



Università degli Studi di Trieste

PhD program in MOLECULAR MEDICINE

PhD Thesis

Titolo della Tesi

**Molecular Mechanisms of Pancreas
Development and Insulin Regulation in β -Cells**

Federico Paroni

Anno Accademico 2005-2007 (XX ciclo)



UNIVERSITÀ DEGLI STUDI DI TRIESTE

Sezione Didattica
Ripartizione Segreterie Laureati – Segreteria Dottorati di Ricerca
P.le Europa, 1 | 34127 – TRIESTE
Tel. +39 040 5583182 Fax +39 040 5583250 e mail Dottorati@amm.univ.trieste.it

UNIVERSITÀ DEGLI STUDI DI TRIESTE

Sede Amministrativa del Dottorato di Ricerca

UNIVERSITÀ DEGLI STUDI DI TRIESTE

ISTITUTO/CENTRO RICERCHE DIP. BIOCHIMICA BIOFISICA E CHIMICA DELLE MACROMOLECOLE

XX CICLO DEL

DOTTORATO DI RICERCA IN

MEDICINA MOLECOLARE

Molecular Mechanisms of Pancreas Development and Insulin Regulation in Beta-Cells

(Settore scientifico-disciplinare BIO12)

DOTTORANDO

Federico Paroni

COORDINATORE DEL COLLEGIO DEI DOCENTI

CHIAR.MO PROF. *Gianni del Sal* UNIVERSITÀ DEGLI STUDI DI TRIESTE

FIRMA: _____

SUPERVISORE

CHIAR.MO PROF. *Sergio Paoletti* UNIVERSITÀ DEGLI STUDI DI TRIESTE

FIRMA: _____

TUTORE

DR. *Amedeo Vetere* UNIVERSITÀ DEGLI STUDI DI TRIESTE

FIRMA: _____

CORRELATORE

CHIAR.MO PROF. *Susan Bonner-Weir* (HARVARD MEDICAL SCHOOL)

FIRMA: _____

Supervisor: Prof. Sergio Paoletti, Università Università degli studi di Trieste, Dip. Biochimica Biofisica e Chimica delle Macromolecole, via L. Giorgieri 1 34127 Trieste.
paolese@units.it

Tutor: Dr. Amedeo Vetere, Università degli studi di Trieste, Dip. Biochimica Biofisica e Chimica delle Macromolecole, via L. Giorgieri 1 34127 Trieste.
amedeo.vetere@joslin.harvard.edu

External Advisor: Dr. Susan Bonner-Weir, Harvard Medical School, Joslin Diabetes Center, 1 Joslin Place, Boston, MA (USA). Susan.bonner-weir@joslin.harvard.edu

Commissione:

Presidente Effettivo: Prof.ssa Anna Savoia, Università degli studi di Trieste, Dip. di Scienze della Riproduzione, via dell' Istria 65/1, 34137 Trieste.

Componente Effettivo: Prof.ssa Silvia Giordano, Università degli Studi di Torino, Dip. Scienze Oncologiche, strada provinciale 142 di Plobesi, 10060 Candiolo Torino.

Componente Effettivo: Prof. Claudio Brancolini, Università degli Studi di Udine, Dip. Scienze Biomediche, Piazzale Kolbe 4, 33100 Udine.

Presidente Supplente: Prof. Franco Vittur, Università degli studi di Trieste, Dip. Biochimica Biofisica e Chimica delle Macromolecole, via L. Giorgieri 1 34127 Trieste.
vittur@units.it

Componente Supplente: Prof. Stefano Gustincich, Scuole di Studi Avanzati di Trieste, Area Science Park, SS 14 km 163.5, 34012 Basovizza.

Componente Supplente: Prof. Massimo Levrero, Università la Sapienza di Roma, Dip. Medicina Interna, via delle Messi d'Oro 156, 00158 Roma.

OVO Homolog-like 1 (Ovo11) Transcription Factor: a Novel Target of Neurogenin-3

Amedeo Vetere^{1,2}, Federico Paroni^{1,2}, Wan-Chun Li¹, Dohoon Kim¹, Sean Burns¹, Xing Dai³, Arun Sharma¹ and Susan Bonner-Weir¹

¹ Boston,MA; ² Trieste, Italy; ³ Irvine,CA

During pancreas development Neurogenin-3 (Ngn3) plays a pivotal role positioning itself at the beginning of the differentiation pathway to endocrine-committed cell fate. Ngn3 activity is required during a proliferative stage at the interface of first and second transition. The specification of different islet cell types and the completion of the differentiation process require Ngn3 to regulate expression of a series of downstream transcriptional regulators. Transcription factors like NeuroD1, IA1, Pax4 and Nkx2.2 are known direct targets of Ngn3. Here, with a ChIP-cloning strategy, new direct targets of Ngn3 were identified in AR42J cells ectopically expressing Ngn3. One novel target was the zinc-finger type transcription factor OVO-like 1 (OVO11). OVO11, originally implicated in hair formation and spermatogenesis, is suggested to promote the transition from a proliferating, less-differentiated state to a post-mitotic more-differentiated state. Published data suggest that OVO11 is activated by LEF1/ β -catenin complex and then down-regulates proliferation by inhibiting expression of c-myc. By RT-PCR, Western Blot and immunostaining analyses we show that OVO11 is expressed in islets, pancreatic ducts and in pancreatic cell lines INS-1 and AR42J. Furthermore, we show that Ngn3 negatively regulates the expression of a -1100 bp OVO11 promoter:-luciferase reporter construct in an E-box dependent fashion. We suggest that Ngn3's ability to inhibit OVO11, a regulator that promotes transition from proliferation to differentiated cells, may permit Ngn3 to control the balance between proliferation and differentiation of endocrine progenitors. The identification and characterization of both positively and negatively regulated targets of Ngn3 will be key to understand the molecular mechanism regulating conversion of progenitors into endocrine cells.

Accepted Poster for the *American Diabetes Association* (ADA) meeting, 2008 (S.Francisco, CA, USA)

TABLE OF CONTENTS

ABSTRACT	1
RIASSUNTO	2
1 INTRODUCTION	3
1.1 THE PANCREAS	3
1.2 THE PANCREATIC β -CELL.....	4
1.3 PANCREAS DEVELOPMENT	5
1.3.1 EARLY PATTERNING OF THE ENDODERM.....	5
1.3.2 EMBRYONIC PANCREAS DEVELOPMENT	6
1.3.3 PRIMARY AND SECONDARY TRANSITION DURING EMBRYONIC PANCREAS DEVELOPMENT	7
1.3.4 THE MESENCHYME CONTROLS THE EXOCRINE AND ENDOCRINE PANCREAS FORMATION	9
1.3.5 NOTOCHORD SIGNALS ARE REQUIRED FOR PANCREAS DEVELOPMENT.....	11
1.4 HIERARCHY OF TRANSCRIPTION FACTORS IN PANCREAS DEVELOPMENT	15
1.4.1 HEPATOCYTE NUCLEAR FACTOR (HNF) 3BETA/FORKHEAD BOX (FOX) A2	16
1.4.2 HEPATOCYTE NUCLEAR FACTOR (HNF) 3ALPHA/FORKHEAD BOX (FOX) A1	17
1.4.3 HLXB9	17
1.4.4 PANCREATIC DUODENAL HOMEBOX 1 (PDX1)/INSULIN PROMOTER FACTOR 1 (IPF1)	18
1.4.5 HEPATOCYTE NUCLEAR FACTOR (HNF) 6	19
1.4.6 NEUROGENIN3 (NGN3)	19
1.4.7 HAIRY ENHANCER OF SPLIT-RELATED PROTEIN (HES1).....	21
1.4.8 PANCREATIC TRANSCRIPTION FACTOR 1 (PTF1-P48)	21
1.4.9 NEUROD1/ β 2 CELL E-BOX TRANS-ACTIVATOR 2 (BETA2).....	22
1.4.10 PAIRED-BOX HOMEBOX GENES (PAX4/PAX6).....	22
1.4.11 NK CLASS HOMEODOMAIN PROTEIN NKX2.2 AND NKX6.1	24
1.4.12 ISLET 1 (ISL1)	24
1.4.12 INSULINOMA-ASSOCIATED FACTOR 1 (INSMA/IA1).....	25
1.4.13 BRN4 BRAIN 4	26
1.5 ENDOCRINE FATE COMMITMENT	26
1.5.1 NGN3 THE MASTER-SWITCH OF β -CELL DEVELOPMENT	26
1.5.2 NOTCH SIGNALING PATHWAY IN ENDOCRINE DIFFERENTIATION	30
1.6 GLUCOSE HOMEOSTASIS:	33
1.7 DIABETES	34
1.7.1 TYPE 1 DIABETES MELLITUS	36
1.7.2 TYPE 2 DIABETES MELLITUS	38

1.7.3	MATURITY ONSET DIABETES OF THE YOUNG:	40
1.8	GLUCOSE RESPONSE IN ISLET CELLS	41
1.9	INSULIN SECRETION:	42
1.10	INSULIN BIOSYNTESIS	43
1.10.1	TRANSCRIPTION FACTORS BINDING THE INSULIN GENE PROMOTER REGION AND ACTIVATING ITS EXPRESSION:	44
1.10.2	ORGANIZATION OF THE INSULIN PROMOTER REGION AND TRANSCRIPTION FACTORS BINDING SITES:	45
1.10.3	TRANSCRIPTION FACTORS REGULATING INSULIN GENE EXPRESSION IN MATURE β -CELL:	47
3	MATERIALS AND METHODS	51
3.1	CELL LINES CULTURE.	51
3.2	(BIO)CHEMICALS.	52
3.3	BUFFER AND SOLUTIONS	54
3.4	PLASMID MINI-PREP	60
3.5	PLASMID MIDI-PREP	60
3.6	PLASMID MAXI-PREP	61
3.7	GEL EXTRACTION.	62
3.8	BACTERIA TRANSFORMATION	62
3.9	RNA EXTRACTION AND REVERSE TRANSCRIPTION	63
3.10	REAL-TIME RT-PCR	64
3.11	AGAROSE-FORMALDEHYDE ELECTROPHORESIS	65
3.12	PROTEIN QUANTIFICATION	65
3.13	MICROARRAYS ANALYSES	66
3.13.1	CELL CULTURE AND DIFFERENTIATION	66
3.13.2	SYNTHESIS OF AMINOALLYL-AMMINOHEXIL-UTP SSDNA	66
3.13.3	HYDROLYSIS AND NEUTRALIZATION OF THE REACTION	67
3.13.4	PURIFICATION OF THE CDNA	67
3.13.5	ETHANOL PRECIPITATION OF CDNA	68
3.13.6	FLUORESCENT DYE LABELING OF CDNA	68
3.13.7	PURIFICATION OF LABELED CDNA	68
3.13.8	ASSESSING LABELING EFFICIENCY	69
3.13.9	PANCHIP 6.0 ARRAY SLIDES	69
3.13.10	PRE-HYBRIDIZATION OF THE ARRAY SLIDES	71
3.13.11	HYBRIDIZATION OF THE ARRAY SLIDES	71
3.13.12	WASHING OF THE ARRAY SLIDES	72
3.13.13	SCANNER PLATFORM FOR CHIP READING	72
3.13.14	TECHNICAL CHARACTERISTICS VERSARRAY CHIPREADER™	72

3.13.15	SCANNING	73
3.13.16	SPOT DETERMINATION	73
3.13.17	SOFTWARE FOR MICROARRAY STATISTICAL ANALYSIS.....	75
3.13.18	COMPARISON OF THE SIGNALS	76
3.13.19	NORMALIZATION OF THE DATA.....	77
3.13.19.1	TOTAL NORMALIZATION	77
3.13.19.2	LOWESS NORMALIZATION	79
3.13.19.3	VARIANCE REGULARIZATION	80
3.13.19.4	REPLICATE FILTERING.....	81
3.13.19.5	AVERAGING OVER THE REPLICATES EXPERIMENTS.....	82
3.13.19.6	SIGNIFICANCE ANALYSIS OF MICROARRAY (S.A.M.)	83
3.14	IMUNOHYSTOCHEMISTRY	84
3.15	PROTEOMIC ANALYSIS	84
3.15.1	CELL CULTURE AND DIFFERENTIATION	84
3.15.2	PROTEIN EXTRACTION	85
3.15.3	FIRST-DIMENSION: ISOELECTRIC FOCUSING	85
3.15.4	SECOND-DIMENSION: SDS-PAGE	86
3.15.5	SILVER STAINING	87
3.15.16	GEL DIGESTION, PROTEIN REDUCTION, ALKYLATION AND TRYPTIC DIGESTION.....	87
3.15.17	PEPTIDE ANALYSIS	88
3.16	PREPARATION OF THE REV-NGN3 VIRAL PARTICLES.....	88
3.16.1	CLONING THE MOUSE NGN3 CDS INTO THE PREVTRE VECTOR.....	88
3.16.2	SELECTION OF STABLE VIRUS-PRODUCING PT67 PACKAGING CELL LINE	93
3.16.3	DETERMINING THE VIRAL TITER	93
3.16.4	ESTABLISHING STABLE REGULATORY TET-ON-F9 CELL LINES.....	94
3.16.5	ESTABLISHING A DOUBLE-STABLE, INDUCIBLE TET-ON/TRE-NGN3-F9 CELL LINE	95
3.16.6	DOXYCYCLINE-INDUCED EXPRESSION OF NGN3.....	95
3.17	PREPARATION OF AN NGN3-TETRACYCLINE-INDUCIBLE SYSTEM (T-REX SYSTEM).....	96
3.17.1	CLONING THE MOUSE NGN3 CDS INTO THE PCDNA5/TO VECTOR	96
3.17.2	ESTABLISHING A DOUBLE-STABLE, INDUCIBLE 6TR/5TO-NGN3-F9 CELL LINE.....	100
3.17.3	DOXYCYCLINE-INDUCED EXPRESSION OF NGN3.....	101
3.17.4	DIFFERENTIATION PROTOCOL.....	102
3.18	PREPARATION OF AN NGN3-RHEOSWITCH® INDUCIBLE SYSTEM	103
3.18.1	CLONING THE MOUSE NGN3-CDS INTO THE PNEBR-X1 HYGRO VECTOR.....	103
3.18.2	CLONING THE MOUSE NGN3-CDS INTO THE PNEBR-X1-PURO VECTOR	106
3.18.3	REMOVAL OF AN ASPECIFIC ATG CODON FROM THE PNEBR-X1-HYGRO- NGN3 VECTOR	109

3.18.4	CLONING THE IRES AND DTOMATO SEQUENCES INTO THE PNEBR-X1- HYGRO-NGN3/2 VECTOR	110
3.18.5	ESTABLISHING OF INDUCIBLE RHEOSWITCH® F9 CELL LINE	116
3.18.6	CLONES SELECTION.....	116
3.18.7	GAUSSIA LUCIFARASE ASSAY	117
3.18.8	ESTABLISHING OF INDUCIBLE RHEOSWITCH®F9-R1-NGN3 CELL LINE.....	118
3.18.9	TRANSFECTION AND ACTIVATION OF STABLE CELL LINE NIH-3T3-47	118
	(GENERAL PROTOCOL).....	118
3.18.10	WESTERN-BLOT ANALYSIS	119
3.19	PREPARATION OF THE ADNGN3-EGFP AND ADNEUROD/BETA2 VIRAL PARTICLES	120
3.19.1	CLONING THE MOUSE NGN3 CDS INTO THE PSHUTTLE-IRES-HREGFP2 VECTOR	120
	120	
3.19.2	TRANSFORMATION OF THE BJ5183AD CELLS TO PRODUCE RECOMBINANT AD-PLASMID	126
3.19.3	PREPARATION OF PRIMARY ADENOVIRUS STOCK.....	128
3.19.4	AMPLIFICATION OF THE PRIMARY VIRAL STOCK	129
3.19.5	PLAQUE ASSAY.....	130
3.20	CHROMATIN IMMUNOPRECIPITATION CLONING (CHIP-CLONING).....	130
3.20.1	CELL CULTURE AND INFECTION.....	130
3.20.2	<i>IN VIVO</i> CROSSLINKING AND LYSIS	131
3.20.3	IMMUNOPRECIPITATION (IP) OF CROSSLINKED PROTEIN/DNA	131
3.20.4	ELUTION OF PROTEIN/DNA COMPLEXES.....	132
3.20.5	REVERSE CROSSLINKS OF PROTEIN/DNA COMPLEXES TO FREE DNA.....	132
3.20.6	DNA PURIFICATION USING SPIN COLUMNS	133
3.20.7	BLUNTING REACTION, LIGATION AND BACTERIA TRANSFORMATION.....	133
3.20.8	BIOINFORMATICS ANALYSIS.....	134
3.20.9	CONFIRMATION OF THE RESULTS	135
3.21	PREPARATION OF THE OVOL1-PROM1-LUCIFERASE REPORTER SYSTEM AND TRANSACTIVATION EXPERIMENTS	135
3.21.1	CLONING OF THE RAT 1.1 KB OVOL1 PROMOTER INTO THE PGL2-BASIC VECTOR	135
3.21.2	SITE-DIRECTED MUTAGENESIS	139
3.21.3	TRANSACTIVATION EXPERIMENTS.....	141
3.21.4	IMMUNOHISTOCHEMISTRY.....	142
3.21.5	WESTERN-BLOT ANALYSIS	144
3.22	IDENTIFICATION OF THE β -CELL SPECIFIC INSULIN GENE COATIVATOR A2.2.....	146
3.22.1	NUCLEAR PROTEIN EXTRACT FROM HIT T-15 CELLS	146
3.22.2	ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)	147
3.22.2.1	OLIGONUCLEOTIDES LABELING	147

3.22.2.2	GEL MOBILITY SHIFT ASSAY	148
3.22.3	PURIFICATION OF A2.2 COACTIVATOR	149
3.22.4	COLLOIDAL GOLD PROTEIN STAINING	150
3.22.5	WESTERN-BLOT ANALYSIS	150
3.22.6	TCA PROTEIN CONTENTRATION.....	152
3.22.7	MASS SPECTROMETRY ANALYSIS	152
4	RESULTS	154
4.1	MIMICKING THE DIFFERENTIATION PROGRAM OF THE PANCREAS AND IDENTIFY NEW GENES INVOLVED IN ENDOCRINE DEVELOPMENT	155
4.1.1	AR42J-B13 DIFFERENTIATION	155
4.1.2	MICROARRAY ANALYSIS	159
4.1.3	PROTEOMIC ANALYSIS	164
4.2	MIMICKING THE EXPRESSION OF NGN3 DURING THE DIFFERENTIATION PROGRAM	168
4.3	PURIFICATION AND CHARACTERIZATION THE β -CELL INSULIN TRANSCRIPTION FACTOR A2.2 (<i>ALL INFORMATION ARE UNPUBLISHED PROPRIETARY OF JOSLIN DIABETED CENTER</i>)	192
4.3.1	CHARACTERIZATION OF THE TRANSCRIPTION FACTOR A2.2.....	193
4.3.1	PURIFICATION OF A2.2 TRANSCRIPTION FACTOR.....	195
5	DISCUSSION	204
5.1	ARE THE TRANSIENT EXPRESSION OF NGN3 AND THE ENDODERM PATTERNING A PREREQUISITE TO HAVE THE PROPER MOLECULAR CONTEXT TO ACHIEVE β -CELLS DIFFERENTIATION IN VITRO?	209
5.2	WHICH ARE THE DIRECT DOWNSTREAM TARGETS OF NGN3?	213
5.3	IDENTIFICATION OF THE TRANSCRIPTION FACTOR A2.2	226
5.3.1	CHARACTERIZATION OF A2.2	229
5.3.2	PURIFICATION OF A2.2	231
6	CONCLUSIONS.....	234
	ACKNOWLEDGEMENTS.....	260

ABSTRACT

Type 1 or “juvenile” Diabetes is an autoimmune disease in which the insulin secreting β -cells, essential for glucose homeostasis, are destroyed by a target immune attack. To date, patients with type 1 Diabetes can rely only on exogenous insulin injection to control the glucose levels, but, unfortunately, the related chronic and devastating complications cannot be prevented. Identification of β -cell-specific genes playing critical roles during the development will help to create new sources of β -cells from stem/progenitor cells. Interestingly, endocrine cell fate appears to be governed by the “peaked” expression of the transcription factor Ngn3. However, Ngn3’s targets are still not well identified. As for pancreas development also insulin gene regulation lacks important information. *Therefore the aims of this Thesis work have been to test an set-up an efficient system for the transient expression of Ngn3, mimicking its physiological behavior; identify the major factors that lie directly downstream the Ngn3 expression and characterize a novel β -cell specific insulin gene regulator.* As result from this work, an innovative system mimicking pancreatic differentiation and Ngn3 pulsed expression has been developed. Furthermore, a novel Ngn3’s downstream target, the zinc-finger protein “OVO_like 1, has been identified opening a new, and interesting, scenario of Ngn3 gene’s regulation. Last, but not least, the β -cell-insulin-gene regulator A2.2 has been characterized and a reproducible purification system has been developed.

