sites of pSlax21 (a shuttle vector whose MCS is flanked by two ClaI sites) to generate pSlax21-DNHNF4α (Fig. 3.15). The ClaI fragment of pSlax21-DNHNF4α was finally subcloned into the RCAS(A) vector to generate RCAS-DNHNF4α (Fig. 3.17). (All three vectors offer Ampicillin selection.)

2.5.2 Primer design rationale

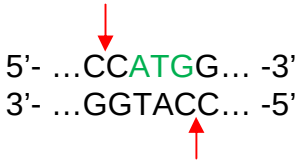
The following primers were designed to PCR-amplify full length HNF4α CDS (1365 bp).

5' Primer: 5'-ATGGAGATGGCCGATTATAGCGC-3'
3' Primer: 5'-CTAGATGACCTCCTGCTTTGTGATGG-3'

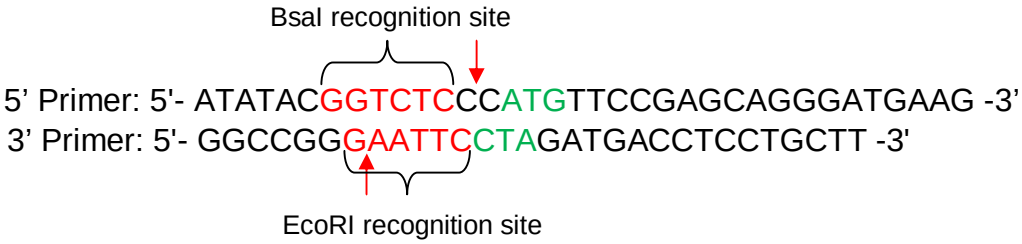
HNF4α mRNA: 1815 bp (NCBI Reference Sequence: NM_001030855.1)
HNF4α CDS: 25-1389 bp

For making DNHNF4α, the first 111 amino acids i.e. the 1st 333 nucleotides of the HNF4α CDS have to be omitted. Also, to ensure translation of the truncated protein, there has to be an in-frame START codon (ATG). The NcoI site of pSlax21 is

particularly useful for this purpose because it helps in introducing an in-frame ATG, as depicted below.



Keeping these considerations in mind, the following primers were designed such that the 5' primer has a BsaI site and an ATG in the overhang and the 3' primer has an EcoRI site in the overhang. This would ensure subsequent cloning into the NcoI/EcoRI site of pSlax21 and translation of the protein *in vivo*. The following primers were used to PCR-amplify DNHNF4α (PCR Product shown in Fig. 2.3). The expected size of this PCR product is 1061 bp.



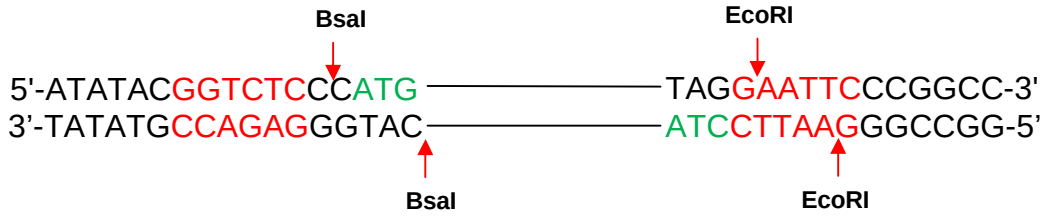


Figure 2.3: PCR-amplified DNHNF4α having BsaI and EcoRI sites (1061 bp) for subsequent subcloning into pSlax21.

After cloning of the PCR product into TA and subsequent digestion with BsaI/EcoRI, the digested insert would have overhangs compatible with those of pSlax21 digested with NcoI/EcoRI (Fig. 2.4).

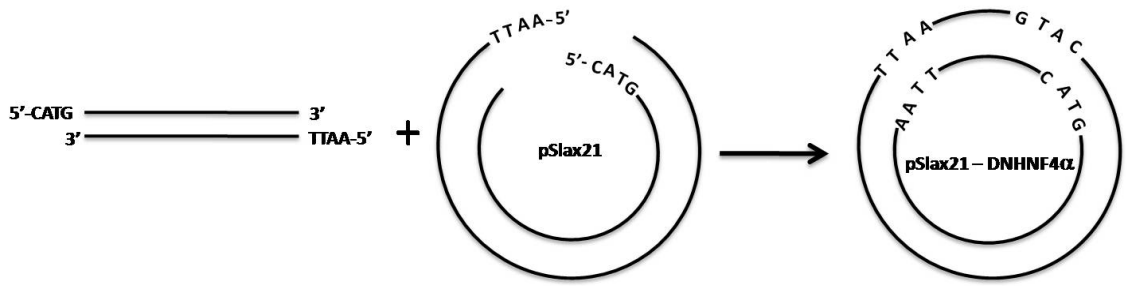


Figure 2.4: A schematic of the ligation of DNHNF4α digested from pCRII-DNHNF4α with BsaI/EcoRI with pSlax21 digested with NcoI/EcoRI.

Table 2.1: List of enzymes used for cloning

Enzyme	Company
BsaI	NEB
EcoRI	NEB
NcoI	NEB
ClaI	NEB
Antarctic Phophatase	NEB
Reverse Transcriptase	Promega, Roche, Invitrogen
Taq polymerase	Invitrogen
T4 DNA Ligase	NEB

2.6 Production of viral stocks of RCAS(A)AP, RCAS(B)AP and RCAS(A)GFP

RCAS(A)AP and RCAS(B)AP are RCAS vectors (http://home.ncifcrf.gov/hivdrp/RCAS) of the A and B subgroups that encode a reporter gene, human placental alkaline phosphatase (PLAP), whose expression can be easily assayed histochemically. RCAS(A)GFP encodes GFP as the reporter gene.

2.6.1 Cell culture and transfection

DF-1 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS, and penicillin-streptomycin (1X) in a humidified incubator at 37°C, 5% CO_2. Nearly 5 µg each of RCAS(A)AP, RCAS(B)AP and RCAS(A)GFP plasmid DNA was transfected into DF-1 cells in 6 cm dishes, via Calcium Phosphate method. Transfection, followed by viral propagation on the producer cell line DF-1, an avian fibroblast cell line, was used to produce high-titer viral stocks of RCAS, as described previously (Yang, 2002).

2.6.2 GFP fluorescence and AP (Alkaline Phosphatase) staining

The cells transfected with RCAS(A)GFP vector were visualized for GFP fluorescence under UV light. The cells transfected with RCAS(A)AP and RCAS(B)AP vectors were processed for AP staining, as follows. Cells were washed with PBS, fixed with 4% PFA/PBS at room temperature, incubated with PBS at 65°C for 45 min to inactivate endogenous AP, followed by washing with NTMT buffer and incubation with the detection solution (125 mg/ml BCIP and 250 mg/ml NBT in NTMT), at room temperature, away from light, till the reaction was judged complete.

2.6.3 Concentration of viral particles and determination of titer

For high infection efficiency, the viral stocks were concentrated and titers were determined through AP staining/GFP fluorescence, as described previously (Yang, 2002).

2.7 Semi-quantitative RT-PCR

The following primers were designed to amplify 500-700 bp products primarily towards the 3'UTR region of each transcript (5' primer chosen from 3' end of CDS and 3' primer from 3' UTR).

1. HNF4α (Product size: 520 bp)
   5' Primer: 5'- TCCTCAGCCATCTCCTCCT -3'
   3' Primer: 5'- ACTGCCTCCTGGTTTGGAG -3'
2. ACSM3 (Product size: 505 bp)
   5' Primer: 5'- CCTGCAGTGCTGGAATCA -3'
   3' Primer: 5'- CATCATCCATTGCACTCTTGA -3'
3. DIO1 (Product size: 619 bp)
   5' Primer: 5'- GAACCCCTTATGTCCAGTGGT -3'
   3' Primer: 5'- GCCTTCTGAAGTGAGGGAAA -3'
4. ACE (Product size: 590 bp)
   5' Primer: 5'- AACACAGAGAACGGGGAGGT -3'
   3' Primer: 5'- ATGGCTCTGCCCATCCTT -3'
5. MOGAT2 (Product size: 560 bp)
   5' Primer: 5'- TATGGTTTCGTGTCCCCTTC -3'
   3' Primer: 5'- GTGACTGTCTTCTGGGACGTT -3'
6. ACAA1 (Product size: 665 bp)
   5' Primer: 5'- GAGCATATGGCGTTGTGTCA -3'
   3' Primer: 5'- CAAAGCTGAAGCACATCCTG -3'
7. FBP1 (Product size: 578 bp)
   5' Primer: 5'- TCTGCTGGAGGTGTCAACTG -3'
   3' Primer: 5'- CAAAGGTCTGGCTGTTCAGA -3'

For these 7 genes, RT-PCR was performed to detect their expression in DF-1 cells and to get an idea of the range of PCR cycles required.

2.7.1 Transfection, RNA preparation and reverse transcription

DF-1 cells in 6 cm tissue culture dishes were transfected with 5 µg DNA each of pRmiR-hairpin1, pRmiR-hairpin2, pRmiR-hairpin3, pRmiR-lacZhairpin (negative control for RNAi), RCAS(A)-DNHNF4α, and RCAS(A)-GFP (negative control for RCAS construct) at ~ 60 – 70 % confluency, using Lipofectamine transfection reagent, according to manufacturer's protocol. 48 hours after transfection, cells transfected with the hairpin constructs and RCAS(A)-GFP were visualized for GFP fluorescence. RNA was isolated and equal amounts of RNA (1-5 µg) for the experimental and the respective control samples was reverse transcribed.

2.7.2 PCR reactions

For each cDNA sample, a series of PCR reactions was set up for all the transcripts whole levels were to be assessed. The tubes in each series varied in only one parameter and that was the number of PCR-amplification cycles (35-50).

The house-keeping gene GAPDH was used as an internal standard with a primer pair combination which yields a 500 bp PCR product (5'-ACCCAGAAGACTGTGGATGG -3'; 5'- CACATTGGGGGTAGGAACAC -3'). PCR cycle conditions were as follows: 95°C for 30 s, 55°C for 30 s, 72°C for 1 min. After PCR, equal volumes of reaction products were electrophoresed.