RIASSUNTO

Il Diabete di tipo 1 è una malattia autoimmune caratterizzata dalla distruzione della popolazione di cellule β del pancreas, compromettendo in questo modo il corretto controllo dei livelli di glucosio nel sangue. Attualmente, i pazienti colpiti da diabete di tipo 1 possono essere trattati solamente mediante somministrazione di insulina esogena. Tuttavia questa soluzione non è una cura al problema in quanto non previene le altre croniche e devastanti complicazioni legate alla patologia. Un approccio promettente è quello di creare una nuova popolazione di cellule β a partire da cellule staminali embrionali o adulte. Tale soluzione è però limitata dalla conoscenza dei fenomeni molecolari alla base dello sviluppo del pancreas e più in particolare delle cellule beta. Molti gruppi di ricerca si sono concentrati nell'identificazione di geni coinvolti nello sviluppo embrionale del pancreas. Da tali studi è emerso come il destino delle cellule endocrine sia regolato dall'espressione, tempo e spazio specifica, del fattore di trascrizione Ngn3 (neurogenina 3). Alla luce di questo, risulta di fondamentale importanza l'identificazione dei geni *target* di neurogenina 3 ancora sconosciuti. Come per la regolazione dello sviluppo embrionale, anche la regolazione beta-specifica del gene dell'insulina manca di informazioni fondamentali. Per questi motivi, lo scopo di questa tesi di dottorato è stato quello di: *settare un sistema per l'espressione transiente di Ngn3 in modo da mimarne l'espressione in vivo; identificare nuovi target a valle di Ngn3; purificare e caratterizzare il cofattore A2.2 responsabile, assieme a MafA, della regolazione dell'espressione del gene dell'insulina nelle β -cellule. E' stato quindi sviluppato e testato un nuovo sistema per l'espressione di Ngn3 che permette di mimarne l'espressione fisiologica. Il gene *ovoll* è stato identificato come *target* di Ngn3 nelle beta cellule, rivelando un nuovo ruolo di neurogenina 3 nella regolazione dell'espressione genica. E' stato inoltre caratterizzato il fattore di trascrizione A2.2 e il sistema usato per la sua purificazione è stato reso riproducibile per ulteriori studi.*

Chapter 1

INTRODUCTION

1 Introduction

1.1 The Pancreas

The pancreas is located next to the duodenum and serves two main functions: controlling blood glucose homeostasis and facilitating nutrient uptake.

The organ consists of three major distinct populations of cells: exocrine, endocrine and ductal cells. In the exocrine compartment the acinar cells are organized into small lobules (acini) that constitute most of the pancreas (90-95%), and synthesize digestive enzymes that are secreted through the highly branched system formed by the duct cells into the duodenum. This secretion of about twenty different enzymes (e.g. amylase, trypsin, chymotrypsin, lipases and nucleases) results in digestion of proteins, carbohydrates, nucleic acids as well as fats in the small intestine (Slack 1995). In addition, the duct cells secrete mucins and bicarbonate to neutralize the stomach acid.

The endocrine cells, which represent 2-5% of the adult pancreas, are clustered into small groups known as Islets of Langerhans. Isolated islets consist of approximately 70% β -cells, 25% α -cells, and <5% each of δ -cells, PP-cells as well as smaller numbers of ghrelin-positive endocrine cells, microglial-like cells and vascular endothelium.

In the mouse the insulin producing β -cells, are located in the inner core of islets. These islets are surrounded by a mantle of glucagon producing α -cells, somatostatin producing δ -cells and PP-cells that produce pancreatic polypeptides.

The hormones, produced by the endocrine cells, are secreted into the bloodstream and are transported to their target tissues to control blood glucose homeostasis. In addition the pancreas is innervated by autonomic nerves that regulate hormone secretion during feeding and fasting (Fig.1).

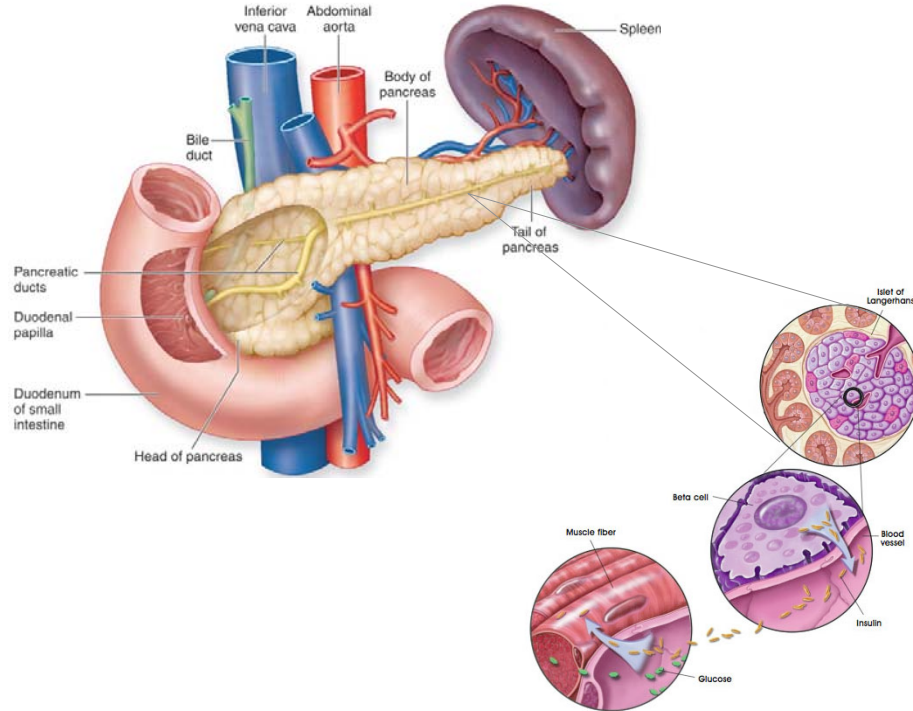


Fig.1: the pancreas is located in the abdomen adjacent to the duodenum. The endocrine portion, islets of Langerhans surrounded by acinar tissue, secrete its hormones production into the bloodstream to control glucose homeostasis. Image modified from <http://academic.kellogg.cc.mi.us/>.

1.2 The pancreatic β -cell

Within the islets of Langerhans, the β -cells are by far the most abundant. Through secretion of appropriate quantity of insulin in response to the glucose levels, the β -cells are mainly responsible for the maintenance of the body's glucose levels within a very narrow range.

Like the other pancreatic endocrine cells, β -cells develop from endodermal progenitor cells. The identity of the progenitor cells and of the transcription factors involved in the generation of functional β -cells is of enormous interest and considerable research effort is spent in this field (Edlund 2001).

The most significant increase in growth of β -cell mass occurs at a period of late fetal life. Unfortunately, assessment of β -cell mass is not a straightforward method and caution must be taken when looking at such data. Any limitations on β -cell replacement are thought to have direct consequences for glucose homeostasis. Throughout the lifetime of a mammal, low levels of β -cell replication (2–3% per day) seem to balance β -cell loss by apoptosis (Habener and Stoffers 1998). This results in a slow increase in mass (Bonner-Weir 2001). Interestingly the most significant increase growth of β -cell mass occurs at a period of late fetal life.

1.3 Pancreas development

1.3.1 Early patterning of the endoderm

During the embryonic development the process of gastrulation subdivides the embryo into three groups of cells or germ layers: the ectoderm that will form the skin and the central nervous system (CNS); the mesoderm that will form blood, the bones and the muscles; the endoderm that will form the respiratory and digestive tracts. The definitive endoderm is recruited from the epiblast during the gastrulation. Following ingression through the primitive streak, it displaces the visceral endoderm to the extra-embryonic yolk sac (Tam and Beddington 1992). The definitive endoderm forms the epithelium of the primitive gut tube from which the epithelial lining of the respiratory and gastrointestinal tract as well as liver, lungs, pancreas, thyroid and thymus develop (Wells and Melton 1999).

At embryonic day 7.5 (e7.5), the endoderm layer in mice is a cup of approximately 500-1000 cells enveloping the mesoderm and ectoderm, both of which send instructive signals to the endoderm making it competent to respond to subsequent permissive signals that will establish the organ domains along the anterior-posterior axis (A-P) (Wells and Melton 1999; Wells and Melton 2000).

For example, Fibroblast Growth Factor(s) (FGF) signaling is required for establishing and maintaining gut tube domains along the A-P axis. FGF4 is expressed in the mesoderm and ectoderm adjacent to the developing midgut and hindgut endoderm, and has been demonstrated to promote a posterior gut cell fate and to repress an anterior endoderm cell fate. Patterning of the early endoderm by FGF4 is necessary to define a window of competence to pancreatic cell fate.

The first sign of pancreatic morphogenesis is the elongation of the endoderm cells so that they assume a columnar shape within the region of incipient development. All three-cell lineages of the pancreas (ducts, acini, islets) develop from a common pool of pancreatic duodenal homeobox 1-expressing (Pdx1/Ipf1) progenitor cells. Pdx1 is a parabox homeobox-containing transcription factor that is turned on right after the endoderm begins to form a primitive gut tube by day 8.5 (e8.5). Also, at this time the region destined to become the pancreas receives signals from two mesoderm-derived structures: the adjacent notochord and the dorsal aorta, which leads to the expression of Pdx1.

Several cell fate choices are important for the development of the nearly homogeneous progenitor cell population into the distinct cells of the complex organ of the mature pancreas. It has been shown by Gu et al. (Gu, Dubauskaite et al. 2002), using lineage tracing, that cell fate choices occur in Pdx1 positive cells at e9.5-e11.5, assigning cells to either a ductal fate or an exocrine/endocrine fate. This window was the only time when ductal differentiation was observed, whereas at e8.5 or e12.5 and anytime later only islets and acini cells are formed.

1.3.2 Embryonic pancreas development

After e10.5 the pancreatic tissue expands rapidly from a compact structure of tightly packed cells and grows several folds in size, forming a branched structure into the surrounding mesenchyme (Guz, Montminy et al. 1995; Miralles, Battelino et al. 1998; Miralles, Czernichow et al. 1998). The branching morphogenesis produces a ductal tree

out of the pancreatic epithelium. This is made through intraepithelial lumen formation in which multiple microlumens eventually coalesce to form continuous lumen from which the ductal network is established (Hogan and Kolodziej 2002; Jensen 2004).

The fusion, at e16.5-e17.5 in mice, of the dorsal and ventral buds made up to the definitive pancreas. In humans and in some other species (rabbit, pig, cat and sheep) the ventral pancreas is formed also from the fusion of two lateral buds. The dorsal and ventral pancreatic genetic program is not completely synchronized. This is due to the fact that they are two independent endoderm domains that receive distinct signals from the surrounding tissues (Fig 2).

1.3.3 Primary and secondary transition during embryonic pancreas development

Rutter et al., in the 1970s (Pictet, Clark et al. 1972) proposed a model for pancreatic differentiation in which two successive regulatory transitions occur: the primary transition refers to the start of the pancreas budding and expansion of the progenitor cell pool, while the secondary transition describes the formation and differentiation of endocrine hormone expressing cells (Fig. 2).

During the primary transition (e8.5), the only differentiated cell type present is glucagon producing α -cells, neither exocrine nor β -cells are identified at this time. These early glucagons expressing cells are Pdx1-independent. Mesenchymal cells accumulate around the primitive gut and form a blanket over the pancreatic diverticulum.

The secondary transition starts at e13.5, peaks at e14.5 and ends at e15.5, during this period the first population of fully differentiated β -cells appears, but there is also an increase of exocrine proteins production. After the secondary transition few progenitor cells are left (Fig.2).

Double Glucagon- and insulin-positive cells have been observed in early pancreatic buds. Nevertheless studies by Chiang and Melton (Chiang and Melton 2003) showed that insulin-positive cells have never active the glucagon gene and vice-versa, so

double positive cells are different compared with the mature endocrine cells. Furthermore studies by Herrera et al. (Herrera, Orci et al. 1998) shown that mature insulin positive cells have the PP-gene turned on, pointing out the possibility that insulin positive cells and PP-cells could develop from a common progenitor cell type.

It has been shown that endocrine cell fate choice occurs temporally with glucagon expressing α -cells being the first to develop, thereafter β -cells (e13.5) δ -cells (15.5) and PP cells (e17). Once differentiated, the endocrine cells delaminate from the squamous epithelium and migrate into the extracellular matrix, recognizing each other and aggregating into a defined structure that is the Islet of Langerhans. This clusterization of endocrine cells and the subsequent Islet formation begin right before the birth (e18.5 in mouse). Finally the formation of capillary networks occurs inside islets due to newly formed islets producing VEGF-a (vascular endothelial growth factor) that ensure that the islets are close to capillaries for the secretion of the hormones into the bloodstream.

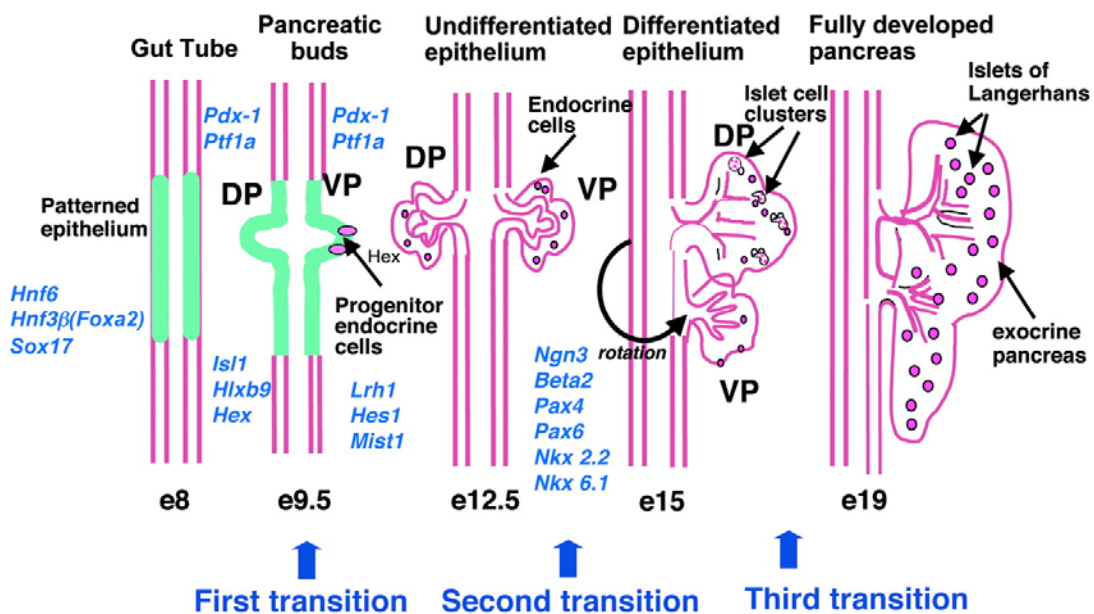


Fig. 2: Schematic diagram of the pancreas development in mouse. The pancreas develops at the foregut/midgut junction from a pre-patterned region of the primitive endoderm. The first morphological indications are around day 9.5 with the emergence of two evaginations, the pancreatic buds, first the dorsal bud, and subsequently the ventral bud (first transition). The pancreatic epithelium proliferates invading the surrounding mesenchyme (second transition) and around the e18 they are going to fuse, after the rotation of the ventral portion and give rise to the mature organ (third transition). Hebner et al, 2005

1.3.4 The mesenchyme controls the exocrine and endocrine pancreas formation

It was long suspected that the pancreas, like other organs such as the lung of the kidney, develops by interactions between the epithelium and its surrounding mesenchyme.

Insulin expressing cells developed *in vitro* from embryonic pancreatic epithelium in the presence of mesenchyme are immature, whereas those developed without mesenchyme become mature (Miralles, Czernichow et al. 1998; Miralles, Czernichow et al. 1999). The cells developed in the presence of mesenchyme resemble to insulin-expressing cells developed *in vivo* before e15 in rat embryos but those developed without mesenchyme resemble to cells developed *in vivo* at a later stage (Roberts, Smith et al. 1998). Thus, it seems that different factors control the development of these cell populations. So, it is important to redefine whether a relationship exists between insulin-expressing cells developed before e15 and those that appear after.

Many studies have confirmed such hypothesis showing that signals derived from the mesenchyme are implicated in the control of pancreatic cell differentiation into exocrine or endocrine tissue.

It is now clear that the differentiation of the epithelium into acinar tissue is mesenchyme-dependent (Pictet, Clark et al. 1972; Gittes, Galante et al. 1996; Miralles, Czernichow et al. 1998; Miralles, Czernichow et al. 1999). Amylase- or carboxypeptidase A-expressing cells develop normally from pancreatic embryonic epithelium growth *in vitro* in presence of mesenchyme but in its absence their growth is strongly repressed. *In vitro* this effect has been shown to be reproduced by the ligands of FGF1, FGF7, and FGF10. Nevertheless, it is still unclear whether FGF1, FGF7, FGF10 induce the differentiation of precursor cells into amylase-expressing cells or act as growth factors for pre-existing amylase-expressing cells. Recently experiments on hamster diabetic models have shown that FGF1 has the highest effect on exocrine cells during the post-natal period.

The implication of signals from the mesenchyme in the control of endocrine tissue development is more controversial. It was shown that when adult ductal epithelial cells are grafted into immuno-incompetent mice in the presence of fetal mesenchyme, insulin- and glucagon-expressing islets develop, whereas such islets did not develop in the absence of mesenchyme.

These experiments strongly suggested that the mesenchyme is necessary for the development of endocrine tissue from epithelial cells, at least during postnatal life (Dudek and Lawrence 1988; Dudek, Lawrence et al. 1991) even though some other evidences shown that the presence of mesenchyme tissue seems not necessary during the embryonic development (Gittes, Galante et al. 1996). Therefore it is possible that the mechanisms controlling endocrine cell differentiation during prenatal and postnatal life could be different.

On the contrary, other experiments have showed that the mesenchyme could be not necessary for the pancreas development, whereas it seems to act as a repressor of endocrine tissue development (Miralles, Czernichow et al. 1998). Based on these data (Gittes, Galante et al. 1996; Miralles, Czernichow et al. 1998) it has been proposed that the default pathway for the epithelial primordium leads to islet and extrinsic signal would exist to exert a negative control on the proportion of endocrine tissue to be generated. A search for specific repressing endocrine-cells differentiation factors expressed in the mesenchyme has been done.

The hypothesis was that the epithelium would produce factors, such activins or bone morphogenic proteins (BMPs), which would act in an autocrine manner to favor differentiation of pancreatic precursor cells into endocrine cells. The hypothesis is supported by the fact that 1) activin and BMP-7 have been found to be expressed in the pancreatic anlage, 2) transgenic mice expressing mutants forms of the activin receptors suffer of hypoplasia of the endocrine portion of the pancreas, 3) AR42J cell line, that is considered the election model of pancreatic precursor cells, needs activin A and HGF or Betacellulin to promote the differentiation onto insulin positive phenotype (Furukawa, Eto et al. 1995; Ritvos, Tuuri et al. 1995; Mashima, Ohnishi et al. 1996; Mashima, Shibata et al. 1996; Yamaoka, Idehara et al. 1998; Shiozaki, Tajima et al. 1999). The possible involvement of activin(s) in the control of the differentiation of precursor cells

into endocrine cells was tested indirectly by investigating the role of follistatin (Nakamura, Takio et al. 1990) in the regulation the relative proportions of endocrine and exocrine tissue during pancreatic development.

It has been shown that exogenous follistatin increases the development of immature pancreatic epithelial cells into amylase-expressing cells while repressing the development of insulin-expressing cells (Miralles, Czernichow et al. 1998). Exogenous follistatin, thus, mimics the effects of the mesenchyme. On the other hand, because mice deficient in follistatin are normal (Edlund 1998), other factors with a function identical to that of follistatin should be implicate in the control of endocrine pancreas development.

1.3.5 Notochord signals are required for pancreas development

The notochord is an axial structure of mesoderm origin and its presence during embryonic development defines members of the Chordate phylum. The notochord is one of the earliest embryonic structures to be formed and functions as structural support for the entire organism, either transiently (higher vertebrates) or persistently (some lower vertebrates). Many different studies have shown that in addition to its structural function, the notochord plays a critical role in the patterning of ectoderm and mesoderm tissues such as the neural tube and somitic derivatives.

Experiments using chick embryo have provided strong evidence that the notochord plays a role in development of the dorsal pancreas at the 6-13 somites stage (Wessells and Cohen 1967; Spooner, Walther et al. 1970; Dieterlen-Lievre and Beaupain 1974; Sumiya and Mizuno 1987). In chick embryo, premature separation of the notochord from endoderm abolishes expression of pancreatic marker genes, while co-culture of isolated notochord with endodermal epithelium restores marker expression. Interestingly, the notochord is only capable of inducing pancreas marker in the anterior but not posterior endoderm, indicating that the endoderm is already pre-patterned at this early stage (Kim, Hebrok et al. 1997).

Studies by Wells and Melton (Wells and Melton 2000) support this pre-patterning hypothesis by showing that factors produced in the overlying mesoderm and ectoderm regionally specify endoderm at the late gastrulation stage. At the time when pancreas formation is initiated, the notochord secretes a variety of potent inducing molecules including the TGF- β family member activin- β B and FGF2 (Fig. 3).