2.8 Tools used for creating vector maps, primer design and microarray data mining

- All vector maps were created using Vector NTI software from Invitrogen.
- Microarray data mining analysis was performed using GEO data mining program (Prasad & Bandyopadhyay, 2012).
- All primers were designed with the help of Primer3 program (http://biotools.umassmed.edu/bioapps/primer3_www.cgi).

3 Results & Discussion

3.1 Dynamic tissue-specific expression patterns of metabolic enzymes during development and subsequent analysis of kidney-specific expression data

As mentioned elsewhere in this thesis I joined an existing project in the laboratory of Dr Bandyopadhyay where two PhD students Ms Priti Roy and Mr Brijesh Kumar were carrying out a genome-wide expression screening of metabolic enzymes encoded in the chick genome. The overall objective of the project was to identify candidate metabolic enzymes which may have a role in the development of a given tissue.

In the context of this project I participated in conducting whole-mount RNA *in situ* hybridization for nearly 40 metabolic genes of HH stage 18, 22, 26 and 28 embryos. For the sake of continuity and coherence the expression patterns obtained for a few of them, are presented in Fig. 3.1.

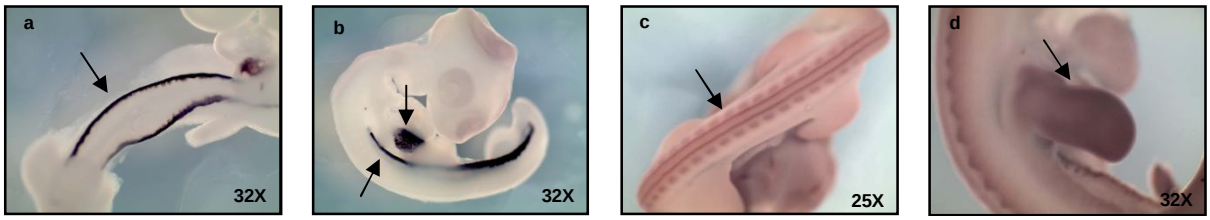


Figure 3.1: Whole-mount *in situ* hybridization. (a) ACSF2 expression in kidney at HH18; (b) ACSF2 expression in kidney and liver at HH22; (c) ATP6V0E2 expression in neural tube and somites at HH22; (d) ALDH9A1 expression in fore limb at HH26. Black arrows indicate the region of expression. Photography credits: Brijesh Kumar.

From here on, the kidney-specific gene expression data set was used as the raw material for further investigation, as mentioned previously. Although a majority of the kidney-specific genes were also found to be expressed in multiple tissues other than kidney during the developmental stages that were looked at, 9 such genes were identified whose expression was restricted exclusively to the kidney, namely, ACE, ACSM3, ALOX5AP, GGT1, GPT2, PTGES, PTPRS, QDPR, and RFX2. These are likely to be specific regulators/determinants of kidney development. For instance, ACE, which is one of these 9 genes and which we believe to be present in glomeruli, has been implicated to play a critical role in glomerular repair and development

(Joles *et al.*, 2006). Furthermore, the data set was used to identify a major co-expression cluster (cluster of genes which are spatially as well as temporally co-expressed) of 106 genes expressed in kidney tubules at HH stage 26 i.e. after 5 days of embryonic development. The expression domain of these genes was verified by performing section *in situ* hybridization for a few of them (Fig. 3.2), namely, DIO1, DPYS, and ISG20L1.

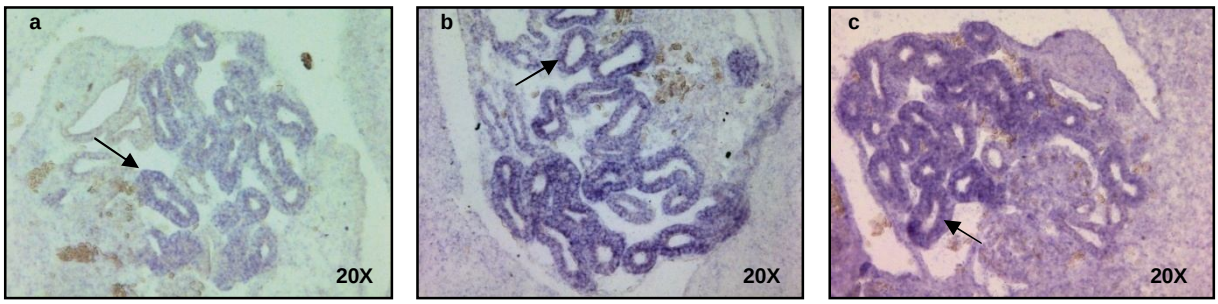


Figure 3.2: Section *in situ* hybridization of day 5 (HH26) chicken embryo. (a) DIO1 expression, (b) DPYS expression, and (c) ISG20L1 expression in kidney tubules (black arrows).

An earlier study has indicated the coordinated expression of enzymes within a metabolic pathway (Miki *et al.*, 2001). Keeping this in mind, all the metabolic enzymes identified within our co-expression group were mapped to their respective biochemical pathways, using the KEGG (Kyoto Encyclopedia of Genes and Genomes) online databse (http://www.genome.jp/kegg/). To the contrary, it was found that there was no common pathway to which most of the genes of the co-expression cluster belonged to (Fig. 3.3). The reason for this apparent discrepancy between the two studies is not immediately clear.

Nonetheless, it was speculated that since these genes share a common spatial as well as temporal expression pattern, it is highly probable that they would also share common upstream transcriptional regulators. In view of this, microarray data mining tool was employed to identify putative upstream regulators of the genes present within this co-expression group. For this purpose, a GEO data mining program has been developed earlier in the laboratory (Prasad & Bandyopadhyay, 2012), through which all the microarray experiments, in which expression of a certain transcription factor or signaling molecule is manipulated in the experimental treatment, have been annotated. The program accesses the experimental data from the GEO (Gene

Expression Omnibus) database of NCBI (http://www.ncbi.nlm.nih.gov/geo/), which is a public repository of microarray data generated by the scientific community.

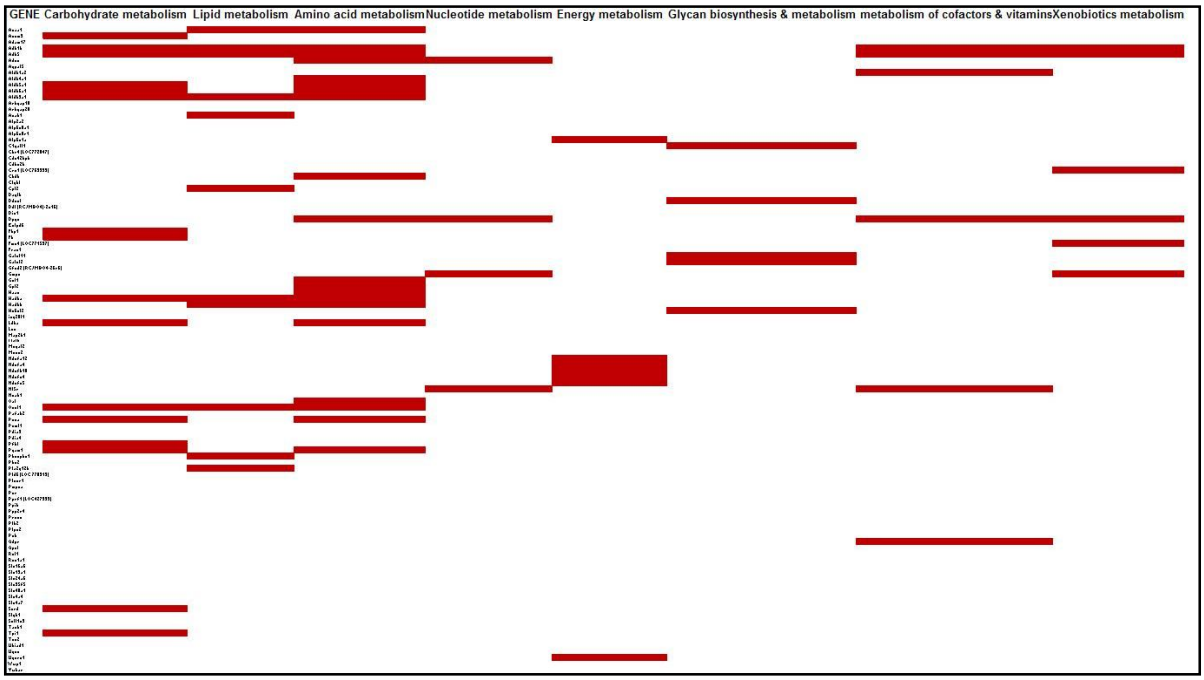


Figure 3.3: Distribution of metabolic genes of the HH26 kidney tubule co-expression group across metabolic pathways.

As a result of this data mining approach, several candidate upstream regulators were identified, namely, HNF4α, FoxO1, FoxO2, FoxO4, Dach1, Bcl3, E47, SOD1, Sox2, Sox7, IFN-α, IFN-γ, IL-4, IL-15, Lef1/β-catenin, TGF-β signaling, p53, Hox-A10, HoxC6, JAK kinase, IkB kinase β, etc. Amongst these, only HNF4α (Hepatocyte nuclear factor 4 alpha) was chosen for further validation of its regulatory role *in vivo*, for the following reasons. A recent study identified 37 anchor genes across 6 anatomical compartments in the developing mouse kidney, the early proximal tubule containing 25 of them. Upon bioinformatic analysis of the minimal promoter regions of these 25 genes, 22 of them were found to be enriched for transcription factor binding sites for HNF4α (Thiagarajan *et al.*, 2011) and some of these 22 genes are present in our kidney tubule co-expression cluster too. HNF4α has also been implicated to have a role in xenobiotics degradation and drug metabolism. Several cytochrome P450 (CYP450) enzymes and flavin-containing monooxygenases (FMOs) have been identified as HNF4α target genes (Verslues & Sladek, 2010), some of which are present in our co-expression group. Additionally, ALDH5A1 and

ALDH6A1 are two genes present in our co-expression group which have been identified as HNF4α target genes by microarray expression profiling in two independent studies in human embryonic kidney cells (HEK293) (Grigo *et al.*, 2008; Lucas *et al.*, 2005). DIO1, another member of our co-expression group, is reported to be regulated by HNF4α (Ohguchi *et al.*, 2008). Taken together, all these findings led us to hypothesize that HNF4α is a putative transcriptional regulator of the HH26 kidney tubule co-expression cluster (the specific hypothesis mentioned in Section 1.4). However, for this hypothesis to be true the minimal condition that has to be satisfied is that HNF4α must be expressed in chicken kidney tubules. The expression of HNF4α mRNA in chicken kidney tubules has not yet been reported. Therefore we first confirmed the expression of HNF4α in developing chick kidney.

3.2 Expression analysis of HNF4α

The expression of HNF4α mRNA, along with 8 other molecules (Eya1, Six1, Wnt4, Wnt9a, Ret, GFRα1, Sema3A, and Gata3) which are well-established regulators of early kidney development in mouse (Dressler, 2006) and for whom ChEST clones were available in the laboratory, was scanned across HH stages 18, 22, 26, and 28 through WM-ISH. HNF4α mRNA was found to be expressed in kidney tubules at HH stages 22, 26 and 28 (Fig. 3.4). Among the other 8, only GATA3 and GFRα1 exhibited notable kidney-specific expression in the developmental time window that was scanned (Fig. 3.5). Since GATA3 is also expressed in HH26 kidney tubules, and is an established regulator of mammalian kidney development, it is likely that it is also an upstream regulator of our co-expression cluster. However, only HNF4α was followed up for the purpose of this study, for reasons stated previously.

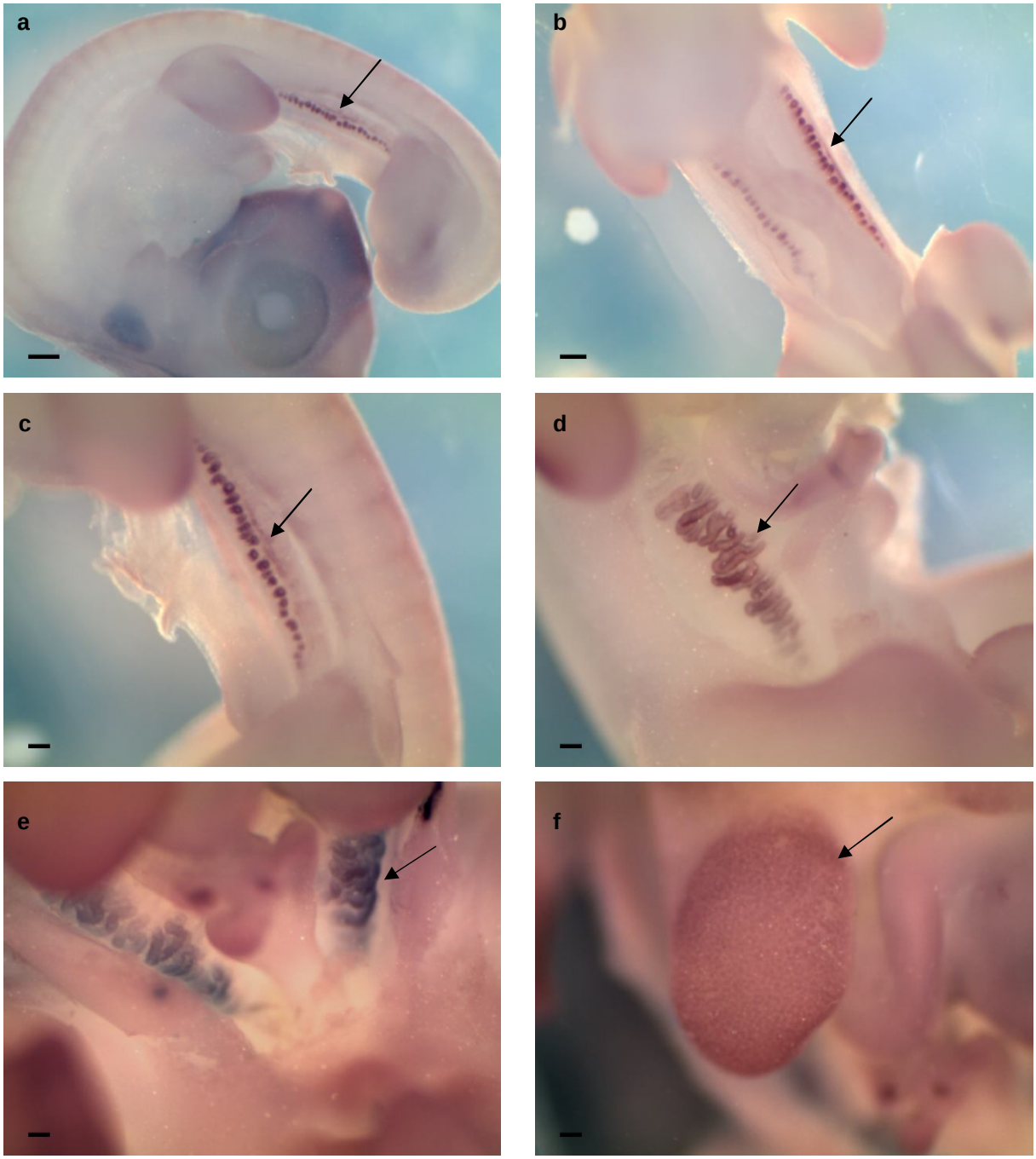


Figure 3.4: Whole-mount *in situ* hybridization for HNF4α. HNF4α mRNA is expressed in kidney tubules at HH22 (a, b, c), HH26 (d), and HH28 (e). Its expression in HH28 liver is also shown (f). Black arrows indicate the region of expression. Scale bar = 4 mm in (a), 2 mm in (b), and 1.4 mm in (c), (d), (e) & (f).

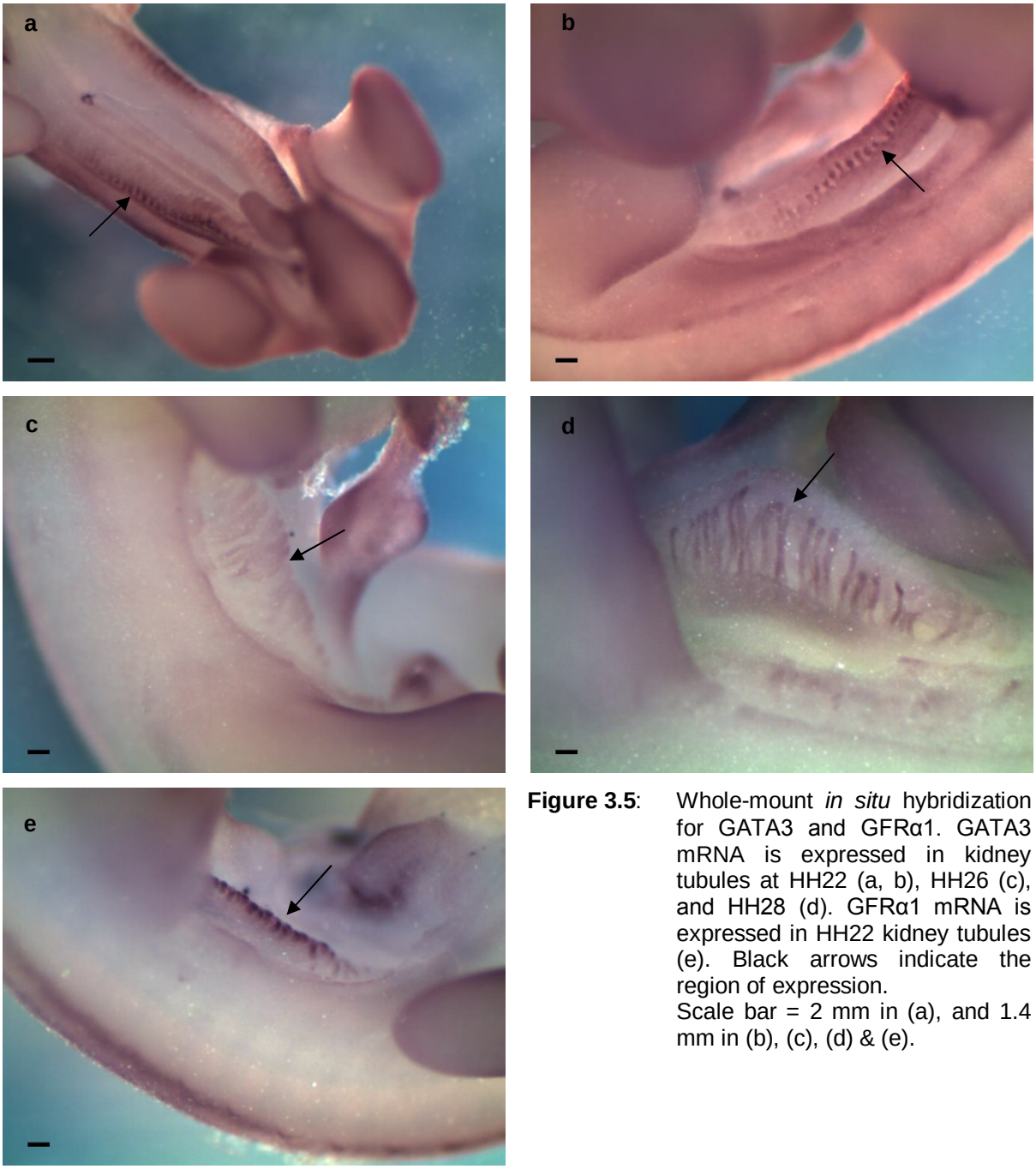


Figure 3.5: Whole-mount *in situ* hybridization for GATA3 and GFRα1. GATA3 mRNA is expressed in kidney tubules at HH22 (a, b), HH26 (c), and HH28 (d). GFRα1 mRNA is expressed in HH22 kidney tubules (e). Black arrows indicate the region of expression.
Scale bar = 2 mm in (a), and 1.4 mm in (b), (c), (d) & (e).

HNF4α expression pattern along with existing information in the literature suggests that HNF4α may be involved in regulating the kidney tubule specific co-expression group. To test this *in vivo* in the context of the developing chicken kidney, we planned to investigate the expression of the genes within this co-expression cluster upon HNF4α loss-of-function. We took two parallel approaches to generate HNF4α loss-of-function in developing chick kidney.