Tissue culture experiments with chick embryos have revealed that both ligands can mimic the notochord effect and induce transcription of pancreatic marker genes in isolated endoderm. Isolated notochord, or purified activin and FGF activate pancreatic gene expression by repressing expression of Sonic hedgehog (Shh), a secreted protein of the Hedgehog family in the dorsal pre-pancreatic epithelium (Hebrok, Kim et al. 1998). The notochord also produces Shh and induces Shh expression in overlying floorplate cells in the neural tube. The apparent embryologic paradox that notochord both induces and represses Shh expression may be explained by the evidence that dorsal endoderm and ventral neural tube have different competence for Hedgehog, FGF and TGF- β signals (Briscoe, Pierani et al. 2000) (Fig. 3). For example expression of activin receptor IIB has been detected in chick endoderm and dorsal neural tube but not ventral neural tube (Stern, Yu et al. 1995) just before morphogenesis. Previous studies have established roles for activin and Hedgehog signaling in patterning foregut gene expression and organogenesis in mice using mutations of the activin receptors (Kim, Hebrok et al. 2000). The pattern of malformations in these mutants was consistent with an anterior transformation of organ fates in the foregut and correspond with antero-posterior shifts in expression genes including Shh and *Isl1*.

The role of Shh in establishing pancreatic development has also been tested extensively. In the mid gestational mouse embryo, Shh is expressed in nearly all epithelial cells lining the alimentary canal (Echelard, Epstein et al. 1993; Bitgood and McMahon 1995; McMahon 2000; Ramalho-Santos, Melton et al. 2000), with the important exception that Shh is excluded from the oral epithelium primordium of the anterior pituitary and from pancreatic epithelium (Hebrok, Kim et al. 1998; Treier, Gleiberman et al. 1998; Apelqvist, Li et al. 1999). In mouse and chicks, ectopic expression of Shh in embryonic pancreatic epithelium results in loss of pancreas marker expression and transformation of pancreatic mesenchyme into the gut mesoderm

(Hebrok, Kim et al. 1998; Apelqvist, Li et al. 1999). Inhibition of hedgehog signaling on isolated endoderm foregut, induce pancreatic gene expression (Hebrok, Kim et al. 1998). Thus, spatially restricted expression of Shh provides permissive condition leading the appropriate organogenesis along the A-P axis.

Anterior-posterior patterning in mesoderm and ectoderm is regulated by spatially restricted expression of Hox transcription factors (Krumlauf 1994) that are encoded in defined genomic regions, the Hox clusters. In the hindgut, endodermal Shh activate expression of Hox genes in overlying mesoderm (Roberts, Johnson et al. 1995; Roberts, Smith et al. 1998). However few Hox genes are expressed in endoderm-derived epithelium and mutations in Hox genes have not yet been shown to result in A-P transformations of endoderm-derived organs in vertebrates. A set of Hox-related genes that are expressed in endoderm make up the ParaHox cluster, first identified in *Amphioxus* (Brooke, Garcia-Fernandez et al. 1998). The expression of *Ipfl/Pdx1*, a ParaHox member, is not restricted to the pancreatic tissue, and a dorsal-ventral bud outgrowth is initiated in *Pdx1* mutant animals (Offield, Jetton et al. 1996), suggesting that other factors are required to specify the pancreas anlage.

A potential candidate is *Hlxb9*, a homeobox transcription factor that is expressed in the dorsal bud before *Ipfl/Pdx1* and whose function is required for the dorsal but not ventral bud morphogenesis (Harrison, Thaler et al. 1999; Li, Arber et al. 1999). Expression of Shh and *Ihh* (Indian Hedgehog) in the pancreas is not affected in the *Hlxb9* mutants (Li, Arber et al. 1999) but remains to be determined if *Hlxb9* expression is regulated by endodermal Hedgehog signals.

While some progress has been made in understanding the signaling pathways that regulate early stages of dorsal pancreas development less is known about the development of the ventral pancreas. The notochord does not touch the lateral endoderm that will give rise to the ventral portion of the future pancreas. Nevertheless, Shh expression is repressed in that portion of lateral endoderm (Kim, Hebrok et al. 1997). During dorsal pancreas formation, Shh remains excluded from dorsal pancreatic tissue even after notochord and endoderm have been separated by the fusion of the paired dorsal aorta (Kim, Hebrok et al. 1997; Hebrok, Kim et al. 1998; Apelqvist, Li et al. 1999). Thus, separate mechanisms may initiate and maintain Shh repression in dorsal pancreatic

endoderm (Fig 3). The ratio of endocrine to exocrine cell mass is approximately 100:1 in mice and in human, but varies in other species, suggesting that genetic mechanisms regulate this ratio. Early studies (Golosow and Grobstein 1962; Wessells and Cohen 1967) showed that pancreatic exocrine development require mesenchymal signals that can be replaced by cell free extracts (Golosow and Grobstein 1962).

Conclusions about the role of cell interactions in endocrine differentiation were limited in these studies because of their dependence on morphogenic markers of development. Soon after, other studies using molecular markers of endocrine and exocrine cells showed that the removal of pancreatic mesenchyme impairs the exocrine development but on the other side promotes the endocrine development (Rutter, Pictet et al. 1978; Gittes, Galante et al. 1996; Miralles, Czernichow et al. 1998). These data support the hypothesis that the epithelial-mesenchymal interactions regulate the appropriate balance of endocrine and exocrine development.

More recent experiments showed that intercellular signaling of TGF- β , Notch and hedgehog pathway are important for appropriate specification of endocrine and exocrine development. TGF- β signaling is a major regulator of pancreatic endocrine and exocrine development. TGF- β signaling components (Massague and Chen 2000) including the TGF- β ligands activin and TGF- β 1, the corresponding receptors, antagonist including follistatin, noggin and gremlin, and intracellular Smads proteins like Smad2 and Smad4 are expressed in embryonic pancreas epithelium and mesenchyme (Ogawa, Abe et al. 1993; Furukawa, Eto et al. 1995; Manova, De Leon et al. 1995; Verschueren, Dewulf et al. 1995; Crisera, Maldonado et al. 2000; Tremblay, Hoodless et al. 2000). It has been shown that *in vitro* exposure of embryonic mouse pancreas to activin of TGF- β 1 promotes the development of endocrine cells and in particular of β -cells and PP-cells (Sanvito, Herrera et al. 1994) but repress the epithelial branching and acinar formation (Ritvos, Tuuri et al. 1995).

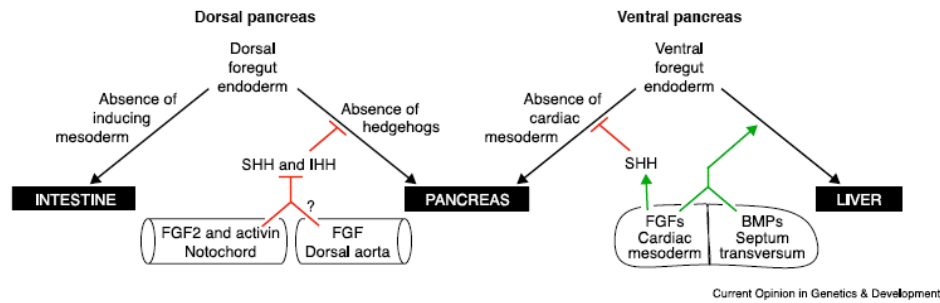


Fig. 3: Signals from mesodermal derived structures such as nothocord, dorsal aorta, cardiac mesoderm and septum transversum regulating early dorsal and ventral pancreas development. *Kim&McDonald, Current Opinion in Genetics & Development, 2002.*

1.4 Hierarchy of transcription factors in pancreas development

The transcription factors involved in pancreas development belong to three different families: the homeodomain family, the basic helix-loop-helix (bHLH) family and the windged helix family of proteins (Schwitzgebel 2001). Several recent gene disruption studies in mice have helped to elucidate the hierarchy of transcription factors in pancreas development. The following paragraph describes each transcription factors, its onset of expression, its function and the studies done in mice knockout models (Fig. 4).

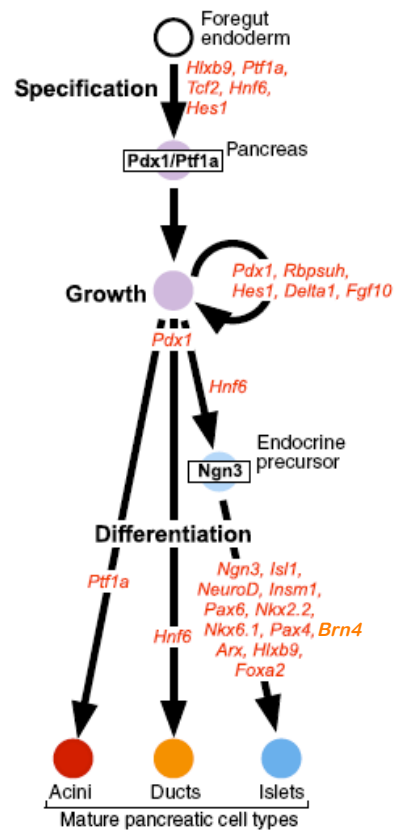


Fig. 4: Simplified hierarchy of transcription factors during pancreas development. The model is based on onset of expression of each factor and function analyzed by experiments of targeted disruption of the corresponding gene. *Murtaugh et al. Development, 2007.*

1.4.1 Hepatocyte nuclear factor (HNF) 3beta/Forkhead box (Fox) a2

HNF-3 β , also called Foxa2, belongs to the winged helix family of transcription factors and it is first expressed at e5.5-e6.5 in the anterior part of the early primitive streak and later in the definitive endoderm (Ang, Wierda et al. 1993; Monaghan, Kaestner et al. 1993; Sasaki and Hogan 1993). Homozygous inactivation of HNF-3 β in mice results in defects in foregut morphogenesis. The embryos die around e11 *in utero* (Ang and Rossant 1994; Weinstein, Ruiz i Altaba et al. 1994) precluding detailed analysis of the role of HNF-3 β in pancreas formation. Recently a specific knockout of Foxa2 in pancreatic β -cells was described. The Foxa2 animals showed hyperinsulemic hypoglycemia revealing that the gene is important for the maintenance of normal

pancreatic histoarchitecture and the preservation of a normal physiological response to glucose. Furthermore it has been suggested for Foxa2 gene an essential role in glucagon-cells development (Sund, Vatamaniuk et al. 2001; Lee, Sund et al. 2005).

1.4.2 Hepatocyte nuclear factor (HNF) 3alpha/Forkhead box (Fox) a1

HNF-3 α , now know as Foxa1, belongs to the same family of transcription factors as HNF-3 α . Its expression in the endoderm starts at e7.5. Targeted disruption of the *HNF-3 α* gene leads to a 70% reduction of glucagon mRNA levels with no reduction in the number of β -cells in the pancreas (Kaestner, Katz et al. 1999). Therefore it has been concluded that the decreased glucagon level is most likely caused by reduced transcriptional activity of the glucagon gene.

1.4.3 Hlxb9

The Hlxb9 gene is a homeobox gene encoding the HB9 protein. The earliest expression of Hlxb9 is detected at e8 in the notochord and gut endoderm (Li, Arber et al. 1999). By e12.5, the expression level decrease but one day later the expression increase again in the dorsal and ventral portion of the pancreas (Harrison, Thaler et al. 1999). Later, by e17.5, the expression of HB9 protein is detectable only in β -cells. Deficiency of Hlxb9 in mouse pancreas causes failure of the dorsal pancreas formation, Pdx1 and IIs1 dorsal expression while the ventral portion forms at the expected time but with a reduced β -cell population (Li, Arber et al. 1999). On the other hand experiments using the Pdx1 promoter to drive expression of Hlxb9 in the pancreas resulted in severely perturbation of pancreatic organization and impaired differentiation. This notion is evidenced by the detection of more elongated, tube-shaped intestinal-like structure in the pancreatic area and decreased insulin-, amylase-, carboxypeptidase A-expressing cells (Li and Edlund

2001). It further confirms the necessity of tightly-controlled temporal expression of Hlxb9 during pancreas development.

1.4.4 Pancreatic duodenal homeobox 1 (Pdx1)/insulin promoter factor 1 (Ipf1)

The Pdx1 (Ipf1) gene belongs to the homeodomain family of proteins and it was the first gene shown to be cell-autonomously required for formation of the pancreas in mice (Jonsson, Carlsson et al. 1994) and humans (Stoffers, Zinkin et al. 1997). Initially it was isolated by different groups as A1-element binding protein of the insulin promoter (Ohlsson, Karlsson et al. 1993) and the TAAT-binding factor of the somatostatin upstream enhancer (Leonard, Peers et al. 1993; Miller, McGehee et al. 1994), only later it emerged as a developmental regulator, in particular when its expression has been deleted. Expression of Pdx1 begins in epithelial cells of the dorsal and ventral pancreatic anlage at the 10 somites stage (e8.5). Lower expression is observed in adjacent duodenal and posterior stomach cells. Pdx1 expression persists in a homogenous way as the pancreatic epithelium branches. The early glucagon-producing cells are not Pdx1 positive. As the secondary transition starts, Pdx1 expression appears more evident. Insulin-positive cells up-regulate Pdx1 expression while differentiating exocrine cells gradually lose the Pdx1 expression, although continuing to express the protein at low level. Ductal-cells also down-regulate Pdx1 and only β -cells and a subset of δ -cells express Pdx1 at levels readily detectable using standard histochemical techniques (Oster, Jensen et al. 1998). Pdx1 has been shown to be important for the transactivation of several pancreatic genes including insulin, glucagon, somatostatin, islet amyloid polypeptide (IAPP) and glucose transporter GLUT-2 (Ohlsson, Karlsson et al. 1993; Miller, McGehee et al. 1994; Waeber, Thompson et al. 1996; Watada, Kajimoto et al. 1996). In Mice Pdx1^{-/-} the pancreas is absent, despite the initial formation of the pancreatic bud (Ahlgren, Jonsson et al. 1996; Offield, Jetton et al. 1996; Johansson, Momma et al. 1999). More recently transgenic animals have been generated to study specifically the absence of Pdx1 in β -cells. Such study has demonstrated the role of Pdx1 in the maintenance of β -cell identity

and glucose sensing machinery (Ahlgren, Jonsson et al. 1998). Finally a selective knockout mouse has been created to study the role of Pdx1 at different times during pancreas development. Such model suggested that Pdx1 expression is essential for the maintenance of β -cell functions, correct pancreatic structure as well as exocrine development in the adult pancreas (Hale, Kagami et al. 2005; Holland, Gonez et al. 2005).

1.4.5 Hepatocyte nuclear factor (HNF) 6

During pancreas development, HNF6 is detected at e9.5 in the epithelial cells and is later restricted to the exocrine and duct cells (Rausa, Samadani et al. 1997). In HNF6^{-/-} embryos, the exocrine pancreas develops normally however the endocrine phenotype is severely inhibited. At birth, fewer endocrine cells were found and the islets of Langerhans are absent. Interestingly, the islets reappear with changed architecture at a later stage (Jacquemin, Durviaux et al. 2000). Therefore, the pancreatic epithelial cells can be committed in the absence of HNF6 but they may fail to give rise to the expected precursor endocrine pool. This notion is supported by the reduced expression of Ngn3, a potential marker of endocrine precursor cells, in HNF6 null mice (Jacquemin, Durviaux et al. 2000).

1.4.6 Neurogenin3 (Ngn3)

Ngn3 is a transcription factor that belongs to the bHLH transcription factor family. Its expression starts at e9.5 (in mice), peaks at e15.5 and diminishes to undetectable levels before the birth (Apelqvist, Li et al. 1999). Experiments using knockout models have shown that those animals lack the endocrine population before birth, on the other end persistent expression of Ngn3 drives the formation of additional

endocrine cells (Apelqvist, Li et al. 1999; Gradwohl, Dierich et al. 2000). For that Ngn3 is thought to be responsible for the pancreatic endocrine commitment. A broader discussion of this fundamental transcription factor will be done in a separate section of this Thesis.

1.4.7 Hairy Enhancer of Split-related protein (Hes1)

Hes1 also encodes a bHLH protein expressed at e9.5 in pancreas. The detection of pancreatic endocrine cells excess in Hes1^{-/-} animals has suggested that Hes1 may have a role in repressing endocrine cell development through the Notch signaling pathway (Jensen, Pedersen et al. 2000). The elevated expression of Ngn3 in the stomach, small and large intestine in Hes1^{-/-} mice confirmed that Hes1 and Ngn3 have opposite effects on the determination of endocrine cell fate (Sumazaki, Shiojiri et al. 2004). Based on the pro-endocrine action of Ngn3, impaired expression of Hes1 would be predicted to increase differentiation of pancreatic endocrine cells, in fact conversion of the developing biliary system to pancreatic tissue in Hes1 deficient mice has been described and further support that idea (Burke, Shen et al. 2004; Sumazaki, Shiojiri et al. 2004). The experiments also claimed out that the conversion from biliary system to pancreatic like cells might be through the re-activation of Ngn3 that is initially restrained by Hes1.

1.4.8 Pancreatic transcription factor 1 (PTF1-p48)

The bHLH transcription factor PTF1-p48 was first described as a DNA-binding subunit of the hetero oligomeric transcription factor PTF-1 that governs the expression of genes in the exocrine pancreas (Krapp, Knofler et al. 1996). Recent results have revised the role of PTF1-p48 to include also the involvement in both exocrine and endocrine cell lineage development. Furthermore a participation in endocrine spatial organization has been shown by several studies in murine and zebrafish models (Krapp, Knofler et al. 1998; Kawaguchi, Cooper et al. 2002; Lin, Biankin et al. 2004).

1.4.9 NeuroD1/ β cell E-Box trans-activator 2 (BETA2)

NeuroD1, also called BETA2 (β -cell E-Box trans-activator-2), is a bHLH transcription factor which expression start at e9.5 and goes throughout the development of β -cells (Naya, Stellrecht et al. 1995) and can be induced by Ngn3 (Huang, Liu et al. 2000). The essential role of NeuroD/BETA2 during pancreatic development has been shown by detection of severely reduced endocrine phenotypes in homozygous mice at birth. NeuroD/BETA2 null mice die after birth because hyperglycemia due to impaired islet formation along with a depletion in β -cell population (Naya, Huang et al. 1997; Chae, Stein et al. 2004). Recently it has been shown that NeuroD/BETA2 activate the insulin promoter in β -cells and represses the somatostatin promoter in human pancreatic δ -cell line (Itkin-Ansari, Marcora et al. 2005).

1.4.10 Paired-box homeobox genes (Pax4/Pax6)

The paired homeobox genes encode key transcription factors involved in the organogenesis of many tissues including eye, brain and endodermal-derived organs such as pancreas and thyroid gland (Mansouri, Goudreau et al. 1999). Among them, Pax4 and Pax6 are associated with pancreatic endocrine cell differentiation (Marsich, Vetere et al. 2003; Smith, Gasa et al. 2003). The expression of Pax4 and Pax6 starts at e9.5 and e9.0, respectively (Turque, Plaza et al. 1994; Sosa-Pineda 2004). Targeted disruption of Pax4 results in the absence of β - and δ -cells in the newborn mice. On the other hand it has been observed, in the same model, an increase in α -cell population (Sosa-Pineda 2004). Insulin and glucagon expressing cells were detected in mutant mice at e10.5, although insulin is no longer detected by e13.5. Thus, Pax4 seems to play a role in maintenance of β -cell differentiation and maturation rather than to initiate it. On the other hand, Pax6 seems to be more crucial for the development of the α -cells (St-Onge, Sosa-Pineda et al. 1997; Sosa-Pineda 2004). Double knockout of Pax4 and Pax6 have also been generated and

shown the complete loss of endocrine cells types (St-Onge, Sosa-Pineda et al. 1997), so Pax4 and Pax6 probably act cooperatively to regulate islet cell identities.

1.4.11 NK class homeodomain protein Nkx2.2 and Nkx6.1

Nkx2.2 and 6.1 belong to the NK class of homeodomain proteins. The expression of the two proteins starts respectively at e9.5 and e10.5 (in mice). For both the proteins the early expression is broadly spread but then became more restricted to defined cell types, in particular during the development the expression of Nkx2.2 become restricted to α -, β - and PP cells while the Nkx6.1 expression become restricted to insulin producing cells in the adult pancreas (Oster, Jensen et al. 1998; Sussel, Kalamaras et al. 1998; Wang, Elghazi et al. 2004). More recently, Prado et al. (Prado, Pugh-Bernard et al. 2004) found that β -cells were replaced by ϵ -cells (ghrelin-producing cells) in both Nkx2.2^{-/-} and Pax4^{-/-} animal models (Prado, Pugh-Bernard et al. 2004). Nkx6.1 null animals display an important decrease in the β -cell population after e12.5 (Sander, Sussel et al. 2000), therefore Nkx6.1 seems to be required for the expansion and final differentiation of β -cell progenitors. The phenotype of the Nkx2.2 and double (2.1 and 6.1) knockout animals is very similar (Sander, Sussel et al. 2000). Since Nkx2.2 is still present in the Nkx6.1^{-/-} model this indicates a possible common pathway with Nkx2.2 expressed upstream of Nkx6.1.