1. retrovirus-mediated RNAi, and
2. retrovirus-mediated misexpression of a dominant negative form of HNF4α (Section 3.4), which would suppress the activity of the endogenous wild-type HNF4α.

3.3 Cloning of miRNAs against HNF4α in pRmiR

After ligation of the double-stranded oligos with pRmiR and transformation, many colonies were obtained for pRmiR-hairpin1, pRmiR-hairpin2, and pRmiR-hairpin3. Positive clones were identified by colony PCR using EmGFP forward and miRNA reverse primers (refer to vector map, Fig. 3.6). Plasmids from two positive colonies per hairpin were isolated through midiprep and presence of the insert was verified again with PCR (Fig. 3.8 and 3.9).

A small but clear difference is apparent between the PCR product formed from empty pRmiR and pRmiR containing hairpin (compare lanes 2 and 3 of Fig. 3.9). The size difference is small because it is equal to the size of the insert i.e. the ds oligo encoding the pre-miRNA, which is only 63 bp in length. The different regions of HNF4α mRNA targeted by the 3 hairpins are depicted in Fig. 3.7.

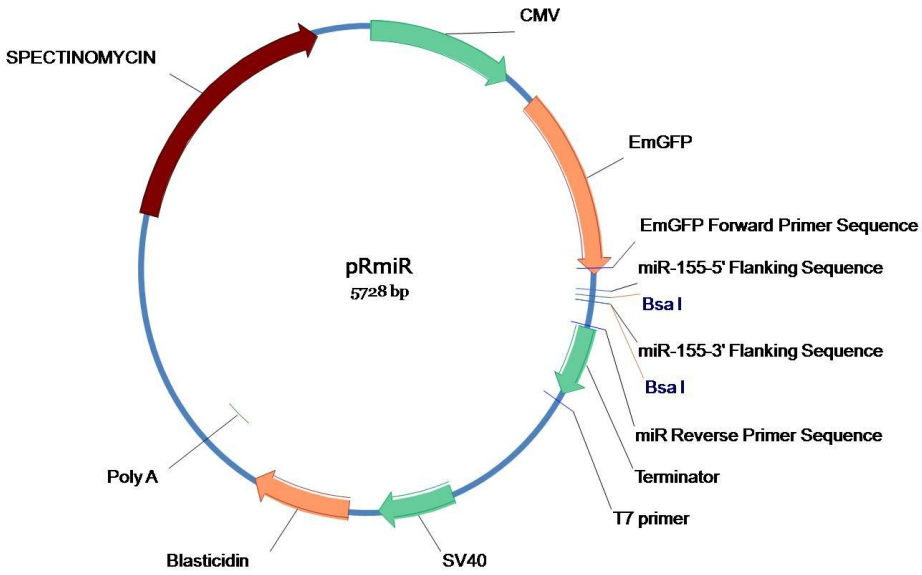


Figure 3.6: Vector map of miRNA expression vector pRmiR (5728 bp) showing important features.

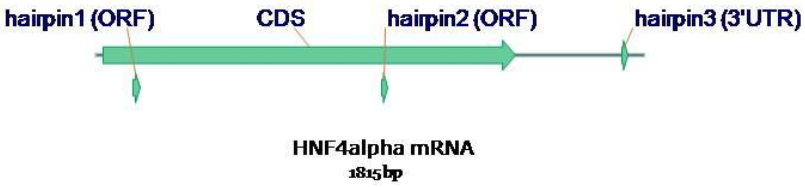


Figure 3.7: Regions of HNF4α mRNA (1815 bp) targeted by the three hairpins.

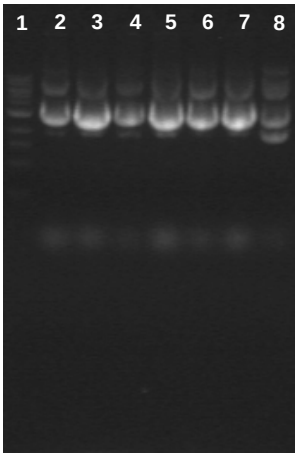


Figure 3.8: Plasmid DNA, isolated from two positive colonies each of the three hairpin constructs.

Lane 1: 1 kb DNA ladder; Lanes 2-3: pRmiR-hairpin1; Lanes 4-5: pRmiR-hairpin2; Lanes 6-7: pRmiR-hairpin3; Lane 8: pRmiR.

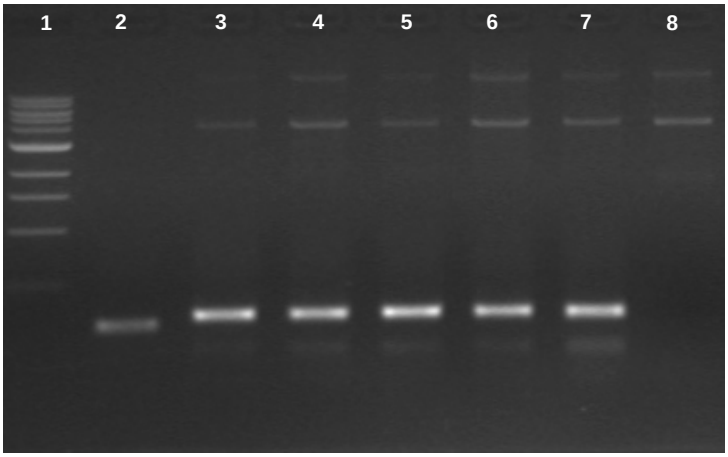


Figure 3.9: PCR from plasmid DNA of the positive clones, with EmGFP forward and miRNA reverse primers.

Lane 1: 1 kb DNA ladder; Lane 2: pRmiR; Lanes 3-4: pRmiR-hairpin1; Lanes 5-6: pRmiR-hairpin2; Lanes 7-8: pRmiR-hairpin3.

3.4 Dominant negative HNF4α

HNF4α (454 amino acids, 50.494 kDa; NCBI Reference Sequence: NP_071516.1) is known to be present in various isoforms formed as a result of splicing variations (Sladek *et al.*, 1999). However, in general, the HNF4α protein has the following functional domains (Cladaras *et al.*, 1997; Sladek *et al.*, 1999):

- An N-terminal ligand-independent transactivation domain.
- A DNA-binding domain, consisting of two zinc fingers.
- A large hydrophobic region containing the dimerization, ligand-binding, cofactor-binding, and ligand-dependent transactivation domains.

The dominant negative mutant of rat HNF4α (Δ111HNF4α) was created in an earlier study by omitting the first 111 amino acids of the wild-type protein, which are a part of the DNA-binding domain (Fraser *et al.*, 1997). This mutant thus lacks the DNA-binding activity but is capable of dimerizing. The mutant upon dimerization with the wild-type HNF4α prevents transcriptional activation by the resultant heterodimer as it also lacks DNA-binding activity (Fraser *et al.*, 1997). As a result, the activity of the endogenous wild-type protein is suppressed. It is important to note that this strategy of making the dominant negative mutant by deleting the DNA-binding domain could be adopted here only because of the unique property of HNF4α of not forming heterodimers with other nuclear receptors. In another study, this dominant negative mutant was used to successfully identify HNF4α-regulated genes in glucose metabolism in a rat insulinoma cell line (Wang *et al.*, 2000). Sequence alignment of the rat HNF4α protein with the chicken protein showed 86% similarity at the amino acid level. Also, there is a perfect overlap between the DNA-binding domains of both the proteins, as annotated on NCBI. Therefore, we adopted an exactly similar strategy for making the dominant negative mutant of chicken HNF4α.

3.5 Isolation and cloning of HNF4α and DNHNF4α into TA vector

Since HNF4α mRNA was found to be expressed in HH22 as well as HH26 kidney tubules, through WM-ISH (Fig. 3.4), embryos of these stages were chosen as the source for isolation of HNF4α (1365 bp) and DNHNF4α (1035 bp) coding sequences through RT-PCR. However, many different attempts had to be made before both the CDS could be amplified. Initially, the RT reaction was carried out at 42°C. But we failed to PCR-amplify both the genes from the cDNA. In order to test the quality of the cDNA, we tried to PCR-amplify GAPDH (a ubiquitous gene) and 6 other genes, which had previously been isolated in the laboratory through PCR from chicken cDNA, including one that was more than 1 kb in length, and all of them were successfully amplified from our cDNA. Then, considering the possibility of optimizing the annealing temperature for PCR, gradient PCR was performed with both sets of primers (HNF4αand DNHNF4α) from 49-61°C, but to no effect. To rule out the possibility of secondary structure formation of template or primers, DMSO was also used in the PCR. Finally, it was speculated that the region to be amplified might be having a higher GC content and therefore higher secondary structure formation in

the RNA which would not allow reverse transcription to happen smoothly at a lower temperature. However, the reverse transcriptase that was used for the cDNA synthesis had an optimal working temperature of 42°C, and a maximum of 55°C. So, an RT gradient was performed from 42-55°C. But it was found that at higher temperatures (>45°C), even GAPDH failed to get amplified from the cDNA. Therefore it was speculated that the reverse transcriptase enzyme was not working efficiently at higher than optimal temperatures. So, a reverse transcriptase that works efficiently at higher temperatures (50-60°C) was procured, reverse transcription was performed at 50°C, and DNHNF4α was successfully amplified at an annealing temperature of 55°C from HH26 cDNA (Fig. 3.10). However, for amplifying HNF4α, a gradient PCR (from 50-65°C) had to be performed and the gene was PCR-amplified at an annealing temperature of 55.5°C from HH26 cDNA (Fig. 3.10). A schematic of both the CDS is shown in Fig. 3.11.

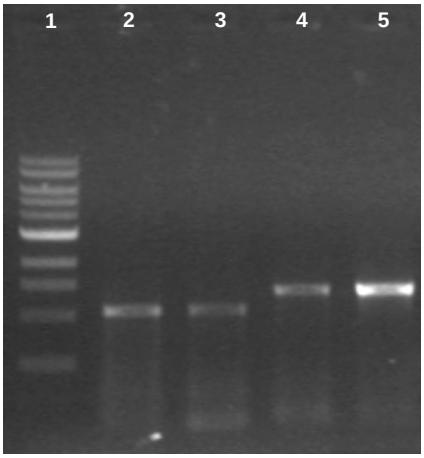


Figure 3.10: PCR-amplification of DNHNF4α (1061 bp) and HNF4α (1365 bp) coding sequences from HH26 cDNA. Lane 1: 1 kb DNA ladder; Lanes 2-3: DNHNF4α; Lanes 4-5: HNF4α.

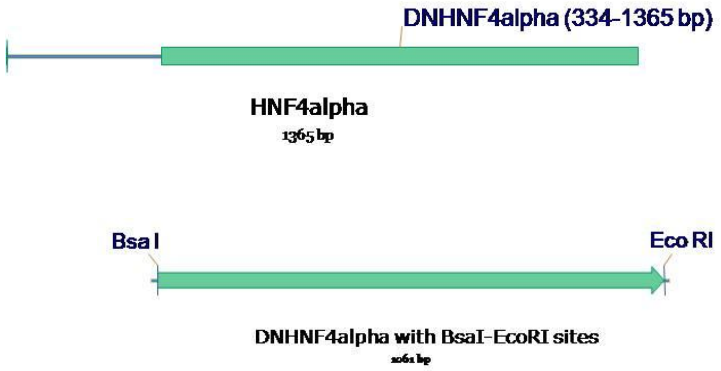


Figure 3.11: Schematic of HNF4α (1365 bp) and DNHNF4α CDS (1031 bp). DNHNF4α is 334 bp short of HNF4α. However, the PCR product is 1061 bp because of incorporation of BsaI and EcoRI sites in the forward and reverse primers, respectively, for subsequent cloning purpose.

After ligation of the HNF4α and DNHNF4α PCR products with pCRII vector, and transformation, one positive clone was identified for each pCRII-HNF4α and pCRII-DNHNF4α through colony PCR with gene-specific and M13 primers (refer to vector maps in Fig. 3.12 and Table 3.1 for expected product sizes). The respective plasmids were isolated through midi prep and the presence and orientation of the

inserts was checked via PCR and digestion (Fig. 3.13). The orientation was found to be as shown in Fig. 3.12.

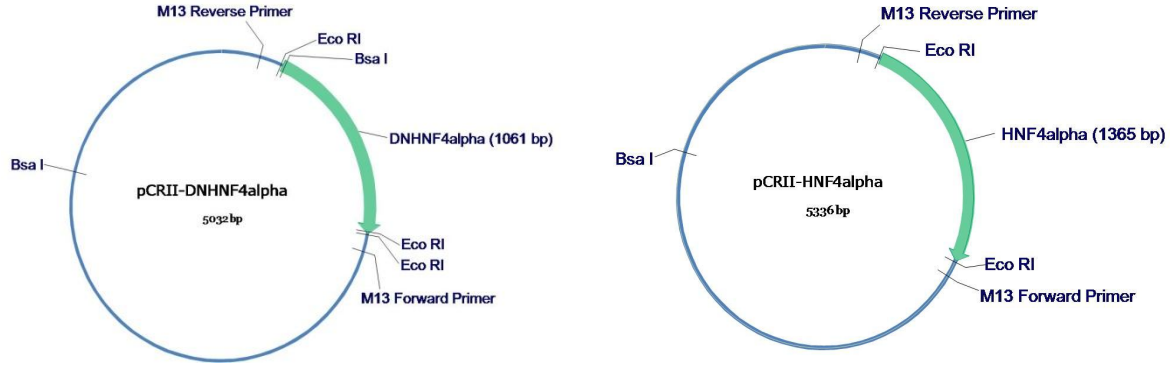


Figure 3.12: Vector maps of pCRII-DNHNF4α (5032 bp) and pCRII-HNF4α (5336 bp) showing relevant features.

Table 3.1: Expected product sizes on PCR and digestion of pCRII-DNHNF4α and pCRII-HNF4α, in accordance with vector maps shown in Fig. 3.12.

	Treatment	**Expected Product Sizes**
pCRII-DNHNF4α	PCR with gene-specific primers	1.06 kb
	PCR with M13 primers	1.30 kb
	PCR with M13 forward and gene forward primers	1.17 kb
	Digestion with EcoRI	1.06 kb
	Digestion with BsaI	1.43 kb + 3.60 kb
pCRII-HNF4α	PCR with gene-specific primers	1.36 kb
	PCR with M13 primers	1.61 kb
	PCR with M13 forward and gene forward primers	1.48 kb
	Digestion with EcoRI	1.38 kb

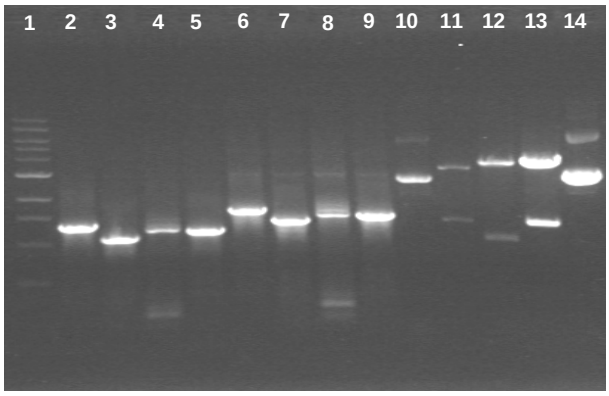


Figure 3.13: Verification of pCRII-DNHNF4α and pCRII-HNF4α plasmid identity through PCR and digestion. Lanes 2-5 are for PCR from pCRII-DNHNF4α and lanes 6-9 from pCRII-HNF4α using different primer combinations. Lane 1: 1 kb DNA ladder; Lane 2: M13 primers; Lane 3: gene-specific primers; Lane 4: M13 reverse and gene forward primers; Lane 5: M13 forward and gene forward primers; Lane 6: M13 primers; Lane 7: gene-specific primers; Lane 8: M13 reverse and gene forward primers; Lane 9: M13 forward and gene forward primers; Lane 10: undigested pCRII-DNHNF4α; Lane 11: pCRII-DNHNF4α digested with BsaI; Lane 12: pCRII-DNHNF4α digested with EcoRI; Lane 13: pCRII-HNF4α digested with EcoRI; Lane 14: undigested pCRII-HNF4α.

3.6 Subcloning of DNHNF4α into pSlax21 and RCAS(A)

For subsequent cloning, DNHNF4α was released from pCRII vector, ligated with pSlax21 vector and transformed. Positive clones of pSlax21-DNHNF4α (vector map in Fig. 3.15) were screened through colony PCR with M13 (expected product size: 1.33 kb) and gene-specific primers (expected product size: 1.03 kb). 20 colonies were screened and all were found to be positive. Plasmids were isolated through midi prep from two positive colonies and verified again through PCR (Fig. 3.14).

Subsequently, the ClaI fragment from pSlax21-DNHNF4α was ligated with RCAS(A). After transformation, colonies were screened for the presence as well as the right orientation of the insert, through colony PCR with RCAS primers (expected product size: 1.3 kb) and that combination of RCAS and gene-specific primers which gives the correct orientation. Out of 27 colonies screened, 23 were positive, of which, 11 had the insert in the correct orientation. RCAS-DNHNF4α plasmid was isolated through midiprep from two of the 11 colonies, and verified again through PCR (Fig.3.16).

Primer combinations which give the right orientation (refer to Fig. 3.17):

1. RCAS forward - gene reverse (expected product size: 1.2 kb)
2. RCAS reverse – gene forward (expected product size: 1.2 kb)

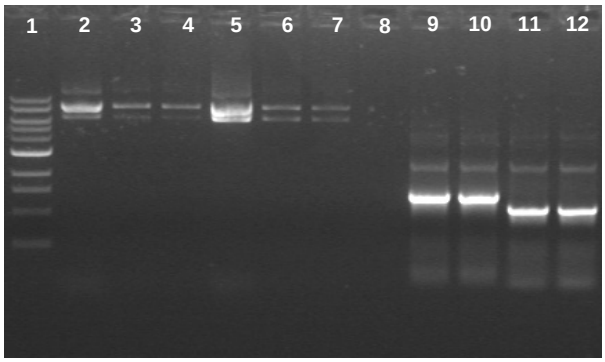


Figure 3.14: Verification of pSlax21-DNHNF4α plasmid identity through PCR. Lane 1: 1 kb DNA ladder; Lanes 9-10: PCR with M13 primers. Lanes 11-12: PCR with gene-specific primers.

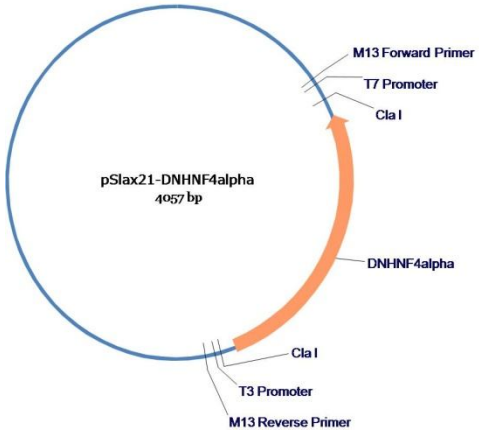


Figure 3.15: Vector map of pSlax21-DNHNF4α (4057 bp) showing relevant features.

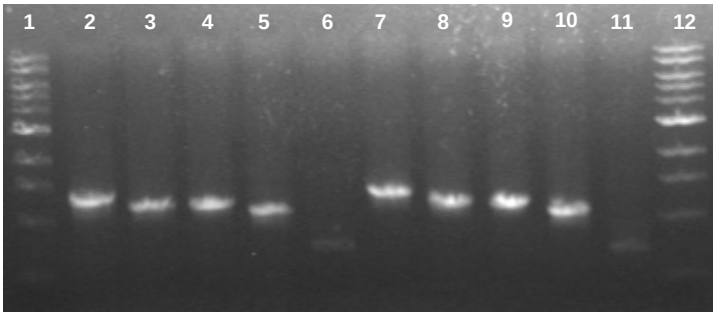


Figure 3.16: Verification of RCAS-DNHNF4α plasmid through PCR with different primer combinations. Lanes 1 & 12: 1 kb DNA ladder; Lanes 2 & 7: RCAS primers. Lanes 3 & 8: RCAS forward and gene reverse primers; Lanes 4 & 9: RCAS reverse and gene forward primers; Lanes 5 & 10: gene-specific primers; Lanes 6 & 11: RCAS forward and gene forward primers.