1.4.12 Islet 1 (Isl1)

Isl1 encodes a transcription factor of the LIM homeodomain family and is involved in the early formation of the pancreas (Ahlgren, Pfaff et al. 1997). The Isl1 expression starts at e9.0 in the dorsal pancreatic epithelium and in the mesenchyme surrounding the dorsal, but not the ventral, evagination of the gut endoderm. Isl1 is particular in that regard because is expressed in fully differentiated cells. In fact, at e9.5 all the glucagon positive cells express Isl1 then later it is detectable in all the islets cells. Mice mutant for Isl1 do not develop after e9.0. The dorsal pancreatic mesenchyme is not formed, endocrine development is abolished and the exocrine component of the dorsal

pancreas is strongly depleted. The importance of the role of Isl1 has been understood by experiments of co-culture of tissue from Isl1 mutant mice with mesenchyme from wild type embryos. In these conditions the endocrine development can be partially restored but not the exocrine one. Furthermore Isl1 has been shown to bind the promoter region of several genes such as glucagon, insulin and somatostatin, suggesting that the Isl1 is also necessary for the maintenance of the islet function in the adult (Brink 2003).

1.4.12 Insulinoma-associated factor 1 (Insma/IA1)

The insulinoma-associated 1 (Insm1/IA1) gene encodes for a zinc-finger protein well conserved during the evolution and was first identified in a subtraction library from an insulinoma (Goto, De Silva et al. 1992). Subsequently it was found to be expressed in a large number of insulinoma, neuroendocrine tumors as well as during nervous system and pancreas development (Goto, De Silva et al. 1992; Zhu, Breslin et al. 2002; Breslin, Zhu et al. 2003). To IA1 has been attributed, for long time, the role of transcription repressor; however recent studies have showed that it can act also as activator of the transcription during pancreas development (Gierl, Karoulias et al. 2006). Activated by Ngn3 (Mellitzer, Bonne et al. 2006) IA1 is expressed in all the cell types of developing and adult pancreatic islet (Gierl, Karoulias et al. 2006), nevertheless its role during pancreas development has not been assessed yet. Generic analysis performed using IA1 mutant mice model shown that the protein is essential for the correct differentiation of the β -cell population. IA1 mutants showed that most of the β -cell precursor are maintained but do not express any hormones. The phenotype of IA1 mutant mice differs from that of Nkx2.2, Nkx6.1 and Pax4 mutant mice (Sander, Sussel et al. 2000; Prado, Pugh-Bernard et al. 2004). Moreover immunohistological and microarray analysis of those animals reveal a down-regulation of Nkx6.1 and Pdx1 in the pancreas, indicating that IA1 affects the transcriptional network that determines β -cell differentiation.

1.4.13 Brn4 Brain 4

Brn4 (Brain 4/Pou3f4) belongs to a POU-homeodomain transcription factor family and directly regulate the glucagon promoter (Hussain, Lee et al. 1997). Its expression starts at e10 in the pancreas just before Pax6, which also appears in the glucagon expressing cells. It was suggested that Brn4 was the first and only gene specifically expressed in glucagon-expressing cells in the pancreas, nevertheless a Brn4 null mouse showed that its expression is not essential for glucagon cell fate determination since this mouse had normal endocrine formation (Heller, Stoffers et al. 2004).

1.5 Endocrine fate commitment

1.5.1 Ngn3 the master-switch of β -cell development

Neurogenin 3 (Ngn3, atoh5) belongs to the bHLH transcription factor family and has the role of proendocrine factor in the development of the pancreas (Gradwohl, Dierich et al. 2000). Ngn3 is expressed in endocrine precursors of both the pancreas and gastrointestinal tract (Jenny, Uhl et al. 2002). Mice deficient for Ngn3 fail to develop any endocrine cells and die 1-3 days after birth because diabetes, whereas exocrine cells develop normally (Gradwohl, Dierich et al. 2000).

Ngn3 has been demonstrated to be necessary and sufficient to drive the formation of islet cells during pancreatic development (Apelqvist, Li et al. 1999). Also using *in ovo* electroporation of early chick endoderm, Ngn3 was shown to induce endocrine differentiation in entire gut epithelium (Grapin-Botton, Majithia et al. 2001).

Ngn3 expressing cells in the pancreas are transient endocrine progenitor cells committed to an endocrine fate. Lineage tracing studies have shown that Ngn3 expressing cells do not give rise to exocrine cells and very rarely to duct cells but give rise to all

endocrine cell types (Gu, Dubauskaite et al. 2002). Ngn3 protein can be detected in very few cells in the early pancreas (e10.5-e12.5), while from e13.5 and onward, a major peak (between e14.5-e15.5) of expression is observed (Apelqvist, Li et al. 1999; Jensen, Heller et al. 2000). Ngn3 expression is only observed as a transient expression in the pancreas and is completely turned off in postnatal mice, as shown by immunohistochemistry (Jensen, Heller et al. 2000).

However, Gu et al. (Gu, Dubauskaite et al. 2002) has detected Ngn3 expression by RT-PCR in handpicked islets from 8-week old mice. In addition, using double transgenic Ngn3-Cre.ERTM;Z/AP mice they observed, upon administration of Tamoxifen (TM), induction, in three or eight-weeks old mice, of Ngn3 expression in 15 out of 25 pancreata (identified by Human Placental Alkaline Phosphatase HPAP staining cells) in islet cells many of which co-express insulin (Gu, Dubauskaite et al. 2002). These data could suggest that Ngn3-expressing cells act as islet progenitor cells not only during embryogenesis but also in adults. This Ngn3 expression is transient and the new Ngn3-positive cells are formed from Ngn3-negative cells during development as showed by the fact that positive staining for Ngn3 disappeared in HPAP+ cells (Gu, Dubauskaite et al. 2002).

In conclusion the majority of Ngn3+ cells are endocrine progenitors that differentiate into endocrine cells without self-renewing (Gu, Dubauskaite et al. 2002). This, suggested also by Jensen et al., shows that once cells enter the endocrine pathway and start to express Ngn3, their proliferation rate decrease (Jensen, Heller et al. 2000). Premature expression of Ngn3 in pancreas (Apelqvist, Li et al. 1999) and endoderm (Grapin-Botton, Majithia et al. 2001) gives rise to endocrine cells consisting mostly of α -cells and it was suggested that α -cells are the default endocrine cell type to be formed. However, an *in vitro* system in which Ngn3-expressing adenovirus was used to infect a mouse pancreatic duct cell line showed that Ngn3 turns on the islet differentiation program in these duct cells (Gasa, Mrejen et al. 2004). Both the genes that act early in differentiation, and those necessary for fully differentiated islet cells were switched on. All islet hormones, with the exception of glucagon, are turned on, although in much lower level than regular islets or β TC3 cell line. This could be explained because the duct

cell line lack positive regulators that allows α -cell differentiation. In fact, Brn4 has not been induced, in that system, by the expression of Ngn3.

An alternative explanation of such results could be the competency of the cells; in particular the duct cells could be not able to respond to Ngn3 as a pancreatic progenitor cell does. Insulin gene expression has been also achieved in a similar study by Vetere et al. (Vetere, Marsich et al. 2003) where pluripotent cells were transfected with an Ngn3-expressing vector (Vetere, Marsich et al. 2003). Interestingly, only after induction of endoderm differentiation the overexpression of Ngn3 has been able to induce insulin gene expression along with many other endocrine transcription factors (Vetere, Marsich et al. 2003).

It is now accepted that Ngn3-positive cells could adapt their responses to specific microenvironmental inductive signals that changes during pancreatic development, and this give rise to the fact that the different endocrine cell types adopt their fate at different time points. This was partly confirmed using pancreatic explants in which forced expression of Ngn3 around e9.5 favors α -cells, whereas at e 12-e13 it favors β -cells (Dominguez-Bendala, Klein et al. 2005). This increase of Ngn3 expression level was achieved by transducing pancreatic explants with a TAT/Ngn3 fusion protein. This TAT/Ngn3 fusion protein was taken by cells by endocytosis and remains in the cells for a short period before to be degraded. Interestingly this short burst of Ngn3 activity, mimicking in some way normal behavior of Ngn3 expression, has been demonstrated to be essential to activate the transcription factors cascade downstream Ngn3 protein. Furthermore it has also been suggested that several other TF are required to deviate from the adult α -cell fate. The Ngn3 gene expression is turned off before terminal differentiation and in the normal pancreas, Ngn3 protein has never been observed co-expressed with mature markers such as insulin. Ngn3 positive cells exit the cell cycle, turn off Ngn3 expression and differentiate into mature endocrine cells. Interestingly, in a recent study by Lee et al., it has been demonstrated that pancreatic transcription factors that are seen to be expressed during the development are re-expressed also during the regeneration, with the only know exception of Ngn3 (Lee, De Leon et al. 2006).

Though Ngn3 protein, which marks all the proendocrine cells during development, is not seen during mouse pancreas regeneration, functional neo-islets are

generated by 4 weeks after 70% pancreatectomy (Lee, De Leon et al. 2006). The explanation for such phenomenon has been given only recently by Joglekar et al. (Joglekar, Parekh et al. 2007). In particular they showed that, after pancreatectomy only transcription factors upstream Ngn3 are activated and expressed at protein level with a pattern of expression similar to that observed during pancreas development whereas both Ngn3 and its direct downstream target (NeuroD/BETA2 and Nkx2.2) were never expressed even though the transcript for Ngn3 could be detected. The hypothesis that comes out from this study is a post-transcriptional regulation of Ngn3 protein by microRNA (Joglekar, Parekh et al. 2007). They carried out microRNA analysis during different stages of pancreatic development and regeneration and identified that, microRNAs that can potentially bind to Ngn3 transcript, are expressed at least 200-fold higher in the regenerating mouse pancreas. Inhibition of these miRNAs in regenerating pancreatic cells using anti-sense miRNA-specific inhibitors induced expression of Ngn3 and of its downstream players: NeuroD/BETA2 and nkx2.2.

Furthermore Johansson et al. (Johansson, Dursun et al. 2007) have generated an inducible form of Ngn3 fusing the mutated form of the estrogen receptor (ERTM) that can be activated by Tamoxifen (Hayashi and McMahon 2002) under the transcriptional control of the mouse pdx1 promoter in order to specifically express this inducible protein in the pancreas. An addition of IRES-nGFP sequence allowed direct tracing of cells expressing the pdx1-ngn3-ERTM-IRES-nGFP transgene (Johansson, Dursun et al. 2007). Transgenic mice expressing the Tamoxifen-inducible Ngn3 fusion protein have been so generated. This work shows how the time of Ngn3 activation is important to differentiate cells into the different endocrine cell types. Early activation almost exclusively induced glucagon⁺ cells, while depleting the pool of pancreas progenitor cells. A activation at e11.5, Pdx1⁺ cells became competent to differentiate into insulin and PP cells. Somatostatin cells were generated from e14.5 while the competence to generate glucagon⁺ cells was dramatically decreased.

Altogether these data show how the timing for Ngn3 expression is important for the development and differentiation of the endocrine components of the pancreas.

1.5.2 Notch signaling pathway in endocrine differentiation

Notch signaling has been shown to play two major functions in the pancreas development: an early role in maintaining progenitor cells into an undifferentiated state and a later role in controlling the lateral inhibition, allowing selection of endocrine cells from non-endocrine cell population. As for the early role it seems that Notch signaling controls directly and indirectly Ngn3 expression. In fact it is known that Ngn3 expression appears to be dependent from positional signals that direct it to the gut and to the pancreatic endoderm and local signals that limit it to specific cells (Fig. 5).

As for the lateral inhibition the mechanism is more straightforward and it has been recognized to be an almost “general” regulator for cell-fate decision in a variety of organisms and tissues (Kageyama and Ohtsuka 1999; Appel, Givan et al. 2001). Lateral inhibition, through the notch signaling pathway, prevents Ngn3 expression in all but a few scattered cells within the pancreatic epithelium (Schroeter, Kisslinger et al. 1998; Apelqvist, Li et al. 1999; Gradwohl, Dierich et al. 2000; Jensen, Pedersen et al. 2000; Schwitzgebel, Scheel et al. 2000). As expected, FGF10 deficient mouse embryos show a hypoplastic pancreas attributed to loss of the proliferative phase of the pancreas epithelial progenitor cells (Bhushan, Itoh et al. 2001).

In the pancreas, all components of the Notch pathway have been shown to be expressed (Jensen, Pedersen et al. 2000; Lammert, Brown et al. 2000). The Notch receptors Notch1 and 2 are expressed in pancreatic epithelial precursors, Notch3 in the surrounding mesenchyme, and Notch4 in the endothelial cells. The Notch ligands Jagged1, 2 and Delta1 (Dll1) are also expressed in pancreas. In addition the Notch in concert with the DNA-binding protein RBP-J κ activate the expression of the target gene Hes1 in the pancreatic bud and later in the pancreatic epithelial precursor cells (Apelqvist, Li et al. 1999).

Interestingly, Hes1, a dominant repressor, dramatically and specifically inhibits the Ngn3 promoter (Lee, Smith et al. 2001). This is achieved when Ngn3 expression in a committed cell activates the Delta1 notch ligand, which in turn activates the receptor Notch1 in the surrounding cells by an intercellular cleavage of the intracellular domain of the Notch receptor. This domain interacts with RBP-J κ activating Hes1 expression and

subsequently downregulates the expression of Ngn3. Hes1 is co-expressed in exocrine precursors but not in amylase expressing acinar cells, along with Ptf1a (Esni, Ghosh et al. 2004). Overexpression of Hes1 repress acinar cell differentiation, but not Ptf1a expression, in a cell-autonomous manner in explant cultures of e10.5 mouse dorsal pancreatic buds. Ectopic Notch activation also delays acinar cell differentiation in developing zebrafish pancreas.

Further evidence of a role for endogenous Notch in regulating exocrine pancreatic differentiation was provided by examination of mutated zebrafish embryos in which Notch signaling is disrupted. Notch-deficient embryos display accelerated differentiation of exocrine pancreas relative to wild-type controls.

Using transient transfection assays involving a Ptf1-responsive reporter gene, it has been showed that Notch target genes directly inhibit Ptf1 activity, independent of changes in expression of Ptf1 component proteins. These results define a normal inhibitory role for Notch in the regulation of exocrine pancreatic differentiation (Apelqvist, Li et al. 1999). This has been confirmed also by the observation that a forced expression of constitutively active Notch1 (Notch1-1C) using the *pdx1* promoter represses both pancreatic endocrine and exocrine differentiation (Hald, Hjorth et al. 2003; Esni, Ghosh et al. 2004).

The pancreatic epithelium appears disorganized with reduced numbers of endocrine cells. Subsequent branching morphogenesis is impaired and the pancreatic epithelium forms structures with ductal phenotype containing a few endocrine cells but completely devoid of acinar cells; on the other hand defects in Notch signaling has been shown to accelerate the endocrine differentiation program leading to a depletion of the precursor cell population (Apelqvist, Li et al. 1999; Jensen, Pedersen et al. 2000; Hald, Hjorth et al. 2003; Hart, Papadopoulou et al. 2003). Mice *Hes1*^{-/-} have showed ectopic expression of Ngn3 in the biliary epithelium and subsequent differentiation into pancreatic-like tissue (Sumazaki, Shiojiri et al. 2004). Interestingly, *Hnf6* null mouse, show a highly reduced expression of Ngn3 and a consequent depletion of endocrine cell population. Nevertheless the level of Notch1, 2, 3 and Hes1, are normal. This suggests that the expression of Ngn3 could be regulated not only by Notch pathway.

The maintenance of Notch activation was proposed to derive from the pancreatic mesenchyme and mediated by FGF10 (Gittes, Galante et al. 1996; Norgaard, Jensen et al. 2003; Miralles, Lamotte et al. 2006). Persistent expression of FGF10 in the pancreatic epithelium resulted in hyperplasia and impaired pancreatic differentiation of pancreatic progenitors caused by a maintained Notch activation. In fact, the FGF10-positive pancreatic cells expressed Notch1 and Notch2, the Notch-ligand genes Jagged1 and Jagged2, as well as the Notch target gene Hes1 resulting in impaired expression of Ngn3 (Hart, Papadopoulou et al. 2003; Norgaard, Jensen et al. 2003). While FGF10 is expressed in the mesenchyme, its receptor FGFR2 is expressed in the pancreatic epithelium during the embryonic life (Orr-Urtreger, Bedford et al. 1993; Scharfmann 2000). This is consistent with the hypothesis that FGF10 is the mesenchymal factor responsible for the maintenance of the Notch signaling in the self-renewal of progenitor cells (Miralles, Lamotte et al. 2006).

In conclusion the Notch signaling pathway can be considered as a “suppressive maintenance”; the Notch target gene Hes1 activation leads to “suppression” of differentiation and “maintenance” of the precursor population. This produces ubiquitous expression of the Notch system in contrast to the scattered expression in the lateral inhibition model.

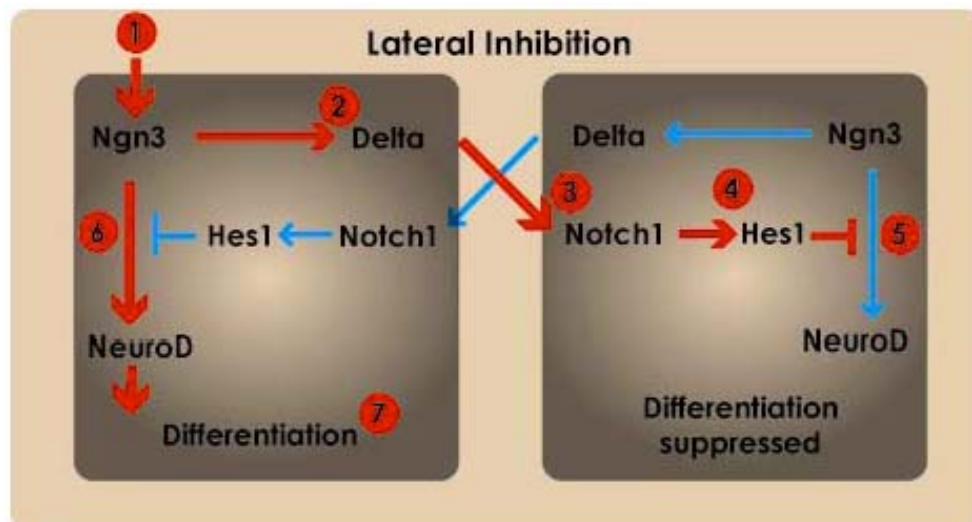


Fig. 5: Lateral inhibition is caused by cell-cell signaling between epithelial neighbors. Cells in which Notch1 is activated by the ligand Deltas express high levels of Hes1, which in turn repress the proendocrine gene Ngn3. However, in ligand expressing cells, hes1 is not regulated, thus allowing robust Ngn3 expression and differentiation toward the endocrine lineage. Jensen J., 2004

1.6 Glucose homeostasis:

Normal glucose homeostasis in the human body represents a balance between glucose appearance and tissue glucose uptake and utilization. This balance is tightly regulated. Therefore, plasma glucose concentrations are maintained within a range of 60–110 mg/dl for the normal fasting state and <140 mg/dl for 2-hour post-glucose-load.

Three physiologic processes need to interact in a highly coordinated way to maintain glucose homeostasis: β -cell insulin secretion, tissue glucose uptake, and hepatic glucose production. In the fasting state, hepatic glucose production via glycogenolysis and gluconeogenesis is almost exclusively responsible for the plasma glucose levels.

During fasting, tissue glucose uptake is predominantly insulin-independent and occurs mostly in those tissues that require glucose such as the central nervous system. In order to maintain fasting glucose concentrations constant, the rate of tissue glucose uptake dictates the rate of hepatic glucose production.

Following a meal (absorptive state), plasma glucose levels rise and promote hepatic glucose uptake. Elevated glucose levels stimulate the release and production of insulin by the β -cell, which acts as a glucose sensor, then the hepatic glucose production is suppressed by increased plasma insulin concentrations. Peripheral glucose uptake, mostly in the muscle, is also stimulated by insulin through GLUT4, a member of the GLUT family of glucose transporters (Foster and Klip 2000). In this way, hyperglycemia is minimized and the return of mealtime glycemic levels to pre-meal values is ensured (Porte and Schwartz 1996).

Once inside the cell, the glucose is converted into glycogen in muscle or converted to lipids in adipose tissue (lipogenesis) (Nonogaki 2000). Free fatty acid (FFA) metabolism also plays an important role in maintaining glucose homeostasis in the fasting and absorptive state. Elevated rates of fat breakdown (lipolysis) lead to an increased release of FFAs. These have a detrimental action on the uptake of insulin by the liver, which in turn results in increased hepatic gluconeogenesis. In addition, elevations of FFA inhibit glucose uptake and utilization in insulin-sensitive tissues. Increased plasma insulin levels after a meal inhibit lipolysis and FFA release from adipose tissue.

1.7 Diabetes

The diabetes mellitus is one of the most common endocrine disorders today, with 180 million people affected in the world in 2003 to approximately 220 million people in 2010 (Zimmet 2000; Zimmet, Alberti et al. 2001).

The diabetes is a group of disorders of multiple etiologies characterized by chronic hyperglycemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action or both (WHO 1999). The two most common types of diabetes mellitus are type 1 and type 2. About the 90% of the patients with diabetes suffer from type 2 diabetes, which is characterized by insulin resistance and impaired insulin release.