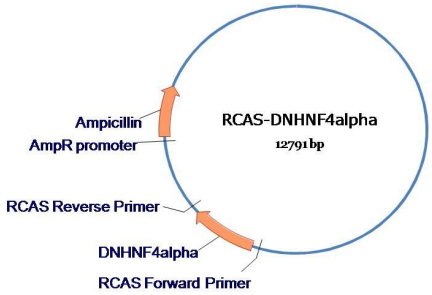


Figure 3.17: Vector map of RCAS-DNHNF4α (12791 bp) showing relevant features. DNHNF4α CDS is 1035 bp long.

Apart from making the RNAi and dominant negative constructs, there was need to optimize the time and site of microinjection of these constructs *in ovo* in the developing chicken kidney, for which high-titer viral stocks of RCAS-AP and RCAS-GFP were made (Fekete & Cepko, 1993).

3.7 AP staining and GFP fluorescence of cells infected with RCAS-reporter (AP/GFP) viruses

Following transfection and viral propagation for a few days, DF-1 cells transfected with RCAS(A)AP, RCAS(B)AP were assayed for AP staining (Fig. 3.18) while cells transfected with RCAS(A)GFP were visualized for GFP fluorescence (Fig. 3.19).

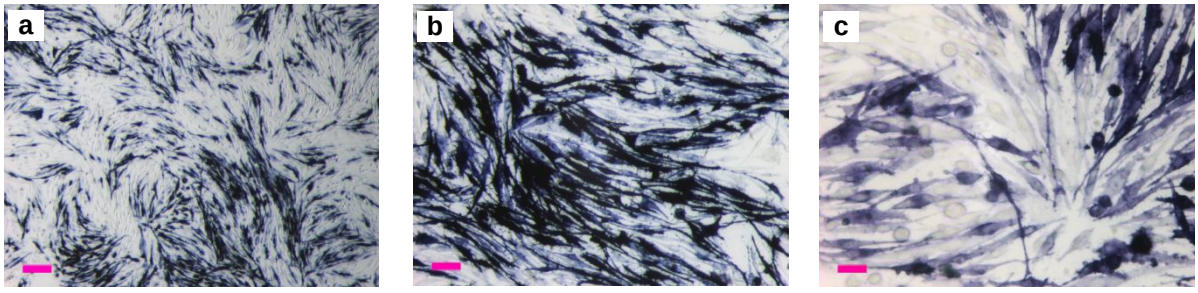


Figure 3.18: Cells infected with RCAS(A)AP. Deep purple staining indicates infected cells. RCAS(B)AP infected cells showed similar staining (image not shown). Scale bar = 140 μm in (a), 60 μm in (b), and 20 μm in (c).

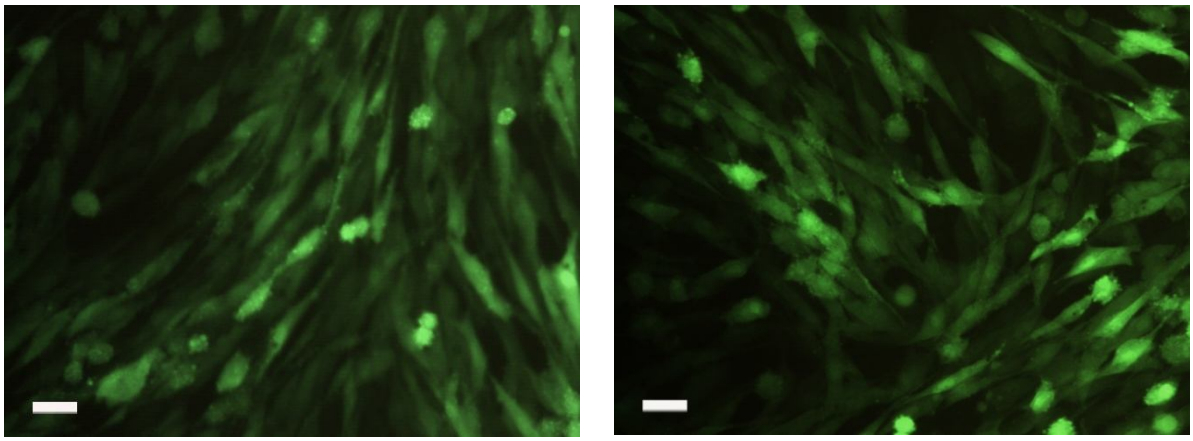


Figure 3.19: Cells infected with RCAS(A)GFP. The green fluorescence indicates infected cells. Scale bar = 20 μm.

The titers of the three viruses, after concentration, were found to be in the range of 10^6-10^8 cfu/ml.

Lastly, semi-quantitative PCR was performed as a preliminary assay to test the efficacy of the pRmiR-hairpin and RCAS-DNHNF4α constructs.

3.8 Testing the efficacy of hairpin constructs and RCAS-DNHNF4α in DF-1 cells through semi-quantitative RT-PCR

Through RT-PCR analysis, using the short-amplicon primers listed in Section 2.7, it was found that HNF4α, along with 3 out of 6 of its putative target genes, namely, ACAA1, DIO1 and FBP1, is expressed in DF-1 cells. Semi-quantitative RT PCR was performed for these 3 genes and HNF4α, following transfection of DF-1 cells with the hairpin constructs, to test the efficiency of knockdown by comparing the changes in expression levels of these transcripts, if any, the results of which are shown in Fig. 3.20.

A very slight lowering in the expression of HNF4α can be seen in pRmiR-hairpin1 - transfected cells as compared to that in the negative control i.e. pRmiR-lacZhairpin - transfected cells (compare lanes 5-7 with lanes 18-20 of Fig. 3.20 (a)). However, it is not a very stark difference to comment on the knockdown efficiency. Also, due to very low levels of PCR product formation of the target genes even at high cycle numbers, it is very difficult to comment on any changes in their expression levels.

The above method was also used to look for any changes in the target gene expression due to dominant negative suppression of HNF4α activity, after transfection with RCAS-DNHNF4α construct. Although a slight increment in expression levels of all genes may be apparent in the cells transfected with DNHNF4α (compare lanes 6-9, 10-13, 14-17 of lower and upper rows in Fig. 3.21) , it is not a very conclusive observation to comment on the efficacy of the dominant negative construct.

The failure to detect any significant depression in the expression levels of HNF4α or appreciable change in target gene levels after transfection with the hairpins may be attributed to a low percentage of cells being transfected with the desired construct. It may also be noted here that this method of testing RNAi efficiency would only detect if there was any significant degradation of the mRNA. However, other means by which miRNAs may effectively act i.e. by causing translational inhibition will not be accounted for through this method. A higher efficiency of knockdown may be achieved by concatamerizing all three hairpins together. Also, for further retrovirus-mediated RNAi studies, these hairpins will have to be cloned into RCAS vectors.

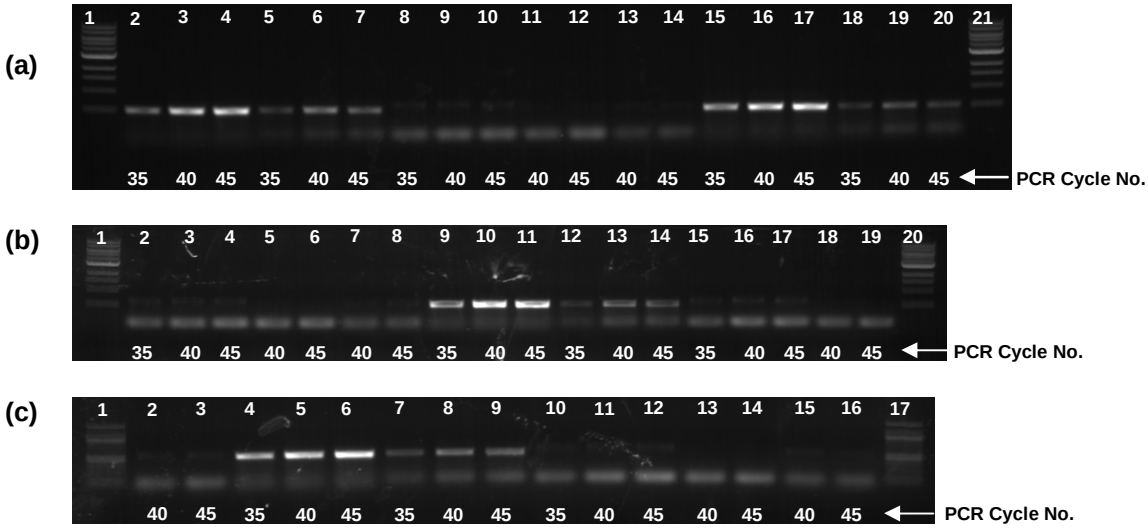


Figure 3.20: Results of semi-quantitative RT-PCR following transfection of hairpin constructs against HNF4α in DF-1 cells. No. of PCR cycles are indicated on the figure.

(a) Lanes 1 & 21: 1 kb DNA ladder; Lanes 2-14: PCR with cDNA from pRmiR-lacZhairpin – transfected cells; Lanes 15-20: PCR with cDNA from pRmiR-hairpin1 – transfected cells. Lanes 2-4: GAPDH; Lanes 5-7: HNF4α; Lanes 8-10: ACAA1; Lanes 11-12: DIO1; Lanes 13-14: FBP1; Lanes 15-17: GAPDH; Lanes 18-20: HNF4α.

(b) Lanes 1 & 20: 1 kb DNA ladder; Lanes 2-8: PCR with cDNA from pRmiR-hairpin1 – transfected cells; Lanes 9-19: PCR with cDNA from pRmiR-hairpin2 – transfected cells. Lanes 2-4: ACAA1; Lanes 5-6: DIO1; Lanes 7-8: FBP1; Lanes 9-11: GAPDH; Lanes 12-14: HNF4α; Lanes 15-17: ACAA1; Lanes 18-19: DIO1.