Type 2 diabetes used to be called “adult onset diabetes”, but today children are also being diagnosed. Type 2 diabetic patients have chronically increased levels of blood glucose, which is toxic and further impairs insulin sensitivity and β -cell function, leading to more severe type 2 diabetes. Chronic hyperglycemia is a major risk factor for the development of complications such as micro and macroangiopathy and neuropathy, which result in blindness, impotence, ischemic foot ulcers and amputations (Porte and Schwartz 1996). One of the possible molecular mechanisms responsible for these complications is the increased oxidative stress induced by chronic hyperglycemia (King and Loeken 2004). Briefly, persistent hyperglycemia drives insulin-independent tissues to assume glucose. The increased glucose inside of the cells will enter in the glycolysis pathway, Krebs cycle and oxidative pathways (King and Loeken 2004). The increased flux will generate an excess of oxygen reactive species (ROS) and the accumulated ROS reduces the increased oxidative stress and may interfere with normal tissue functions. For example the increased levels of ROS affect: blood flow to retina (Kowluru and Kennedy 2001) (retinopathy) and contractility of vascular smooth muscle cells (Sharpe, Liu et al. 1998) (vascular disorders). ROS decrease neural conductivity in peripheral nerves (Hounsom, Corder et al. 2001) (neuropathy).

Another major factor in the development of the diabetic complications could be the increased glycosylation of proteins. This effect can lead to the accumulation of advanced glycosylation end-products and lipoxidation end-products (AGEs/ALEs) (Wada and

Yagihashi 2005). AGEs/ALEs can induce cross-linking of proteins or induce specific cellular responses including the release of profibrogenic and proinflammatory cytokines thus causing diabetic retinopathy and nephropathies such as glomerulosclerosis, interstitial fibrosis, and tubula atrophy (Bohlender, Franke et al. 2005). In addition to ROS and AGEs/ALEs, VEGF in diabetic retinopathy (Aiello 2005) and microalbuminuria for cardiovascular and renal disorders in type 2 diabetes has also been shown (Lane 2004).

Patients with type 2 diabetes can often be treated with diet and exercise alone or in combination with insulin sensitizers, which act to reduce the insulin resistance in target tissues, or sulphonylureas which stimulate the β -cells to secrete more insulin.

Type 1 diabetes is caused by an autoimmune reaction, which leads to destruction of β -cells and subsequent insulin deficiency. These patients are treated with insulin injections to compensate for the absence of natural insulin production. Whole organ pancreas transplants have been performed in type 1 diabetic patients since the 1970s.

Now with the recent success of islet transplantation using the Edmonton protocol (Shapiro, Lakey et al. 2000), type 1 diabetic patients have been shown to maintain normal blood glucose levels without daily insulin injections. This procedure/cure unfortunately relies on islets available from pancreata from cadavers, a very limited supply. Each type 1 diabetic patient needs islets isolated from 2-3 pancreata and could need several transplants before achieving insulin independence.

These results are promising, but a shortage of donors has lead to an explosion in the number of investigations of methods to generate islets surrogates suitable for transplantation. Some of the roads taken so far have been expansion of existing β -cell lines, encapsulated islet xenografts, conversion of a stem cell population (either pancreatic or non pancreatic adult stem cells) to β -cell, in vitro differentiation of embryonic stem cells to β -cells (Soria 2001) or transdifferentiation from a different cell type (Nygaard Jensen and Jensen 2004).

1.7.1 Type 1 Diabetes Mellitus

Type 1 diabetes mellitus (T1DM) results from an absolute deficiency of insulin due to autoimmune attack and destruction of the β -cells. Type 1 diabetes mellitus usually appears suddenly in non-obese children or young adults. This is why it is referred to as juvenile-onset diabetes. This occurs as a result of direct attack of an aberrant b-specific autoimmune antibodies (Narendran, Estella et al. 2005) with involvement of cytokines and T-lymphocytes specific for β -cell antigens (Nerup, Mandrup-Poulsen et al. 1988), which only became symptomatic when the β -cell mass is reduced by approximately 90%.

T1DM is a complex genetic trait, with multiple genetic loci contributing to susceptibility, but also with environmental factors playing a role in determining a risk. It has been shown to exist a significant familial clustering of T1DM with an average risk in siblings of 6% compared to 0.4% in the general population (Karvonen, Tuomilehto et al. 1993; Karvonen, Viik-Kajander et al. 2000). At least one locus that contributes strongly to this familial clustering resides within the major histocompatibility (MHC) region on the chromosome 6p21.3. Other established T1DM risk genes include the insulin gene region on chromosome 11p15 and the cytotoxic T-lymphocyte associated 4 gene (CTLA4) region on the chromosome 2q33 (Pociot and McDermott 2002). Most attention to environmental factors has been paid to virus infections and early exposure to cow's milk protein as triggers of autoimmunity in individuals genetically susceptible to T1DM (Dahlquist 1998; Jun and Yoon 2003) although no consensus has been reached for this hypothesis.

Another epigenetic factor is the influence of the intrauterine environment on the β -cell development and later development of diabetes. Low protein diet given to pregnant rats during gestation induces permanent changes in the offsprings (e.g. altered islet proliferation, islet size, pancreatic insulin, apoptotic rates and increased sensitivity to interleukin 1b (IL-1b) and nitric oxide (NO) (Snoeck, Remacle et al. 1990; Petrik, Reusens et al. 1999; Merezak, Hardikar et al. 2001). In the offspring of T1DM mothers, it has been suggested that the exposure to a diabetic environment in utero is associated with

increased prevalence of IgT and defective insulin secretory response in the adulthood (Sobngwi, Boudou et al. 2003).

The present concept of the pathogenesis of T1DM is mainly based on evidence from studies on the animal models of human T1DM. The diabetes-prone BioBreeding (BB-DP) rat spontaneously develops a diabetic syndrome with characteristics similar to the human T1DM. In the non-obese diabetic mice model (NOD) autoimmune diabetes can be transferred from a diabetic to a non-diabetic animal via T lymphocytes. This confirms that the T lymphocytes are the mediators of autoimmune diabetes in NOD mice. It seems that both glutamate decarboxylase-(GAD) and insulin-reactive T cells are diabetogenic in the NOD mice. Antibodies to the above antigens can be observed long before the manifestation of hyperglycemia. The major histocompatibility complex class II allele, HLA-DQ, is the major genetic element, conferring susceptibility to or protection from human T1DM.

However, the mechanisms determining the contribution of the HLA-DQ allele remain to be elucidated (Moustakas and Papadopoulos 2002). Accumulating evidence suggests that the regulation of the gut immune system may be aberrant in T1DM, possibly due to environmental factors (Akerblom, Vaarala et al. 2002). Many autoimmune diseases are characterized by an overproduction of cytokines. In the case of T1DM, interferon- γ (IFN γ), tumor necrosis factor- α (TNF α) and interleukin 1- β (IL-1 β) have been shown to contribute to the pathology. In addition, the quality of the immune reactivity is an important factor in beta-cell destruction. In the animal model, only infiltration of the islets by helper T cells of the subtype Th1, which secrete high levels of IFN γ , caused beta-cell damage (Healey, Ozegbe et al. 1995).

According to the so-called “Copenhagen model” of T1DM, beta-cells are destroyed by cytokine-induced free radical formation before cytotoxic helper T cells and/or autoantibody-mediated cytolysis (Nerup, Mandrup-Poulsen et al. 1988; Freiesleben De Blasio, Bak et al. 1999). A consequence of this model is that proinflammatory cytokines induce both protective and deleterious mechanisms in all cells expressing cytokine receptors. In genetically predisposed individuals, T1DM development occurs because the deleterious events in the β -cells prevail.

1.7.2 Type 2 Diabetes Mellitus

The glucose homeostasis deregulation in type 2 diabetes mellitus (T2DM) is mainly due to a combination of two physiological defects: impaired insulin secretion and reduced peripheral and hepatic insulin sensitivity (Beck-Nielsen and Groop 1994; Bajaj and DeFronzo 2003).

The development of T2DM is a gradual process, and the relative roles of the two defects in the etiology of the disease remain controversial. However, it is now generally agreed that overt diabetes arises when β -cell s can no longer compensate for the insulin resistance, that because a prolonged exposure of pancreatic β -cells to glucose. This eventually leads to β -cells exhaustion and glucose toxicity. The effect of β -cell exhaustion is the early form of the disease and could be reversible whereas the glucose toxicity occurs later in the disease and is less reversible (Rossetti, Giaccari et al. 1990).

Recently, a detrimental effect of elevated levels of free fatty acids on β -cell s function has also been studied. The suggestion is that diabetes is not only the carbohydrate disease but also associated with disordered lipid metabolism (Unger 1995).

In addition, the deposits of islet amyloid present in T2DM diabetic pancreas may lower the functional β -cell mass (Kahn, Andrikopoulos et al. 1999; Marzban, Park et al. 2003). Several lines of evidence indicate that hyperglycemia induces β -cell apoptosis by inducing apoptotic molecules including glucose-induced NF- κ B, apoptotic-associated protein caspase 3 and ROS. Once the apoptosis-related factors have accumulated, β -cell mass is reduced, insulin secretion is impaired and eventually diabetes occurs (Mandrup-Poulsen 2003).

Insulin resistance and β -cell dysfunction are two major clinical characteristics found in T2DM diabetic subjects. As will be described after, insulin acts to maintain euglycaemia and this process involves several steps. A significant disruption of any of these actions regardless of the presence of normal or even high circulating levels of insulin indicates the presence of “insulin resistant” phenotype (Reaven 1988).

Insulin resistance is the first defect to be detected in T2DM and must be accompanied by β -cell deficiency for full development of the disease (Martin, Warram et al. 1992). The early insulin resistance in insulin-targeted organs including liver, muscle

and adipose tissue starts to be detected during pre-diabetic stage and therefore elevated insulin is subsequently produced from β -cell s to compensate the require of insulin to maintain glucose homeostasis. Over a period of time, however, overexpression of insulin does not match the requirement of insulin to uphold euglycemia and the next pathological condition, β -cell failure, occurs. Furthermore the expansion of β -cell mass (called islet hyperplasia), is another notable consistent feature associated with insulin resistance in T2DM animals (Tokuyama, Sturis et al. 1995).

The working model for the β -cell dysfunction in T2DM can be discussed using animal model, 90% pancreatectomy (Px) rats. It has been found that insulin secretion is enhanced (hypersensitivity) during the period when the blood glucose is as stable as the level in control animals (within 2 weeks post surgery) in Px rats (Porte 2001). This can be explained by the idea that, in order to compensate for the demands of insulin, β -cells overwork to produce sufficient insulin to reach normoglycaemia.

Another effect that accompanies the hypersecretion of insulin from β -cell s in Px rats is the increase of glucokinase activity, probably to increase glucose sensitivity. Additionally an increase in the low-Km hexokinase activity is also observed in 90% of diabetic rats in comparison with the non-diabetic control animals (Narendran, Estella et al. 2005). Higher activity of hexokinase results in increased sensitivity of β -cell s even when the blood glucose concentration is low. Thus established hyperglycaemia and enhanced β -cell hexokinase activity lead to very high level of insulin output (Fajans, Bell et al. 2001). Since the insulin targeted tissues are insulin-resistant, the hyperglycaemia is more serious in spite of constitutive insulin secretion. Finally overworked β -cells become fatigued and insulin granules are depleted.

The occurrence of dysfunctional insulin receptor (IR) and downstream kinase provided another mechanism to explain why insulin resistance occurs. In the peripheral organ (e.g. muscle and liver) of insulin resistant subjects, a defect or IR and insulin receptor kinase have been observed (Froguel, Zouali et al. 1993; Yamagata, Furuta et al. 1996). Defects in IR function were originally determined by the detection of IR-specific autoimmune antibody, the mutation on IR and the decreased IR numbers (Horikawa, Iwasaki et al. 1997; Stoffers, Ferrer et al. 1997). However, this defect is unlikely to be significant contributing to the pathology of diabetes while the antibody titre is generally

low and large changes of insulin-IR binding have never been seen (Stoffers, Ferrer et al. 1997). More attempts therefore were turned to investigate the effect of deficient IR kinase for the onset of insulin resistance. Insulin-IR binding activates insulin receptor substrate-1 (IRS-1)-dependent phosphatidylinositol 3-kinase (PI3K) which activates atypical protein kinase C (aPKC or PKC λ or ζ) and protein kinase B (PKB/Akt) pathways and thereby regulates glucose transport (Malecki, Jhala et al. 1999).

The results exhibiting the reduced Akt kinase activity in the skeletal muscle tissue of T2DM patients (Busch and Hegele 2001) and defective activation of aPKC in muscles and adipocytes of rats with T2DM (Guz, Montminy et al. 1995) suggested that IR-related kinase plays a major role for the development of insulin resistance.

1.7.3 Maturity Onset Diabetes of the Young:

MODY is a specialized type of T2DM since it is characteristic of an onset before the age of 25 years and usually in childhood and adolescence (Fajans, Bell et al. 2001). MODY can result from the mutation in any one of at least six genes.

One of them, MODY 2, encodes the glycolytic enzyme glucokinase (Froguel, Zouali et al. 1993). Whereas the other five types of MODY are associated with pancreatic transcription factors: HNF4 α (related MODY1)(Yamagata, Furuta et al. 1996), HNF1 α (associated with MODY3) (Stoffers, Ferrer et al. 1997), HNF-1 β (associated with MODY5) (Horikawa, Iwasaki et al. 1997), and NeuroD/BETA2 (associated with MODY6) (Malecki, Jhala et al. 1999). These genes are all expressed in pancreatic β -cells and the disruption of each gene lead to β -cell dysfunction, and eventually diabetes. In comparison to late-onset T2DM, the onset of MODY is more likely to be a genetic issue.

This statement is supported by the fact that the screening of family pedigrees of MODY diabetic patients showing a history of diabetes in three of four generations (Busch and Hegele 2001; Fajans, Bell et al. 2001).

1.8 Glucose response in islet cells

During fasting carbohydrates are processed by digestive enzymes and absorbed through the intestinal epithelium. Insulin is then released, as said before, when the glucose level become elevated and the β -cells are “turned on” to produce the hormone.

Glucose sensing is the first step leading up to the release of insulin from the β -cells. In this step, both the high-K_m glucose transporter GLUT2 and the high-K_m glucose phosphorylating enzyme glucokinase (GCK) are important (Efrat, Tal et al. 1994).

In order to respond to an elevated glucose concentration, glucose must enter into β -cells. Glucose entry occurs by facilitated diffusion via GLUT2. GLUT2 was originally cloned from a rat liver cDNA library and is expressed specifically in liver and in pancreatic islets (Unger 1991). A parallel correlation has been shown between reduced GLUT2 levels and loss of glucose-induced insulin secretion in Zucker diabetic mice (Johnson, Ogawa et al. 1990). In addition, hyperglycaemia has been observed in transgenic mice over-expressing GLUT2 antisense RNA. This reveals that GLUT2 might be not only responsible for the transportation of glucose into cells but also essential for maintaining the glucose homeostasis.

However, different observations were presented in other animal models. For example, in RIP-Hras transgenic mice (Hras protein is specifically expressed in insulin producing cells), there is a reduction of GLUT2 but this change did not affect glucose homeostasis (Tal, Wu et al. 1992). Furthermore, in human another member of the glucose transporter family, GLUT1, is more abundantly expressed than GLUT2 in pancreatic β -cells (De Vos, Heimberg et al. 1995).

These results suggested the function of GLUT2 may be primarily for the transport of glucose into the cells (Schuit 1997). GCK is a member of the hexokinase family of enzymes and expressed in pancreatic β -cells and hepatocytes. GCK catalyses the ATP-dependent phosphorylation of glucose after glucose entry into the β -cells (Weir, Sharma et al. 1997) generating glucose-6-phosphate (G-6-P) (Tirone and Brunicardi 2001). It functions as glucose sensor in β -cells controlling the rate of entry of glucose into glycolysis and the tricarboxylic acid (TCA) cycle (Matschinsky, Glaser et al. 1998). A

decrease in β -cell glucokinase activity has been observed in several situations associated with decrease insulin release (Meglasson and Matschinsky 1986).

It was also reported that the impaired GCK activity is related to one type of autosomal dominant mode inheritance mature-onset diabetes of the young, MODY2 (Froguel, Zouali et al. 1993). This further confirms the importance of GCK for insulin secretion and glucose homeostasis.

1.9 Insulin secretion:

The metabolism of glucose through the glycolytic pathway produce two molecules of ATP (MacDonald, Fahien et al. 2005). As a result of the increase of ATP production, the intracellular ATP/ADP ratio increased; this change leads to closure of K_{ATP} channels which results in depolarization of the cell membrane. Ca^{2+} accumulates in the membrane-depolarized cells via activation of the VGCC (Voltage-Gated Calcium Channel) or release of Ca^{2+} from the ER. In addition to the K_{ATP} -dependent pathway, protein kinase A (PKA) and protein kinase C (PKC) pathways are activated by sensing the glucose metabolic signaling in a K_{ATP} -independent manner although the molecular basis is not yet known (Bratanova-Tochkova, Cheng et al. 2002).

A Biphasic insulin secretion has been observed in the pancreas following rapid and sustained stimulation of β -cells with glucose (Henquin, Ishiyama et al. 2002). The immediate reaction, so-called first phase insulin release, is mediated by rapid release of the insulin granule pools in a Ca^{2+} -dependent manner (Straub, Shanmugam et al. 2004).

The insulin granule populations are complex, consisting of three pools including a reserve pool, a docked pool and a ready releasable. The increase of the intracellular Ca^{2+} concentration is accelerated not only by the closure of K_{ATP} channel but also by several K_{ATP} -independent mechanisms (Straub, Shanmugam et al. 2004). The orchestration of K_{ATP} -dependent and K_{ATP} -independent signals positively augments the pre-existing insulin granules to release insulin.

Under prolonged exposure to high glucose environment, the demand of insulin to maintain normal physiological glucose homeostasis drives the next round of insulin biosynthesis.

1.10 Insulin biosynthesis

Insulin is synthesized in response to prolonged elevated blood glucose level. Briefly the sequence of events is as follow: 1) production of preproinsulin transcripts in the cell nucleus; 2) packaging of preproinsulin into proinsulin in the rough endoplasmic reticulum (RER); 3) synthesis of mature insulin from proinsulin in the Golgi apparatus and 4) precipitation of insulin protein in the secretory granules.

The preproinsulin is the original single polypeptide used to form mature insulin. It consists of four regions, from N-terminal to C-terminal, a hydrophobic signal sequence (SS), B-chain polypeptide, connecting peptide (C-peptide) and A-chain polypeptide.

To process the preproinsulin to folded proinsulin the transcribed mRNA moves to a ribosome and sequentially binds to the signal recognition particle (SRP) initiating protein translation. The translated SS peptide (also called nascent chain) tags onto the SRP and migrates to a high-affinity receptor for SRP named the SRP docking protein (SRPdp) that is located at the external surface of the RER (Gilmore, Walter et al. 1982). Subsequently, the ribosome dissociates from the SRP and SRPdp and is restrained by the ribosome receptor (RR) on the ER membrane. The natural hydrophobic property of SS (22 of 24 amino acids in SS are hydrophobic) facilitates and enhances the binding affinity between ribosome-mRNA complex and the phospholipids bilayer of the ER membrane (Do, Falcone et al. 1996).

In addition to SRPdp, other membrane proteins such as translocating-chain associated membrane (TRAM) and Sec61 proteins are essential for directing the ribosome-mRNA complex to the ER membrane (High, Martoglio et al. 1993; Voigt, Jungnickel et al. 1996; Knight and High 1998). Through the joint interaction of membrane proteins the complex is placed at the site on the ER membrane termed

translocon (McCormick, Miao et al. 2003) where the SS peptide guides the extension of the preproinsulin polypeptide chain into the ER lumen. The SS peptide is excised by signal peptidase positioned on inner ER membrane and proinsulin is formed. The native proinsulin is then folded by a molecular chaperone (e.g. protein disulfide isomerase, PDI) by the formation of the disulfide bonds (Winter, Klappa et al. 2002). The vesicles filled with proinsulin bud out from the RER towards the Golgi apparatus where the zinc/calcium environment around the Golgi favors the formation of the zinc-containing proinsulin hexamer (Dodson and Steiner 1998). Proinsulin passes through cis, medial and trans Golgi cisternae and is converted to insulin at the stage of vesicle formation. This transformation is driven by membrane-associated prohormone convertase enzymes (PCs). The catalytic sites targeted by PCs are called the dibasic sites (Goodge and Hutton 2000). Subsequently the granules are packed with mature insulin coupled with the precipitation of removed C-peptide and are maintained in the cytoplasm as an “insulin secreting vesicles pool”.

Once the β -cells sense certain physiological variations of glucose level (high glucose) the insulin will be secreted through the release of these primary granules in order to maintain the normoglycaemia.

1.10.1 Transcription factors binding the Insulin gene promoter region and activating its expression:

Short-term regulation of insulin secretion, as can be in response to a meal occurs mainly at the level of exocytosis. However the maintenance of adequate intracellular stores of insulin on a long-term maintenance relies on the transcriptional and translational regulation of insulin biosynthesis. The two more important nutrients for the mammals, glucose and fatty acids, profoundly affect the preproinsulin gene under physiological and pathological (glucotoxicity and lipotoxicity) status.

Insulin, synthesized by the β -cells of the pancreas, is of major physiological importance for metabolic homeostasis. While mature insulin consists of two polypeptide

chains joined by disulphide bridges, the gene encodes for a highly conserved single chain precursor (Chan, Keim et al. 1976). In most species preproinsulin exists as a single gene, whereas in the mouse and the rat two non-allelic insulin genes are present.