(c) Lanes 1 & 17: 100 bp DNA ladder; Lanes 2-3: PCR with cDNA from pRmiR-hairpin2 – transfected cells; Lanes 4-16: PCR with cDNA from pRmiR-hairpin3 – transfected cells. Lanes 4-6: GAPDH; Lanes 7-9: HNF4α; Lanes 10-12: ACAA1; Lanes 13-14: DIO1; Lanes 15-16: FBP1.

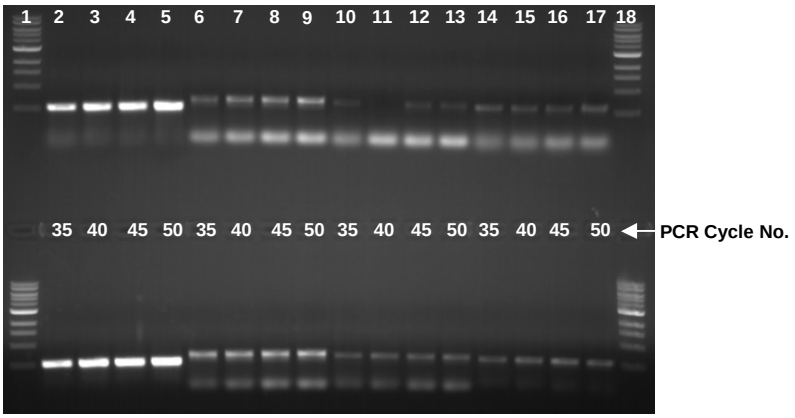


Figure 3.21: Results of semi-quantitative RT-PCR following transfection of RCAS-DNHNF4α in DF-1 cells. No. of PCR cycles are indicated on the figure.

Lower row: Lanes 1 & 18: 1 kb DNA ladder; Lanes 2-17: PCR with cDNA from RCAS-DNHNF4α – transfected cells; Lanes 2-5: GAPDH; Lanes 6-9: ACAA1; Lanes 10-13: DIO1; Lanes 14-17: FBP1.

Upper row: Lanes 1 & 18: 1 kb DNA ladder; Lanes 2-17: PCR with cDNA from RCAS-GFP – transfected cells; Lanes 2-5: GAPDH; Lanes 6-9: ACAA1; Lanes 10-13: DIO1; Lanes 14-17: FBP1.

Similarly, no conclusive difference could be seen in expression of target genes even after transfection with the RCAS-DNHNF4α. This may also be attributed to a lower percentage of cells infected with the retrovirus. Nonetheless, this was a very preliminary test and the constructs need to be tested *in vivo* through microinjections in the developing chicken kidney, followed by *in situ* hybridization analysis of the putative target genes. Many such retroviral-dominant negative constructs have been successfully employed to inhibit protein activity *in vivo* (Sen *et al.*, 2005). Also, successful targeted microinjections have been carried out in the intermediate mesoderm (region of the embryo which gives rise to the urogenital system) of stage 10 embryos (Obara-Ishihara *et al.*, 1999). Knowledge from both these kinds of studies needs to be integrated to achieve the goals stated above.

The expression analysis of metabolic enzymes across four different developmental stages in the chicken revealed dynamic and spatio-temporally regulated expression patterns of these enzymes during development and opened up new avenues for further analysis and hypothesis-testing. The observation that the enzymes sharing a common spatio-temporal domain of expression do not map on to a common biochemical pathway may implicate a novel role of these enzymes which is not related to the biochemical pathway that they are traditionally known to belong to, but, which is critical for the development of the tissue. There are a few evidences to support this speculation. Components of the glycolytic pathway have been implicated in transcriptional regulation (Kim and Dang, 2005). Another enzyme HSD10 has been shown to have a non-enzymatic function important for cellular survival (Rauschenberger *et al.*, 2009). Retinoic acid signaling is a well-established example where metabolic events lie upstream of transcription factor-mediated gene regulation.

Further, microarray data mining analysis of the HH26 kidney tubule co-expression cluster has identified a diverse range of putative upstream regulators of these genes, most of which are transcription factors, HNF4α being one of them. HNF4α, also known as nuclear receptor subfamily 2 group A member 1 (NR2A1), is a highly conserved member of the nuclear hormone receptor superfamily of ligand-dependent transcription factors. It was first identified and isolated from rat liver nuclear extracts, as a protein that bound to the promoters of two liver-specific genes, transthyretin (TTR) and apoliprotein C3 (APOC3) (Sladek *et al.*, 1990).

Nuclear hormone receptors (NHRs) are well-established regulators of endocrine signaling and metabolism (Chawla *et al.*, 2001). They play key roles in embryonic development as well as adult homeostasis. They act as metabolic sensors and respond by interacting with their ligands and regulating expression of specific genes. They bind to specific sites in the promoters of their target genes and nucleate the assembly of transcriptional complexes by recruiting tissue- and gene-specific cofactors (Lin *et al.*, 2005; Yang *et al.*, 2006). In a recent study that produced the first experimentally mapped metabolic gene regulatory network in *C. elegans*, it was found that the network was enriched for NHRs, suggesting a central role for these in metabolic gene regulation (Arda *et al.*, 2010).

HNF4α defines a new subfamily of nuclear receptors since it resides primarily in the nucleus and binds to DNA sequences containing direct repeats (Fraser *et al.*, 1998), exclusively as homodimers, not forming heterodimers with other nuclear receptors like retinoic X receptor α, β, γ, retinoic acid receptor α or thyroid hormone receptor α (Jiang *et al.*, 1995). The ligand-binding domain of the protein is thought to be critical for preventing heterodimerization with other receptors (Jiang *et al.*, 1997).

Apart from liver, HNF4α mRNA has been shown to be expressed in the pancreas, intestine, stomach and kidney (Sladek & Seidel, 2001). Ever since the discovery of HNF4α, its role in the context of liver development and function has been extensively studied with the help of tissue-specific knockouts in mice. It is known as the master regulator of hepatocyte differentiation and expression of liver-specific genes (Li *et al.*, 2000). It is shown to regulate genes encoding enzymes of carbohydrate and lipid metabolism in liver and pancreas (Wang *et al.*, 2000; Hayhurst *et al.*, 2001). In humans, heterozygous mutations in HNF4α are associated with type 1 maturity-onset diabetes of the young (MODY1) characterized by defective insulin secretion (Hattersley, 1998).

Immunohistochemistry studies have detected the expression of HNF4α protein in adult as well as developing mouse kidney proximal tubules (Suh *et al.*, 2006). HNF4α mRNA is also found to be expressed in tubules of E23 murine kidney (GUDMAP (Genitourinary Development Molecular Anatomy Project) ID 12531) (http://www.gudmap.org/). A recent study suggests a possible role of HNF4α in the regulation of cell survival in condensed mesenchyme of the developing mouse

kidney (Kanazawa *et al.*, 2010). Nonetheless, the role of HNF4α in the context of kidney development and kidney-specific gene regulation is largely unexplored. This further underscores the significance of the present study.

A long-term implication of this study would also be in integrating metabolic networks with transcriptional networks, at a systems level, through a bottom-up approach, to create tissue-specific gene regulatory networks (GRNs) which would provide a comprehensive picture of all regulatory interactions between transcriptional regulators and their targets (Fig. 3.22). Further, correlating these regulatory networks with key morphogenetic events marking the development of the tissue will help us in acquiring a holistic view of organogenesis.

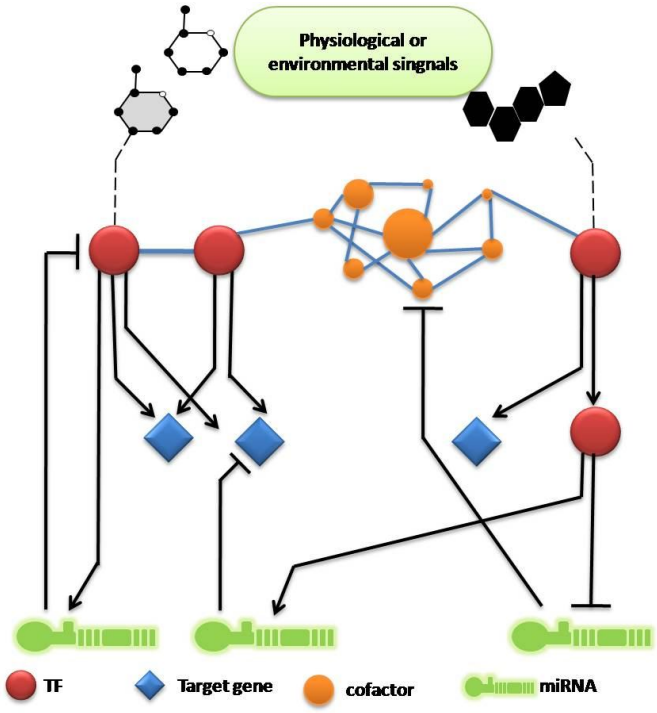


Figure 3.22: Cartoon illustrating hypothetical regulatory interactions in a GRN (Adapted from Arda & Walhout, 2009).

References

Arda, H. E., Taubert, S., MacNeil, L. T., Connie, C. C., Tsuda, B., Gilst, M. V., Sequerra, R., Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Sys. Biol. *6*, 367.

Bosman, F. T., and Stamenkovic, I. (2003). Functional structure and composition of the extracellular matrix. J. Pathol. *200*, 423-428.

Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001). Nuclear receptors and lipid physiology: opening the X-files. Science *294*, 1866–1870.

Cladaras, M., Kistanova, E., Evagelopoulou, C., Zeng, S., Cladaras, C., and Ladias, J. A. A. (1997). Functional domains of the nuclear receptor hepatocyte nuclear factor 4. J. Biol. Chem. *272*, 539-550.

Dressler, G. R. (2006). The cellular basis of kidney development. Annu. Rev. Cell Dev. Biol. *22*, 509-529.

Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., and Griffin, J. L. (2011). Chem. Soc. Rev. *40*, 387-426.

Dzeja, P. P., Chung, S., Faustino, R. S., Behfar, A., and Terzic, A. (2011). Developmental enhancement of adenylate kinase – AMPK metabolic signaling axis supports stem cell cardiac differentiation. PLoS ONE *6*, e19300.

Fekete, D. M., and Cepko, C. L. (1993). Replication-competent retroviral vectors encoding alkaline phosphatase reveal spatial restriction of viral gene expression/transduction in the chick embryo. Mol. Cell Biol. *13*, 2604-2613.

Fraser, J. D., Keller, D., Martinez, V., Mere, D., Straney, R., and Briggs, M. R. (1997). Utilization of recombinant adenovirus and dominant negative mutants to characterize hepatocyte nuclear factor 4-regulated apolipoprotein AI and CIII expression. J. Biol. Chem. *272*, 13892, 13898.

Fraser, J. D., Martinez, V., Straney, R., and Briggs, M. R. (1998). DNA binding and transcription activation specificity of hepatocyte nuclear factor 4. Nuc. Acids Res. *26*, 2702-2707.

Gehring, W. J., and Ikeo, K. (1999). Pax6: mastering eye morphogenesis and eye evolution. Trends Genet. *15*, 371–377.

Gelse, K., Poschl, E., and Aigner, T. (2003). Collagens – structure, function, and biosynthesis. Adv. Drug Deliv. Rev. *55*, 1531-1546.

Grigo, K., Wirsing, A., Lucas, B., Hitpass, L., and Ryffel, G. U. (2008). HNF4 alpha orchestrates a set of 14 genes to down-regulate cell proliferation in kidney cells. Biol. Chem. *389*, 179-187.

Hamberger, V., and Hamilton, H. L. (1951). A series of normal stages in the development of the chick embryo. J. Exp. Morphol. *88*, 49-92.

Hattersley, A. T. (1998). Erratum: Maturity-onset diabetes of the young: clinical heterogeneity explained by genetic heterogeneity. Diabet Med. *15*, 15-24.

Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M., and Gonzalez, F. J. (2001). Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Mol. Cell Biol. *21*, 1393-1403.

http://biotools.umassmed.edu/bioapps/primer3_www.cgi

http://home.ncifcrf.gov/hivdrp/RCAS

http://rnaidesigner.invitrogen.com/rnaiexpress/

http://www.genome.jp/kegg/

http://www.gudmap.org/

http://www.ncbi.nlm.nih.gov/geo/

Huxley-Jones, J., Robertson, D. L., and Boot-Handford, R. P. (2007). On the origins of the extracellular matrix in vertebrates. Matrix Biol. *26*, 2-11.

Jiang, G., Lee, U., and Sladek, F. M. (1997). Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. Mol. Cell Biol. *17*, 6546-6554,

Jiang, G., Nepomuceno, L., Hopkins, K., and Sladek, F. M. (1995). Exclusive homodimerization of the orphan receptor hepatocyte nuclear factor 4 defines a new subclass of nuclear receptors. Mol. And Cell Biol. *15*, 5131-5143.

Joles, J. A., Braam, B., and Verhaar, M. C. (2006). ACE inhibition and glomerular repair: restructuring or regeneration? Kidney Int. *69*, 1105-1107.

Kalluri, R. (2003). Basement membranes: structure, assembly and role in tumor angiogenesis. Nat. Rev., Cancer *3*, 422-433.

Kanazawa, T., Konno, A., Hashimoto, Y., and Kon, Y. (2010). Hepatocyte nuclear factor 4 alpha is related to survival of the condensed mesenchyme in the developing mouse kidney. Dev. Dyn. *239*, 1145-1154.

Kim, J. W., and Dang, C. V. (2005) Multifaceted roles of glycolytic enzymes. Trends Biochem. Sci. *30*, 142-150.

Li, J., Ning, G., and Duncan, S. (2000). Mammalian hepatocyte differentiation requires the transcription factor HNF-4α. Genes and Dev. *14*, 464-474.

Lin, J., Handschin, C., and Spiegelman, B. M. (2005). Metabolic control through the PGC-1 family of transcription coactivators. Cell Metab. *1*, 361–370.

Lozano, P. R., Smith, S. M., Perkins, G., Kubalak, S. W., Boss, G. R., Sucov, H. M., Evans, R. M., and Chien, K. R. (1998). Energy deprivation and deficiency in downstream metabolic target genes during the onset of embryonic heart failure in RXRα$^{-/-}$ embryos. Development *125*, 533-544.

Lucas, B., Grigo, K., Erdmann, S., Lausen, J., Hitpass, L., and Ryffel, G. U. (2005). HNF4α reduces proliferation of kidney cells and affects genes deregulated in renal cell carcinoma. Oncogene *24*, 6418-6431.

Miki, R., Kadota, K., Bono, H., Mizuno, Y., Tomaru, Y., Carninci, P., Itoh, M., Shibata, K., Kawai, J., Konno, H., *et al.* (2001). Delineating developmental and metabolic pathways in vivo by expression profiling using RIKEN set of 18,816 full-length enriched mouse cDNA arrays. Proc. Nat. Acad. Sci. *98*, 2199-2204.

Miller, M. S., Juchau, M. R., Guengerich, F. P., Nebert, D. W., and Raucy, J. L. (1996). Drug metabolic enzymes in developmental toxicology. Fundam. Appl. Toxicol. *34*, 165-175.

Minyon, J., Baek, I., Lee, B. J., Wonyun, Y., and Nam, S. (2011). Dynamic expression of manganese superoxide dismutase during mouse embryonic organogenesis. Int. J. Dev. Biol. *55*, 327-334.

Obara-Ishihara, T., Kuhlman, J., Niswander, L., and Herzlinger, D. (1999). The surface ectoderm is essential for nephric duct formation in intermediate mesoderm. Development *126*, 1103-1108.

Ohguchi, H., Tanaka, T., Uchida, A., Magoori, K., Kudo, H., Kim, I., Daigo, K., Sakakibara, I., Okamura, M., Harigae, H., *et al.* (2008). Hepatocyte nuclear factor 4alpha contributes to thyroid hormone homeostasis by cooperatively regulating the type 1 iodothyronine deiodinase gene with GATA4 and Kruppel-like transcription factor 9. Mol. Cell Biol. *28*, 3917-3931.

Prasad, S., and Bandyopadhyay, A. (2012). GEO data mining program. Department of Biological Sciences & Bioengineering, IIT Kanpur (Unpublished Work).

Rauschenberger, K., Scholer, K., Sass, J. O., Sauer, S., Djuric, Z., Rumig, C., Wolf, N. I., Okun, J. G., Kolka, S., Schwarz, H., *et al.* (2009). A non-enzymatic function of 17β-hydroxysteroid dehydrogenase type 10 is required for mitochondrial integrity and cell survival. EMBO Mol. Med. *2*, 51-62.

Riddle, R. D., Johnson, R. L., Laufer, E., and Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. Cell *75*, 1401-1416.

Roy, P., Singh, A., Kumar, B., and Bandyopadhyay, A. (2012). Developmental metabolomics: in search of correlation between metabolic enzymes and morphogenesis. Department of Biological Sciences & Bioengineering, IIT Kanpur (Unpublished Work).

Sainio, K., Suvanto, P., Davies, J., Wartiovaara, J., Wartiovaara, K., Saarma, M., Arumae, U., Meng, X., Lindahl, M., Pachnis, V., and Sariola, H. (1997). Glial-cell-line-derived neurotrophic factor is required for bud initiation from ureteric epithelium. Development 124, 4077-4087.

Sajithlal, G., Zou, D., Silvius, D., and Xu, PX. (2005). Eya1 acts as a critical regulator for specifying the metanephric mesenchyme. Dev. Biol. 284, 323-336.

Sen, J., Harpavat, S., Peters, M. A., and Cepko, C. L. (2005). Retinoic acid regulates the expression of dorsoventral topographic guidance molecules in the chick retina. Development *132*, 5147-5159.

Shi, J., Mei, W., and Yang, J. (2008). Heme metabolism enzymes are dynamically expressed during Xenopus embryonic development. Biocell *32*, 259-263.

Sladek, F. M., and Seidel, S. D. (2001). Nuclear Receptors and Genetic Disease. Burris, T. B., McCabe, E. R. B. ed. (Academic Press: San Diego), pp. 309–361.

Sladek, F. M., Ruse, M. D., Nepomuceno, L., Huang, S., and Stallcup, M. R. (1999). Modulation of transcriptional activation and coactivator interaction by a splicing variation in the F Domain of nuclear receptor hepatocyte nuclear factor 4α1. Mol. Cell Bol. *19*, 6509-6522.

Sladek, F. M., Zhong, W. M., Lai, E., and Darnell, J. E. Jr. (1990). Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. Genes Dev. *4*, 2353-2365.

Smith, C. A., Roeszler, K. N., Ohnesorg, T., Cummins, D. M., Farlie, P. G., Doran, T. J., and Sinclair, A. H. (2009). The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature *461*, 267-271.

Suh, J. M., Yu, C. T., Tang, K., Tanaka, T., Kodarna, T., Tsai, M. J., and Tsai, S. Y. (2006). The Expression Profiles of Nuclear Receptors in the Developing and Adult Kidney. Mol. Endocrinol. *20*, 3412-3420.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–72.

Thiagarajan, R. D., Georgas, K. M., Rumballe, B. A., Lesieur, E., Chiu, H. S., Taylor, D., Tang, D. T. P., Grimmond, S. M., and Little, M. H. (2011). Identification of anchor genes during kidney development defines ontological relationships, molecular subcompartments and regulatory pathways. PLoS ONE 6, e17286.

Verslues, W. W., and Sladek, F. M. (2010). HNF4α – role in drug metabolism and potential drug target? Curr. Opin. Pharmacol. *10*, 698-705.

Wang, H., Maechler, P., Antinozzi, P., Hagenfeldt, K. A., and Wollheim, C. B. (2000). Hepatocyte nuclear factor 4α regulates the expression of pancreatic β-Cell genes implicated in glucose metabolism and nutrient-induced insulin secretion. J. Biol. Chem. *275*, 35953-35959.

Yang, F., Vought, B. W., Satterlee, J. S., Walker, A. K., Jim, Z. Y., Watts, J. L., DeBeaumont, R., Saito, R. M., Hyberts, S. G., Yang, S., *et al.* (2006). An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. Nature *442*, 700–704.

Yang, X. J. (2002). Retrovirus-mediated gene expression during chick visual system development. Methods *28*, 396-401.