The human insulin gene is located on the short arm of the chromosome 11 (p15.5) (Harper, Ullrich et al. 1981) the rat insulin I and II gene are colocalized on chromosome 1 (Soares, Ishii et al. 1985) while the mouse genes are positioned on two different chromosomes: insulin I on chromosome 19 (Davies, Poirier et al. 1994) and insulin II on chromosome 7 (Duvillie, Bucchini et al. 1998).

The structure and organization of the insulin gene has been reviewed in detail over two decades ago (Steiner, Chan et al. 1985), briefly the insulin gene in humans and the insulin II gene (in rodents) containing three exons and two introns while the insulin I (in rodents) lacks the second intron (Steiner, Chan et al. 1985). In adult islets the nonallelic genes appear to be coordinately expressed and regulated (Giddings and Carnaghi 1988; Koranyi, Permutt et al. 1989).

1.10.2 Organization of the insulin promoter region and Transcription factors binding sites:

In mouse embryo, insulin mRNA was first detected at e9 in the foregut area from which the pancreas develop (Gittes and Rutter 1992). In humans, insulin-positive cells are already present by 8 weeks of embryonic development (Polak, Bouchareb-Banaei et al. 2000) during post-natal life, insulin gene expression is restricted to the β -cells. Most studies using humans and rat genes have showed that the insulin gene expression is controlled by the 5'-flanking region of the gene (German, Ashcroft et al. 1995) (Fig. 6).

A highly conserved region (~340bp) immediately upstream of the transcription-starting site confers both tissue-specific expression and metabolic regulation of the insulin gene. Many of the transcription factors act upon this region, forming a highly sophisticated network, ensuring the precise regulation of this gene. Deletion and mutational analyses of various insulin promoter regions, conducted in transfected cell

lines and transgenic animals, led to the identification of *cis*-acting elements necessary to specific insulin gene expression.

Three are the more critical binding site identified in the promoter region of the insulin gene: A3, C1 and E1 (Poitout, Stein et al. 2004).

Pancreatic/duodenal homeobox-1 (Pdx1) binds the region called A3 box of the promoter region of the insulin gene (Melloul, Ben-Neriah et al. 1993) and functionally interacts with proteins of the bHLH family that bind the E1 box (Peers, Leonard et al. 1994). The E1 activator has been shown to be a heterodimer consisting of on ubiquitous class A member (E12/E47 and E2/5) and a cell restricted class B member (B2) member of the bHLH family (Aronheim, Ohlsson et al. 1991). The cooperation between PDX1 and bHLH proteins and coactivators is forming then a unique network of protein-protein and DNA-protein interactions (Ohneda, Mirmira et al. 2000; Qiu, Guo et al. 2002).

Recently the C1 (RIPE3b) activator has been identified as MafA (Olbrot, Rud et al. 2002). Although a number of hepatic nuclear factors and Pax family members contribute to the general network of transcription factors that control the transcription of the preproinsulin gene.

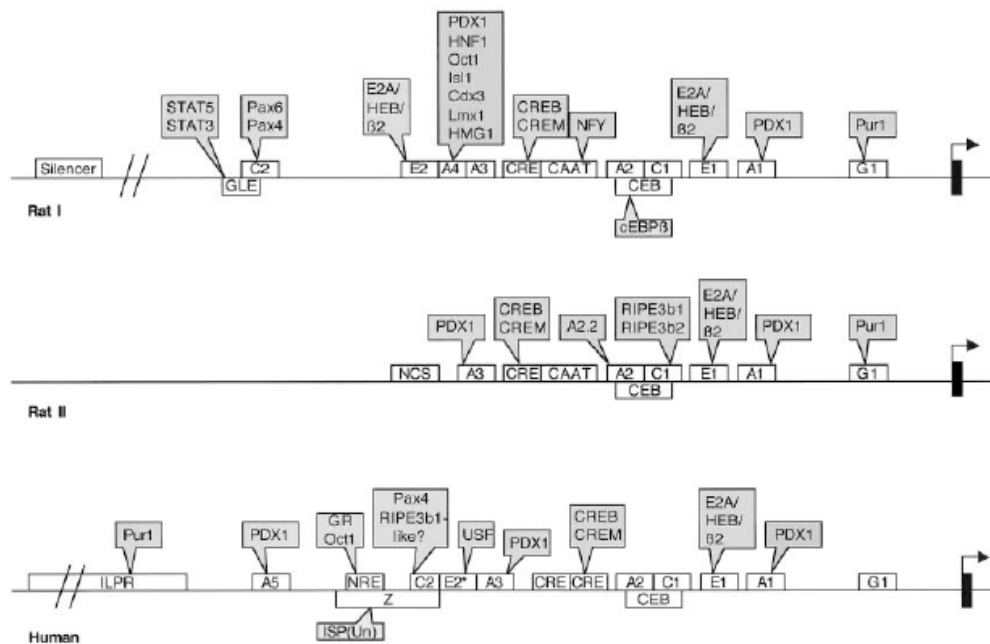


Fig 6: Schematic organization of the human insulin gene and rat insulin I and II genes. Highlighted are the *cis*-regulatory elements.

1.10.3 Transcription factors regulating insulin gene expression in mature β -cell:

Glucose controls the steps of insulin gene expression including transcription, preRNA splicing and mRNA stability. The binding site A3, E1 and C1 are the major glucose-responsive elements of the insulin gene. In addition to those, a more distal glucose-responsive element, present in primary islets, appears to bind the glucose-sensitive complex, nevertheless it is still not identified (Sander, Griffen et al. 1998).

The most studied cooperative interactions are those between the proteins binding the juxtaposed E and A elements. For that purpose, the rat insulin I promoter, and in particular a short part of the promoter region called Far-Flat region (FF), composed of an E2 motif and of the unique AT-rich A3/A4 sequence elements and considered the election region of such studies has been extensively used.

Glucose promotes the binding of PDX1 to the A3 site and its transactivation efficiency (Melloul, Ben-Neriah et al. 1993; Petersen, Serup et al. 1994). Also it is now well accepted that PDX1 action on insulin promoter region requires the recruitment of p300. The role of the ubiquitous coactivator p300 and the close homologue CBP has been investigated extensively. Those are histone acetyltransferases that stimulate transcription of specific genes by interacting either directly or through co-factors with a number of transcription factors and with the basal transcription machinery (Imhof, Yang et al. 1997).

The Homeodomain-containing factor that bind to A3/A4 region interacts also with the E2A-encoded proteins in transfected islet cells. PDX1 has been demonstrated acting in a synergistic manner with E47 to activate transcription (Serup, Petersen et al. 1995; Catasti, Chen et al. 1996; Edlund 2001). Such synergism requires the DNA binding and activation domains of both PDX1 and bHLH proteins. It was also demonstrated that the synergistic transactivation activity results not only from the coexpression of both E2A and PDX1 but also of BETA2 involving its C-terminal activation domain and a region within the N-terminus (Glick, Leshkowitz et al. 2000).

The transactivation activity of PDX1 involves also, to form an activation complex, the recruitment of the high mobility group protein 1 (HMG1(Y)) that binds

himself directly to the A3/A4 region increasing the PDX1/bHLH transactivation activity (Ohneda, Mirmira et al. 2000). PDX1 stimulation of insulin gene transcription involves recruitment of coactivators such bHLH proteins but also proteins affecting chromatin structure through methylation and/or acetylation of the histones (Mosley and Ozcan 2004; Francis, Chakrabarti et al. 2005). Furthermore, glucose appears to promote the translocation and modification of cytoplasmatic inactive PDX1 form of 31kDa to the nuclear 46kDa active form (Rafiq, Kennedy et al. 1998). Even if this translocation requires phosphorylation, the large increase in apparent molecular mass suggests that PDX1 undergoes through multiple post-translational modification like O-linked N-acetylglucosamine or small ubiquitin-related modification (Gao, Miyazaki et al. 2003; Kishi, Nakamura et al. 2003). Many kinases have been proposed to mediate PDX1 phosphorylation, e.g. p38 mitogen-activated protein kinase, phosphatidylinositol-3 kinase and extracellular signal-regulated kinases (Macfarlane, McKinnon et al. 1999; da Silva Xavier, Varadi et al. 2000; Khoo, Griffen et al. 2003). MafA (RIPE3b1) that binds the C1 element and B2/E47, binding the E1 box element, are also stimulated by glucose. Nevertheless it is still unclear how B2/E47 is regulated and how MafA expression and binding activity are regulated in response to glucose (Sharma and Stein 1994; Zhao, Guo et al. 2005).

Like the rat FF region, the C1 (RIPE3b) enhancer in rat insulin II promoter contains an E1 (ICE) box that cooperates with the RIPE3b1 element (Hwung, Gu et al. 1990; Emens, Landers et al. 1992). In such context the E47:BETA2 heterodimers require the presence of the beta-cell specific RIPE3b1-Act/C1 complex to activate gene insulin transcription (Sharma and Stein 1994; Naya, Stellrecht et al. 1995). In the human gene as well the synergism E1-A1 or E1-A2/C1 has been reported between (Odagiri, Wang et al. 1996).

The mechanisms responsible for efficient activation of the insulin promoter are still not fully clarified. Importantly, PDX1, B2 and MafA are not acting in an isolated manner but interact to each other to induce a synergic transcription of the insulin gene (Zhao, Guo et al. 2005). Also no β -cell specific activators and coactivators, beside MafA and a recently E2A interacting protein recently discovered (Thomas, Yao et al. 1999), have been identified.

Therefore glucose enhances insulin gene transcription by many complementary mechanisms that include transcription factors recruitment, histone modifications, and start of transcription.

Tissue specific expression seems at the end to be dependent of different combinations and concentration of transcription factors, only some of which is β -cell specific. Understanding of the molecular mechanisms underlying the insulin expression will then require detailed analysis of interactions between the various transcription factors and coactivators.

Chapter 2

AIMS

2 AIMS

Type 1 diabetes is an autoimmune disease in which the insulin secreting β -cells in the islet of Langerhans, that are essential for the glucose homeostasis, are destroyed by a target immune attack.

Patients affected by Type 1 diabetes are dependent of exogenous insulin, even though this is not a cure because the chronic and devastating complications cannot be prevented. Identifying and characterizing β -cell-specific genes playing critical roles during the differentiation and insulin gene expression could allow to better control the use of stem/progenitor cells to create new β -cells for diabetes treatment and improve or induce the expression of insulin.

Among the genes important for differentiation Neurogenin-3 (Ngn3) seems to play a pivotal role positioning itself at the beginning of the true differentiation pathway that discriminates an endocrine-committed cell fate from cells that are still (multi)potent. Nevertheless Ngn3's targets are still not well identified.

As for insulin gene regulation, recent findings showed that two transcription factors, A2.2 (β -cell specific transcription factor) and MafA bind the insulin promoter onto an overlapping region. The two transcription factors are not competing for the binding region so it has been hypothesized that they cooperatively activate the insulin transcription in β -cells.

Therefore the aims of my research work have been:

1. *to identify new genes involved in pancreas development*
2. *to test if an ectopic transient expression of Ngn3, mimicking as much as possible its physiological expression, is sufficient to act as generic switch to specify the beta-cell fate;*
3. *To identify the major factors that lie directly downstream the expression of Ngn3;*
4. *To identify the transcription factor A2.2.*

Chapter 3

MATERIALS AND METHODS

PRIMER TABLE

Primer Sets	Forward	Reverse
RevNgn3	5'-AAGGATCCATGGCGCCTCAT-3'	5'-GCAAGCTTTCACAAGAAGTCTGAGAAC-3'
pRevTRE-SEQ	5'-TCCACGCTGTTTTGACCTCC-3'	5'-CCCCCTTTTTCTGGAGAC-3'
mNgn3	5'-CACGAAGTGCTCAGTTCCAA-3'	5'-GCAGGGTCTCGACCTTTGTA-3'
rTubulin	5'-ATTGACCTTGTCTTGGACAG-3'	5'-GTCATAGATGGCCTCATTGT-3'
T-RexNgn3	5'-GAAGCTTGAAATGGCGCCTCATCCCTTGG-3'	5'- GGGATCCTCACAAGAAGTCTGAGAACACCAGTG C-3'
GATA4	5'-CGCCGCTGTCCGCTTCC-3'	5'-TTCGGCTTCCGTTTTCTGGTTTGA-3'
NeuroD/BETA 2	5'-CTTGGCCAAGAACTACATCTGG-3'	5'-GGAGTAGGGATGCACCGGGAA-3'
Pax4	5'-GAGATCCAACACCAGCTTTGCACTG-3'	5'-GGAGAAGATAGTCCGATTCCTGTG-3'
Pax6	5'-TGAAGCGGAAGCTGCAAAGAAA-3'	5'-TTTGGCCCTTCGATTAGAAAACC-3'
Nkx2.2	5'-ATGTCGCTGACCAACACAAAG-3'	5'-GGGAGTATTGGAGGCCCTC-3'
Isl1	5'-ATGGGAGACATGGGCGAT-3'	5'-TCATGCCTCAATAGGACTGGC-3'
PDX1	5'-CTCGCTGGGAACGCTGGAACA-3'	5'-GCTTTGGTGGATTTTCATCCACGG-3'
Insulin	5'-AAGTCCCGCCGTGAAGTG-3'	5'-TTAGTTGCAGTAGTTCTCCAGCTGG-3'
Glucagon	5'-ATGAAGACCATTTACTTTGTGGCT-3'	5'-GGTGTTTCATCAACCACTGCAC-3'
Somatostatin	5'-ATGCTGTCCTGCCGTCTCC-3'	5'-CTAACAGGATGTGAATGTCTTCCAGA-3'

AdNgn3	5'-GAGTACTGAAATGGCGCCTCATCCCTTGG-3'	5'-GCTCGAGTCACAAGAAGTCTGAGAACACCAGTGC-3'
AdNeuroD	5'-GAGTACTGCCACCATGACCAAATCATACAGC-3'	5'-GCTCGAGCCTAATCGTGAAAGATGGCATTG-3'
M13 (-20)	5'-GTAAAACGACGGCCAG-3'	
M13		5'-CAGGAAACAGCTATGAC-3'
OVO11ChiP	5'-CGGGAGGCTCTGATGC-3'	5'-GGACACGGGTTGCAACC-3'
OVO11Prom1	5'-P-GGTACCGGACACGGGTTGCAACCGGC-3'	5'-P-GAGCTCGAAAGAGTCGCCGAGGCGTG-3'
OVO11EboxMut	TACGTCCACCGATAGCCCAAGGTGCT	TTCGCCAAGACTCTTCTGGTCAAAAACAGC
NeuroDChiP	5'-CTGTCTTCTGCTGCCACAAA-3'	5'-TCACCCCTCCCTAGAACTCC-3'
OVO11	5'-CTTCTGCCACCGCAGCCCTACCTCG-3'	5'-CCTCGCAGGGTCCTGAGGCCACACCAT-3'
pGL2SEQ-5'	5'-TGTATCTTATGGTACTGTAAGT-3'	
pGL2SEQ-3'		5'-CTTTATGTTTTTGGCGTCTTCCA-3'
Amilase rat	5'-GGTTGGGGTTTTGTGCCTACTG-3'	5'-TGCCATCACTGCCAACATTTACTT-3'
GAPDH mouse	5'-ATGGCCTTCCGTGTTCTTAC-3'	5'-GGCCCTCCTGTTATTATGG-3'
GAPDH rat	5'-CGCGAAATTCAAGAGAGTAGGGAG-3'	5'-CGCGGATCCTGCCAAGTATGATGAC-3'
dTomato	5'-ATTCGTCGACACCATGGTGAGCAA-3'	5'-CTGAGTCGACTTACTTGTACAGCTCGTCCATG-3'

IRES	5'-TAATCTCCGAGCGCGTCCGAGCATGCA-3'	5'- ATTCGTCGACGGGTTGTGGCAAGCTTATCATC-3'
HMG1	5'-GGGAGGCGCATAAGAAGAAG-3'	5'-TTTTCATAAGGCTGCTTGTCATC-3'
RNA cold inducible binding protein	5'-CGCTGGAGCAGGTCTTCT-3'	5'-ACTGTAGCCACCACTCTGACTC-3'
Hystone deacet.	5'-GCGCGTTACCCACCCAGGAGGAA-3'	5'-CTGAAGTAGGTGGTGCCCGATTTG-3'
RNA binding protein 8	5'-AACTTGTTTCTTCAGGCTCTTC-3'	5'-CTGCATCATCAATGTCAGAAATC-3'
WWDB	5'-GAGGGCGGCGGAGTGAT-3'	5'-GAGGGCAGGGGTACATTCCATT-3'
TEFB	5'-CTATGGTGGCTGTGAAGGAC-3'	5'-AGCCATCAGCAGTTCCAG-3'
WTAP	5'-AAAGTTATGGCAAGAGATGAGTTA-3'	5'-GGCTTGTTCCAGTTTGTCTTTAG-3'
PHD	5'-CAGCGCTTAGCCACCTTACCC-3'	5'-TGCTGCAGCTCCTTTTACATCACT-3'
BATF	5'-CAGAAGAGCCGACAGAGACAGACACAGAAA-3'	5'-TGGCAGCACGGAGCACAGG-3'
p21SNFT	5'-CCAGCCCGCAGAGCCCCAAGGACGAT-3'	5'-GTTTCATAGGACACAGCAGCAGCGGGCACAT- 3'
Csdc2	5'-AGGGTCTGGGAGCGGGGTGGTGTC-3'	5'-AGGCCCGGGCTGTCGCTGAATAGGT-3'
CIRBP	5'-AGCAGCAGCAGAGATGA-3'	5'-AGCTGGCTCCATTGTTC-3'
RNATF	5'-CGCTGGAGCAGGTCTTC-3'	5'-GCTGGCATAGTAGTCTCTGGA-3'
HMGB4	5'-GCCCCATCCTCCTTCTGCTCTTCT-3'	5'-ATTCCCCTTCTGCCTTGACATTGGTTG-5'
RIKEN	5'-GTGTCCACTTGCTTCAGAT-3'	5'-CTAGGACACCCGTAACAG-3'

-139-113	5'-CCAGGCAAGTGT <u>TTTGGAA</u> ACTGCAGCT-3'	
-119-20m	5'-CCAGGCAAGTGT <u>TTTGGAAA</u> AGGCAGCT-3'	
-121-122m	5'-CCAGGCAAGTGT <u>TTTGGAC</u> CTGCAGCT-3'	
-124-125m	5'-CCAGGCAAGTGT <u>TTT</u> TAAACTGCAGCT-3'	
-127-128m	5'-CCAGGCAAGT <u>GGT</u> GGAAACTGCAGCT-3'	
-139-101	5'-CCAGGCAAGTGTGGAAACTGCAGCTCAGCCCCTCTG-3'	
-109.10LM	5'-CCAGGCAAGTGTGGAAACTGCAGCTC <u>CT</u> CCCCCTCTG-3'	
-113.114LM	5'-CCAGGCAAGTGTGGAAACTGCAG <u>AG</u> TCAGCCCCTCTG-3'	
-121.22LM	5'-CCAGGCAAGTGTGGAC <u>CC</u> CTGCAGCTCAGCCCCTCTG-3'	
-124.25LM	5'-CCAGGCAAGTGT <u>TTT</u> TAAACTGCAGCTCAGCCCCTCTG-3'	
-126-101	5'-CCAGGCAAGTGTGTGAAACTGCAGCT-3'	
C/EBP	5'-CACACTTGGAAATTAGCAATAGATGCAATTTGGGACTTG-3'	
C/EBPm	5'-CACACTTGGAAATTAG <u>CG</u> CTAGATGCAATTTGGGACTTG-3'	
AP1	5'-CGCTTGATGAGTCAGCCGAA-5'	
CREB	5'-ACGCTGCTGCGTCAGCAAAT-5'	

3 Materials and Methods

3.1 Cell lines culture.

NIH/3T3, F9, HIT T-15, α -TC, RIN and AR42J cell lines were from ATCC (Manassas, VA, USA). INS-1 cell line was provided by Prof. CB Wollheim (University of Geneva, Switzerland). NIH/3T3-47 were from New England Biolabs (Ipswich, MA, USA). PT67 packaging cell line was from Clontech (Mountain View, CA, USA). 293Ad cell line was from Stratagene (La Jolla, CA, USA). MIN-6 cells were from Dr. J.-I. Miyazaki (Osaka University Medical School, Japan). NIH/3T3, F9 and 293Ad were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 units/ml penicillin, and 100 mg/ml streptomycin, supplemented with 10% (v/v) heat-inactivated fetal bovine serum. NIH/3T3-47 were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 units/ml penicillin, 100 mg/ml streptomycin, and 800 μ g/ml G418 supplemented with 10% (v/v) heat-inactivated fetal bovine serum. PT67 were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 units/ml penicillin, and 100 mg/ml streptomycin, supplemented with 10% (v/v) heat-inactivated Tet System Approved fetal bovine serum. AR42J were cultivated in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose and 1mM sodium pyruvate, 25 mM HEPES, 2 mM L-glutamine, 100 units/ml penicillin, 100 mg/ml streptomycin, supplemented with 10% (v/v) heat-inactivated fetal bovine serum, and maintained at 37°C in 5% CO₂. RIN and INS-1 were cultivated in RPMI 1640 containing 4.5 g/L glucose and 1 mM sodium pyruvate, 2 mM L-glutamine, 10 mM HEPES, 50 mM β -mercaptoethanol, 100 units/ml penicillin, 100 mg/ml streptomycin, supplemented with 10% (v/v) heat-inactivated fetal bovine serum and maintained at 37°C in 5% CO₂. HIT T-15 were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 units/ml penicillin, and 100 mg/ml streptomycin, supplemented with 5% (v/v) heat-inactivated fetal bovine serum and

15% (v/v) heat-inactivated horse serum. α -TC were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 50mM β -mercaptoethanol, 100 units/ml penicillin, and 100 mg/ml streptomycin, supplemented with 10% (v/v) heat-inactivated fetal bovine serum. MIN-6 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 units/ml penicillin, and 100 mg/ml streptomycin, supplemented with 20% (v/v) heat-inactivated fetal bovine serum.

3.2 (Bio)chemicals.

Invitrogen (Carlsbad, CA, USA): DMEM, RPMI 1640, streptomycin/penicillin solution, sodium pyruvate solution, HEPES solution, glutamine solution, TOPO TA cloning kit, ZERO BLUNT PCR cloning kit, Lipfectamine2000 transfection reagent, OptiMEM medium, Trizol reagent, DEPC-treated water, UltraPure Agarose, Superscript Indirect cDNA Labeling System, Labeling Purification Module S.N.A.P. Columns, precast 10-20% Tris-glycine gel.

Amersham Biosciences (Milan, Italy): Cy3TM Mono-Reactive Dye Pack and Cy5TM Mono-Reactive Dye Pack, Mini Quick Spin Columns.

Beta Cell Biology Consortium (www.betacell.org): PanChip 6.0 array slides, Anti-Ngn3 antibody.

Stratagene (La Jolla, CA, USA): AdEasy XL Adenoviral Vector System (containing pShuttle-IRES-hrEGFP2 vector), BJ5183-AD-1 Electroporation Competent Cells.

Clontech (Mountain View, CA, USA): RevTet-On Gene Expression System (containing pRevTet-On, pRevTRE and pRevTRE-Luc vectors), Doxycycline, heat-inactivated Tet System Approved fetal bovine serum.

Santa Cruz Biotechnology (Santa Cruz, CA, USA): Rabbit anti-Neurogenin-3, goat anti- β -actin, mouse anti-MafF/G/K (C-18), mouse anti-MafA, rabbit anti-C/EBP α , rabbit anti-C/EBP β

Vector Laboratories (Burlingame, CA, USA): Rabbit normal IgG, goat anti-rabbit-IgG HRP conjugate, rabbit anti-goat-IgG HRP conjugate, Vecstastain Elite ABC.

Aviva System Biology (San Diego, CA, USA): Rabbit anti-OVO11.

Millipore (Billerica, CA, USA): EZ-ChiP kit, Immobilon-P PVDF membrane.

New England Biolabs (Ipswich, MA, USA): Protein G magnetic beads, magnetic separation rack, the blunting enzymes mix (T4 DNA polymerase and T4 polynucleotide kinase), dNTP mix, T4 DNA ligase, competent *E. coli* cells, *BamHI*, *HindIII*, *XhoI*, *SacI*, *PmeI*, *PacI*, *NotI*, *Sall*, *ApaI*, *BstEII*, SOC, Phusion Site-Direct Mutagenesis kit TransPassD2, RheoSwitch Expression system, Blue Loading Buffer Pack.

Cell Signaling Technology (Beverly, MA, USA): rabbit anti-HA tag, 20X LumiGLO® Reagent and 20X Peroxide, anti-biotin HRP-linked antibody, anti-mouse IgG HRP-linked antibody, anti-rabbit IgG HRP-linked antibody, rabbit anti-NeuroD/BETA2 antibody.

QIAGEN (Valencia, CA, USA): QIAprep miniprep, midiprep and maxiprep kits and QIAquick gel extraction kit.

Promega (Fitchburg, WI, USA): GoTaq DNA polymerase, RQ1 (RNA-Qualified) RNase-Free DNase, M-MLV Reverse Transcriptase, random hexadeoxynucleotides, FuGENE6 transfection reagent, Passive Lysis Buffer, Luciferase Assay kit, Dual Luciferase Assay kit, pGL2-Basic, pRL-TK, Sequencing Grade Modified Trypsin”.

Novagen (San Diego, CA, USA): CytoBuster Protein Extraction Reagent.

BioRad (Hercules, CA, USA): 30% Acrylamide/Bis Solution, 37.5:1 mixture (30%T, 2.67% C), BioRad Protein Assay reagent.

Applied Biosystem (Foster City, CA, USA): SYBR Green PCR Master Mix, Western Lightning Chemiluminescence Reagent, Galacto-Light Plus System.

Stop&Shop Supermarket (Brookline, MA, USA): Instant non fat dry milk.

SIGMA (St Louis, MO, USA): Paraformaldehyde, Protease Inhibitor Cocktail and β -mercaptoethanol, agarose, ethidium bromide, phenol:chloroform:isoamyl alcohol (24:25:1), 3 M Sodium Acetate pH 5.2 solution, chloroform, isopropyl alcohol, ampicillin, kanamycin, G418, hygromycin, blasticidin, LB medium, LB agar, hexadimethrine bromide (polybrene), N,N,N',N'-Tetramethylethylenediamine (TEMED), ammonium persulfate (APS), Tris base, HCl, Sodium dodecyl sulfate (SDS), Glycine, Retinoic Acid (RA), dibutyryl-c-AMP (db-cAMP), KOH, xylene, 100% ethanol. 95%

ethanol, methanol, Triton X-100, Tween-20, NaCl, KCl, 3, 3'-Diaminobenzidine (DAB), H₂O₂, Harris's Hematoxylin solution, sodium borate, citrate, paraformaldehyde, polyornithine, glycine, Triton X100, goat anti-Rabbit FITC conjugated, CHAPS, protease inhibitor cocktail, Bradford reagent, urea, dithiothreitol (DTT), iodoacetamide, silver nitrate, sodium carbonate, formaldehyde, sodium bicarbonate, potassium ferricyanide trifluoroacetic acid (TFA), 3-(N-morpholino)propanesulfonic acid (MOPS), Sodium Saline Citrate Buffer (SSC), Sarcosyl, 3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPS), Protease Inhibitor cocktail, trichloroacetic acid.

Amersham Bioscience (Arlington Heights, IL): PD-10 desalt column, IPG buffer, Immobiline DryStrip pH 3-10 and pH 4-7, HiTrap Heparin HP columns, Cy3TM Mono-Reactive Dye Pack and Cy5TM Mono-Reactive Dye Pack.

PIERCE (Rockford, IL, USA): TrichromRanger Molecular Weight mix.

Carlo Erba Reagenti (Rodano, MI, Italy): Acetonitrile.

Roche Applied Biosciences (Indianapolis, IN, USA): Mini Quick Spin Columns.

Strathmann Biotech AG (Dengelsberg, Germany): Activin A.

RELIATech GmbH (Braunschweig, Germany): Hepatocyte Growth Factor (HGF).

3.3 Buffer and Solutions

Sorensen's buffer pH 7.4

50 ml dH₂O

40.5 ml 0.2 M Sorensen's phosphate solution A

9.5 ml Sorensen's phosphate solution B

Sorensen's phosphate solution A:

35.61 g Na₂HPO₄ · 2H₂O in 1000 ml of dH₂O

Sorensen's phosphate solution B

27.6 g NaH₂PO₄ · H₂O in 1000 ml of dH₂O

Pre-hybridization buffer

5X SSC

0.1% SDS

1% BSA

Washing buffer 1

1X SSC

0.25% SDS

Washing buffer 2

1X SSC

0.1% sarcosyl

Sample Buffer (SDS Reducing Buffer)

3.55 ml dH₂O

1.25 ml 0.5 M Tris-HCl, pH 6.8

2.5 ml glycerol

2.0 ml 10% (w/v) SDS

0.2 ml 0.5% (w/v) bromophenol blue

9.5 ml Total Volume

10% SDS

Dissolve 10 g SDS in 90 ml water with gentle stirring and bring to 100 ml with dH₂O

1.5 M Tris-HCl, pH 8.8

27.23 g Tris base

80 ml of dH₂O

Adjust to pH 8.8 with 6 N HCl. Bring total volume to 150 ml with dH₂O and store at 4°C.

0.5 M Tris-HCl, pH 6.8

6 g Tris base

60 ml of dH₂O

Adjust to pH 6.8 with 6 N HCl. Bring total volume to 100 ml with dH₂O and store at 4°C.

10X Electrode (Running) Buffer, pH 8.3

30.3 g Tris base

144.0 g Glycine

10.0 g SDS

Dissolve and bring total volume up to 1,000 ml with dH₂O. Do not adjust pH with acid or base. Store at 4 °C.

Streptavidin-agarose binding buffer

12% (v/v) glycerol

12 mM HEPES-NaOH, pH 7.9

4 mM Tris-HCl, pH 7.9

60 mM KCl

1 mM DTT

Store at -20°C upon addition of DTT

Streptavidin-agarose elution buffer

12% (v/v) glycerol

20 mM Tris-HCl, pH 6.8

5 mM MgCl₂

1 M KCl

1 mM EDTA

1 mM DTT

200 µg/ml BSA

Store at -20°C upon addition of DTT

10X Binding Buffer (EMSA)

100 mM HEPES, pH 7.9

1 M NaCl

20 mM EDTA

80 mM DTT

50% glycerol

1X Buffer A

10 mM HEPES, pH 7.9

1.5 mM MgCl₂

10 mM KCl

0.5 mM DTT

1X Buffer B

20 mM HEPES, pH 7.9

25% (v/v) glycerol

0.42 M NaCl

1.5 mM MgCl₂

0.2 mM EDTA

0.5 mM PMSF

0.5 mM DTT

1X Buffer C

20 mM HEPES, pH 7.9

20% (v/v) glycerol

0.1 M KCl

0.2 mM EDTA

0.5 mM PMSF

0.5 mM DTT

Rehydration Buffer

8 M Urea,
2% CHAPS
60 mM DTT
IPG buffer pH 3-10 or pH 4-7
0.0002% bromophenol blue.

Reduction Buffer

2% SDS
50 mM Tris-HCl pH 8.8
6 M Urea
30% glycerol
0.002% bromophenol blue
65 mM DTT

Alkylation Buffer

2% SDS
50 mM Tris-HCl pH 8.8
6 M Urea
30% glycerol
0.002% bromophenol blue
135mM iodoacetamide

STE Buffer

10 mM Tris pH 8.0
50 mM NaCl
1 mM EDTA

HiTrap Binding Buffer

10 mM sodium phosphate, pH 7

HiTrap Elution Buffer

10 mM sodium phosphate, pH 7

1-2 M NaCl

Streptavidin-Agarose Binding Buffer

12% glycerol

12 mM HEPES-NaOH, pH 7.9

4 mM Tris-HCl, pH 7.9

60 mM KCl

1 mM EDTA

1 mM DTT

Streptavidin-Agarose Elution Buffer

12% Glycerol

20 mM Tris-HCl, pH 6.8

5 mM MgCl₂

1 mM EDTA

1 mM DTT

200 µg/ml BSA

3.4 Plasmid Mini-Prep

For the plasmid DNA Mini-Prep the QIAprep miniprep kit was used. The QIAprep miniprep procedure uses the modified alkaline lysis method of Birnboim and Doly (2). Bacteria are lysed under alkaline conditions, and the lysate is subsequently neutralized and adjusted to high-salt binding conditions in one step. After lysate clearing, the sample is purified on the QIAprep silica membrane.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing the appropriate antibiotic, and grown with vigorous shaking for 12–16 h. 1.5 ml of the bacterial cells were harvested by centrifugation at 6800 x g in a conventional, table-top microcentrifuge for 3 min at room temperature. After harvesting the pelleted bacterial cells were resuspended in 250 µl of Buffer P1. 250 µl Buffer P2 were then added and the tubes mixed thoroughly by inverting the tube 4–6 times. 350 µl Buffer N3 were then added, the tubes mixed immediately and thoroughly by inverting the tube 4–6 times and centrifuged for 10 min at ~17,900 x g in a table-top microcentrifuge. The cleared supernatants from step above were applied to the QIAprep spin column centrifuged for 1 min. The flow-through was discarded. The QIAprep spin column were washed by adding 750 µl of Buffer PE and centrifuged for 1 min as above. The flow-through was discarded, and the column centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep columns were placed in a clean 1.5 ml microcentrifuge tube and the DNA eluted adding 50 µl of water to the center of each QIAprep spin column, let stand for 1 min, and centrifuged for 1 min.

3.5 Plasmid Midi-Prep

For the plasmid DNA midi-prep the QIAprep midiprep kit was used. To 50 ml of LB medium containing the appropriate antibiotic, 100 µl of positive bacterial suspension were added and grown with vigorous shaking for 12–16 h. The bacterial cells were harvested by centrifugation at 600 x g for 15 min at room temperature. After harvesting

the pelleted bacterial cells were resuspended in 4 ml of Buffer P1. 4 ml Buffer P2 were then added and the tubes mixed thoroughly by inverting the tube 4–6 times and incubated 5 min at room temperature. 4 ml of cold Buffer P3 were then added, the tubes mixed immediately and thoroughly by inverting the tube 4–6 times, incubated 15 min on ice and centrifuged for 30 min at $\sim 20,000 \times g$. The cleared supernatants from the above step were applied to the QIAprep column previously equilibrated with 4 ml of buffer QBT and allowed to enter the resin by gravity flow. The flow-through was discarded. The QIAprep column was washed by adding 2 x 10 ml of buffer QC. The flow-through was discarded, the column placed in a clean 50 ml tube and eluted with 5 ml of buffer QF. The DNA was then precipitated by adding 3.5 ml of isopropyl alcohol at room-temperature and centrifuged immediately at $\geq 15,000 \times g$ for 30 min at 4°C. The supernatant was carefully decanted and the pellet washed with 2 ml of 70% ethanol, centrifuged at 15,000 x g for 10 min. The supernatant was carefully decanted and the pellet air dried and then 150-200 μl of dH_2O . The purified plasmid DNA was quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.6 Plasmid Maxi-Prep

For the plasmid DNA maxi-prep the QIAprep maxiprep kit was used. To 500 ml of LB medium containing the appropriate antibiotic, 100 μl of positive bacterial suspension were added and grown with vigorous shaking for 12–16 h. The bacterial cells were harvested by centrifugation at 600 x g for 15 min at room temperature. After harvesting the pelleted bacterial cells were resuspended in 10 ml of Buffer P1. 10 ml Buffer P2 were then added and the tubes mixed thoroughly by inverting the tube 4–6 times and incubated 5 min at room temperature. 10 ml of cold Buffer P3 were then added, the tubes mixed immediately and thoroughly by inverting the tube 4–6 times, incubated 15 min on ice and centrifuged for 30 min at $\sim 20,000 \times g$. The cleared supernatants from the above step were applied to the QIAprep column previously equilibrated with 10 ml of buffer QBT and allow it to enter the resin by gravity flow. The

flow-through was discarded. The QIAprep column was washed by adding 2 x 30 ml of buffer QC. The flow-through was discarded, the column placed in a clean 50 ml tube and eluted with 15 ml of buffer QF. The DNA was then precipitated by adding 10.5 ml of isopropyl alcohol at room-temperature and centrifuged immediately at $\geq 15,000 \times g$ for 30 min at 4°C. The supernatant was carefully decanted and the pellet washed with 5 ml of 70% ethanol, centrifuged at 15,000 x g for 10 min. The supernatant was carefully decanted and the pellet air dried and then 150-200 μ l of dH₂O.

3.7 Gel Extraction.

For the DNA gel extraction the QIAquick gel extraction kit was used. After the run the bands were excised from the agarose gel with a clean, sharp scalpel, weighed and added with 3 volumes of Buffer QG (v/w). The gel slices were incubated at 50°C until completely dissolved, added with 1 gel volume of isopropyl alcohol placed onto QIAquick spin columns and centrifuged at 10,000 x g for 1 minute. The QIAquick spin columns were then washed with 750 μ l of buffer PE and centrifuged as above. The flow-throughs were discarded and the QIAquick columns centrifuged for an additional 1 minute at 10,000 x g. The QIAquick columns were placed into a clean 1.5 ml microcentrifuge tube and the DNA eluted with 50 μ l of dH₂O and centrifuged for 1 min at 10,000 x g.

3.8 Bacteria transformation

For bacteria transformation, to the competent *E. coli* cells, thawed on ice, up to 50 ng of DNA (vector or ligation reactions) are added and the mixture incubated on ice for 30 min. The cells were then heat shocked at exactly 42°C for exactly 30 seconds and placed on ice for 5 min. After addition of 950 μ l of room temperature SOC into the mixture the cells were incubated at 37°C for 60 min with vigorous shake (250 rpm). Ten

μl for vectors and up to 100 μl of ligation reaction were plated onto a LB-agar plate containing the appropriate antibiotic and incubated overnight at 37°C. The plasmid DNA was then extracted as described above.

3.9 RNA Extraction and Reverse Transcription

The RNA was extracted by adding directly to the plate 1 ml of Trizol reagent and passing the cell lysate several times through a pipette. The homogenized samples were then transferred in a microcentrifuge tube and incubated for 5 min at room temperature. 200 μl of chloroform per 1 ml of Trizol were then added and the tube vigorously shaken for 15 sec incubated at room temperature for 5 min and centrifuged at $12,000 \times g$ for 15 min at 4°C. Following centrifugation, the upper aqueous phase is recovered and transferred in a new microcentrifuge tube. The RNA from the aqueous phase was then precipitated by mixing with 500 μl of isopropyl alcohol per 1 ml of Trizol Reagent. The samples are then incubated at room temperature for 10 min and centrifuged at $12,000 \times g$ for 10 min at 4°C. The supernatant is removed and the RNA precipitated washed once with 1 ml of 70% ethanol. The sample are mixed by vortexing and centrifuged at $7,500 \times g$ for 5 min at 4°C. At the end of the procedure, the RNA pellet was briefly dried and the RNA pellet dissolved in DEPC-treated water and stored at -70°C.

The purified DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

Prior to proceed to the Reverse Transcription reaction the RNA extracted as above is DNase treated. In a 200 μl tube the following reagent were added:

RNA	1.0 μl (1 mg)
RQ1 RNase-Free DNase 10X Reaction Buffer	1.0 μl (1X final)
RQ1 RNase-Free DNase	1.0 μl (1 Unit)
DEPC-treated H ₂ O	To 10 μl

The mixture is incubated for 1 h at 37°C. To the reaction 1 µl of RQ1 DNase Stop Solution is added to terminate the reaction and the mixture incubated at 65°C for 10 min to inactivate the DNase.

To the DNase-treated RNA 2 µl of a 50 ng/µl solution of random hexadeoxynucleotides were added and the solution incubated for 5 min at 70°C. The tubes were then rapidly cooled on ice and added with the following reagent:

M-MLV 5X Reaction Buffer	5.0 µl (1x final)
dNTP mix 10 mM	5.0 µl (2 mM)
M-MLV RT	1.0 µl (200 Units)
DEPC-treated H ₂ O	To 25µl

The mixture is incubated for 1 hr at 37°C followed by incubation at 70°C for 10 min to inactivate the M-MLV RT.

3.10 Real-Time RT-PCR

For the real-time PCR reaction the following mixture is prepared in 96-optical tube plate:

SYBR Green Mix (2X)	12.5 µl
cDNA	3.0 µl
10 µM 5'-Primer	2.5 µl (0.5 nM)
10 µM 3'-Primer	2.5 µl (0.5 nM)
Sterile dH ₂ O	To 25 µl

For the quantification the comparative C_t method has been used. This involved comparison of the C_t values of the samples of interest (samples transfected) with a calibrator (samples transfected with the empty vector). The C_t values of both the

calibrator and the samples of interest are normalized to an appropriate endogenous housekeeping gene.

The comparative C_t method is also known as the $2^{-[\Delta][\Delta]C_t}$ method, where $[\Delta][\Delta]C_t = [\Delta]C_{t,\text{sample}} - [\Delta]C_{t,\text{reference}}$. Here, $[\Delta]C_{t,\text{sample}}$ is the C_t value for any sample normalized to the endogenous housekeeping gene and $[\Delta]C_{t,\text{reference}}$ is the C_t value for the calibrator also normalized to the endogenous housekeeping gene.

3.11 Agarose-Formaldehyde Electrophoresis

RNA electrophoresis under denaturing conditions in 2.2 M formaldehyde is performed according to Maniatis et al., (1982) using the MOPS buffer system. RNA under these conditions is fully denatured and migrates according to the \log_{10} of its molecular weight. Agarose has been prepared by melting the required amount of agarose in distilled water, cooling to approximately 60°C (hand hot) and adding 37% formaldehyde and 10X MOPS to give 2.2 M formaldehyde and 1X MOPS, respectively. RNA samples have been prepared by adding up to 25 μg of RNA in a maximum of 5 μl sterile H_2O , to 15 μl RNA denaturation buffer. One μl of 10 $\mu\text{g}/\mu\text{l}$ ethidium bromide is added to aid visualization of RNA after electrophoresis. Samples are loaded into the gel and electrophoresed at no more than 5V/cm, with occasional buffer recirculation, until the leading bromophenol blue dye front has migrated approximately three quarters of the length of the gel. Visualisation of RNA is achieved by irradiation with short wave (254nm) UV light.

3.12 Protein Quantification

For the protein quantification the BioRad Protein Assay reagent was used. In a 96-well plate 5 and 10 μl of cell extracts and 10 μl of solutions containing 0.1, 0.2, 0.3, 0.4 and 0.5 $\mu\text{g}/\mu\text{l}$ of BSA were pipetted. To the well 200 μl of 1X dye reagent were

added, the plates incubated at room temperature for 5 min and the absorbance read at 595 nm in a microplate reader.

3.13 MICROARRAYS ANALYSES

3.13.1 Cell culture and differentiation.

Approximately 2×10^7 AR42J cells were plated in 100 mm cell culture dish in the appropriate medium and treated with 2 nM Activin A and 100 pM of HGF for 2 days changing medium every other day. After the differentiation period the cells were washed with cold PBS. The cell were recovered by a cell scraper centrifuged and the pellet resuspended in 1 ml of Trizol reagent. The total RNA was then extracted as described above.

3.13.2 Synthesis of Aminoallyl-Amminohexil-UTP ssDNA.

For the synthesis of aminoallyl-amminohexil-UTP ssDNA the Superscript Indirect cDNA Labeling System has been used. In 200 μ l RNase-free tubes were mixed:

1X Reaction mixture 1:

Total RNA	up to 10 μ l (up to 20 μ g)
Anchored oligo(dT) ₂₀ primer (2.5 μ g/ μ l)	2 μ l

The mixture has been incubated 5 min at 70°C then the reaction was placed on ice for 1 min. At the same time the reverse transcription reaction mixture has been prepared as follow:

1X Reaction mixture 2:

5X First strand buffer	6.0 μ l
0.1M DTT	1.5 μ l
10mM dNTPs	1.5 μ l
RNAse out (40U)	1.0 μ l
SuperScriptII	2.0 μ l
DEPC-treated H ₂ O	To 18 μ l

The above mixture (Reaction mixture 2) was combined with the RNA solution previously prepared (Reaction mixture 1) and the final mixture incubated for 2 h at 46°C.

3.13.3 Hydrolysis and Neutralization of the Reaction

To the reverse transcription reaction from above, 1.5 μ l of NaOH 1 N (fresh made from 10 N stock solution) has been added and incubated for 10 min at 70°C. Fifteen μ l of 1 N HCl have been added to neutralize the solution and gently mixed followed by addition of 20 μ l of 3 M Sodium Acetate pH 5.2 and gently mixed.

3.13.4 Purification of the cDNA

For the purification of the cDNA the Labeling Purification Module S.N.A.P. Columns have been used. Five hundred μ l of Loading Buffer have been added to the cDNA mixture previously neutralized, loaded on the S.N.A.P.TM column and centrifuged at 14,000 x g at room temperature for 1 min. The column was then washed with 700 μ l of Washing Buffer, centrifuged as above and washed again. After a further centrifugation,

the cDNA was eluted twice with 50 μ l of DEPC-treated water. The purity cDNA was measured spectrophotometrically.

3.13.5 Ethanol Precipitation of cDNA

To the cDNA purified with the S.N.A.P. column, 10 μ l of 3 M sodium acetate pH 5.2, 2 μ l of 20 mg/ml glycogen and 300 μ l of ice-cold 100% ethanol have been added and incubated overnight at -80°C . The following day the mixture was centrifuged at 14,000 x g for 20 min at 4°C . The pellet was then washed with 70% ethanol, centrifuged again for 5 min at 4°C and resuspended with 5 μ l of DEPC-treated water.

3.13.6 Fluorescent Dye Labeling of cDNA

The Cy3TM Mono-Reactive Dye Pack and Cy5TM Mono-Reactive Dye Pack were resuspended with 10 μ l of DMSO and 5 μ l of each have been used for the labeling reaction. 5 μ l of DMSO/dye and 5 μ l of cDNA previously resuspended were combined and incubated in the dark for 1 h at room temperature.

3.13.7 Purification of Labeled cDNA

For the purification of the Cy3- and Cy5-labeled cDNA the Labeling Purification Module S.N.A.P. Columns have been used. Briefly 200 μ l of 3 M sodium acetate pH 5.2 and 500 μ l of Loading Buffer plus isopropanol were added to the dye-coupled cDNA mixture, mixed by vortexing, loaded on a S.N.A.P. column and centrifuged at 14,000 x g

for 1 min at room temperature. The column was washed twice with 700 μ l of Washing Buffer and the cDNA eluted twice with 50 μ l of DEPC-treated water.

3.13.8 Assessing Labeling Efficiency

The sample, previously purified with S.N.A.P. column was assessed spectrophotometrically to confirm the labeling reaction. All the samples have been transferred into a cuvette and the absorbances measured at the following wavelength: 260 nm, 280 nm, 310 nm, 550 nm 650 nm and 750 nm. The coupling efficiency has been calculated as follow:

$$\text{Cy3}^{\text{TM}} [\text{pmol}] = [(A_{550} - A_{650}) \times 55] / 0.15$$

$$\text{Cy5}^{\text{TM}} [\text{pmol}] = [(A_{650} - A_{750}) \times 55] / 0.25$$

After the spectrophotometer reading, the samples were mixed (e.g.: Sample differentiated-Cy3 vs. Sample undifferentiated-Cy5), added with 12 μ l of sodium acetate 3M pH 5.2, 12 μ l of 20 mg/ml of glycogen, 300 μ l of ice-cold 100 % ethanol and incubated overnight at -80°C. The following day the labeled-cDNA was precipitated by centrifugation and the pellet resuspended with 1.17X Hybridization Buffer.

3.13.9 PanChip 6.0 array slides

The Mouse PanChip 6 provided by the Beta Cell Biology Consortium contains approximately 13,000 mouse cDNAs chosen for their expression in various stages of pancreatic development, many of which are not found on commercially available arrays. Many of these clone where chosen to represent every assembly found to be expressed in

the pancreas by sequence analysis of Consortium Libraries EPConClones. In addition, well characterized probes amplified from full-length sequence verified transcripts known to be involved in pancreas development and pathways relating to glucose homeostasis are included. The high level of homology between mouse and rat genes makes this array suitable for use with tissues or cells derived from the rat. The PanChip 6.0 has been printed with 306 spots per grid. As control each subgrid has a Cy3 anchor on the right bottom corner (Fig. 2).

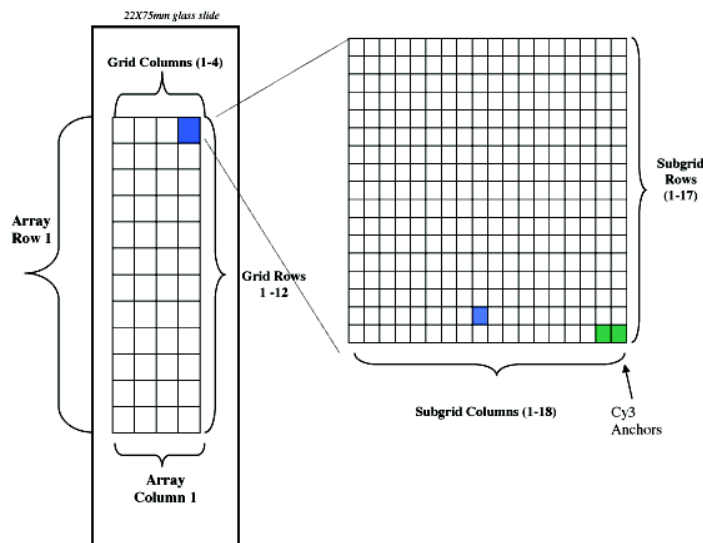


Fig. 2: The Array has been spotted on Corning ULTRA Gaps slides at the University of Pennsylvania. Each Array consists of grid of 48 subgrids (12x4) while each subgrid consists of 306 spots (17x18). The anchor Cy3 control has been placed at the right bottom corner.

The PanChip 6.0 is supported with excel and text files containing all the information about the genes spotted on the surface:

- Internal unique identifier for sample or clone
- Data source for the External_Source_ID (e.g. GeneBank accession number)
- EntrezGene gene symbol and unofficial symbols and description used for the gene and its products represented by the clone
- EntrezGene description of the clone
- Mouse Genome Information accession number for the clone
- NCBI Reference Sequence collection ID associated with the Verified_EntrezGeneID for the clone

- Chromosome information
- Map location
- PCR failure flag to determine if the PCR reaction has failed
- Sequence_Verified flag to determine if the spotted clone has been verified
- GeneBank accession number for the 5' sequence end of the spotted clone
- GeneBank accession number for the 3' sequence end of the spotted clone
- DoTS (Database Of Transcribed Sequences) for the spotted clone and description
- ID for mouse PromoterChip element corresponding to the Mouse PanChip genes
- Description of the controls:

Cy3 Anchors: random 70mer oligo labeled with Cy3

SpotReport™: positive (β -actin) and negative (*Arabidopsis thaliana* PCR products)

3.13.10 Pre-Hybridization of the array slides

For the hybridization reaction the Pre-Hybridization Buffer has been warmed for 30 min at 42°C and the microarray slides were submerged in the pre-warmed Pre-Hybridization Buffer for 45 min. After incubation the microarray slides were washed for 2 min with 0.5 SSC buffer and dried by centrifugation at 1,000 x g at room temperature.

3.13.11 Hybridization of the array slides

To the labeled-cDNA, 5 μ l of oligo(dT)-NoTI (5 μ g/ μ l) have been added in order to minimize non-specific hybridization. The mixture was first incubated at 95°C for 3 min, centrifuged and then placed on ice. The microarray slides were placed into the hybridization chamber and the labeled-cDNA mixture spread on the spotted surface of

the slide. After covering the slide with a coverslip the hybridization chamber has been closed and incubated for 16 h at 40°C in a water-bath.

3.13.12 Washing of the array slides

The day after the microarray slides were removed from the hybridization chamber and washed with: Washing Buffer 1 (pre-warmed at 42°C) for 4 min, with Washing Buffer 2 for 5 min, with 0.2X SSC for 5 min, with 0.1X SSC for 5 min and then centrifuged at 1,000 x g for 1 min. The slides were then scanned and the intensity signals were acquired for the statistical analysis.

3.13.13 Scanner Platform for Chip Reading

The **VersArray ChipReader System** scanner is a dual-laser confocal system designed for rapid imaging of microarrays. The BioRad scanner as been chosen because it allows differentiating between a spot response and background level, and optimizing the signal on the bottom as well as on the top end of the range for increased sensitivity but also for its multi-cross platform properties. The VersArray ChipReader shows low background, enhanced weak signal detection, and virtually no photobleaching for a long dye lifetime.

3.13.14 Technical characteristics VersArray ChipReader™

Chip Size: 25×75 (1×3") or other similar sizes and depths.

Scan Area: Up to 22×65 mm (0.87×2.56"), defined by user; other scan areas can be selected.

Read Time: Approx. 8 min per slide at 10 µm resolution (scanning speed 5-30 lines/sec).

Read Sequence: Simultaneous or independent.

Scanning Resolution: 3 μm model : 3 - 24 μm in 3 μm increments.

Dynamic Range: 16 bits per pixel; linear detection over the entire range.

Data Storage: 16-bit TIFF images.

Number of Lasers: 2 standard.

Excitation Wavelengths: 532 nm and 635 nm (standard).

Laser Power: 10 mW for each channel; independently controlled from 0-100%.

3.13.15 Scanning

The first scanning of the microarray slides has been done at low resolution (40 μm per pixel). Such scanning allows setting all the parameters of the photomultiplier and the laser power in order to balance the signal between the two lasers and to optimize the system. In order to get that, the count ratio (ratio between the area under the curves in the gap of the histogram defined by the two vertical cursors of minimum and maximum considered) must be as much as possible close to the unit and the histogram must be as large as possible, without outclassing the sensitivity of the instrument. With the first low-resolution scanning it is possible also decide the area that will be scanned with the high-resolution scan.

3.13.16 Spot Determination

After the first scanning at low-resolution, the software VersArray Image Acquisition needed to be set with defined parameters for the spot position determination. Those parameters were the grid and sub-grid number and their position and the software has centered automatically the laser on the spots on the array. After that the array has been checked (Fig. 1a) and the laser position on each spot has been corrected (Fig 1b), if

necessary, so the laser was able to read exactly the spot surface without external interferences such as dust on the array surface (Fig. 1c) or be miss-centered because the wrong shape of the spotted cDNA (Fig. 1d).

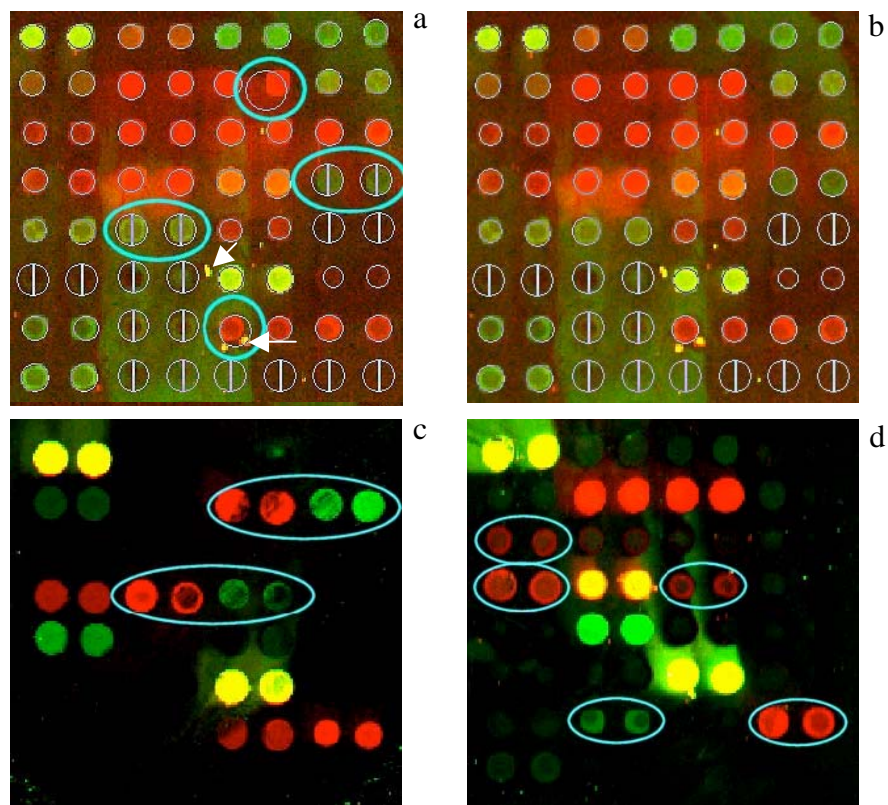


Fig.1: after the low-resolution scan and after settled the grid and subgrid parameters the laser was not focused exactly on all the spots (a) that because sometime the shape (c and d) of the spots are not perfect but also because dust on the glass surface (white arrow). For that after the first scan the laser position needs to be adjusted (b).

Once that the parameters have been chosen a high-resolution scan ($3\ \mu\text{m}$) has been performed and **.tif** images have been produced. The images are composed of two separated monochromatic images for both the channels that the software compare and analyse as a single, colored, image. The scan gives a **.txt** file as output with different information about the slide. The informations are: *GridName*, *Column*, *Row*, *CentroidX*, *CentroidY*, *Signal Average*, *Signal Median*, *Signal Std*, *Signal Pixels*, *Background Average*, *Background Median*, *Background Std*, and *Background Pixels*. The Signal

Average and the Background Average have been used for the following statistical analysis.

3.13.17 Software for Microarray Statistical Analysis

TIGR-MIDAS (from TIGR-Software <http://www.tigr.org/>)

R (from The R project for statistical computing, <http://www.r-project.org/>)

mArray (from Bioconductor, <http://bioconductor.org/repository/release1.5/package/html/index.html>)

S.A.M. (significance of microarray analysis, from Stanford Univeristy, <http://www-stat.stanford.edu/~tibs/SAM/>)

Several software tools that combine data normalization are currently available. Those software packages were analyzed by Dudoit et al. (Dudoit, Genteman et al. 2003) with particular emphasis on the TIGR-MIDAS microarray software suite, Bioconductor in R and BioArray software Enviroment (BASE) system. For this Thesis project the analysis has been done with the TIGR-MIDAS software suite, the Bioconductor in R packege and the S.A.M. package for Microsoft Excel. Briefly the TIGR Microarray Data Analysis System is a member of a suite of microarray data management and analysis applications developed at The Institute for Genomic Research (TIGR). TIGR-MIDAS is an application that allows performing normalization and other statistical data analysis such background subtraction, total normalization, Lowess normalization, and S.A.M. analysis. The Bioconductor in R (Genteman, Carey et al. 2004; Team 2007) is an open source language environment for statistical computing. The expression ratio for each element on the array was calculated in term of $\log_2(\text{Red}/\text{Green})$. The data were normalized by total and Lowess method using the Bioconductor package mArray

(Genteman, Carey et al. 2004). The analysis was carried out with RNA from five replicate experiments at the same time point of treatment.

Analyzing a microarray slide means to manage some thousands or tens of thousands different values for each parameter for each of the two colors used to label the different samples. That means that in order to analyze the data deriving from an array analysis it must be chosen a well-defined approach in order to get high-quality and reproducible measures of the hybridization intensities. For that purpose before comparing the levels in a “secure way”, a number of transformations must be carried out on the data to eliminate questionable low-quality measurements, to adjust the measured intensities to facilitate comparison and to select genes that are significantly differentially expressed between classes of samples. The first step has been the primary comparison of the signals (Red and Green).

3.13.18 Comparison of the Signals

The hypothesis is that we have data from an array of a certain number of spotted probes (*e.g.* from 1 to 100,000 and beyond). So the case is, the array has N_{array} distinct elements, and query elements and reference elements have to be compared. The comparison of those elements derives from the ratio of the corresponding element. If the query sample has been labeled with the green dye (G) and the reference sample has been labeled with the red dye (R) the ratio (C) will be:

$$C_i = \frac{G_i}{R_i}$$

where i is the index for all the arrayed genes from 1 to N_{array} . Even if the ratio gives a measure of the expression change, it has the drawback to consider differently up- and down-regulated genes. In fact, genes up-regulated by a factor of 2 have an expression ratio of 2, whereas those down-regulated by the same factor have an expression ratio of -

0.5. For that reason is required a further transformation, instead to use the simple ratio between the query and reference sample, a logarithmic transformation has been introduced:

$$C_i = \log_2 \frac{G_i}{R_i}$$

The base 2 logarithmic transformation ensure that all the expression changes are weighted in the same way. In fact the logarithms of the expression ratios are, in such way, treated in a symmetrical way: 2 fold up-regulation, the $\log_2 (2/1) = 1$, on the other hand if the query has a 2 fold down-regulation, the $\log_2 (1/2) = -1$. Furthermore, a first transformation has been applied to the data to adjust the individual hybridization intensities to balance them appropriately. There are many of reasons why the data must be normalized: quantities of starting RNA, differences in labeling or detection efficiencies between the fluorescent dye used, and also systematic biases in the measured expression levels.

3.13.19 Normalization of the Data

3.13.19.1 Total Normalization

There are many approaches to normalize the expression levels; one of those is the *Total Normalization*, which is based on simple assumptions. The first is that the starting amount of RNA is the same between the samples, that the mass of each single molecule is the same, and therefore the number of molecules in the samples is the same; The second is that the spotted genes in the array are representing a random sampling of the gene organism or the gene pool of a particular organ. Based on these two simple hypotheses a

normalization factor can be calculated by the sum of the intensities of both channels in the following way:

$$C_{total} = \frac{\sum_{i=1}^{C_{array}} G_i}{\sum_{i=1}^{C_{array}} R_i}$$

where the G_i and R_i are the measured intensities of the i^{th} elements of the array and C_{array} is the total number of the elements on the array. One or both intensities are properly scaled in the follow way:

$$R'_K = C_{total} R_K \text{ and } G'_K = G_K$$

then:

$$M_i = \frac{G'_i}{R'_i} = \frac{1}{C_{total}} \frac{G_i}{R_i}$$

That expression corrects each ratio such that the mean ratio is equal to 1, which is as to subtract a constant from the logarithm of the expression ratio:

$$\log_2(M_i) = \log_2(M_i) - \log_2(C_{total})$$

That result is a mean value of the logarithm ratio equal to 0. After the total normalization a second normalization called Lowess has been applied to the array values.

3.13.19.2 Lowess Normalization

In addition to the total intensity normalization many other functions can be applied to the analysis of microarray data: *Linear Regressions* (Chatterjee and Price 1991), *Rank Invariant Methods* (Tseng, Oh et al. 2001) and *Chen's Ratio Statistics* (Chen, Dougherty et al. 1997), just to mention some of those. Nevertheless the main limitation of such approaches is that none of them consider the systematic biases that may appear in the data. In fact several reports have indicated that the \log_2 ratio values can have systematic dependence on the intensities (Yang, Chen et al. 2002; Yang, Dudoit et al. 2002), that is the most common cause of deviation from zero in spots with low-intensity values. For that purpose a local weighted *Lowess Normalization* (or *Linear Regression Analysis*) has been applied after the total intensity normalization in the microarray analysis. Lowess uses a weight function that does not increase the contribution of data from array elements that are far from each point. The effects of the Lowess normalization are evident when the data from the array are plotted in $\log_2(G_i/R_i)$ as function of $\log(G_i \cdot R_i)$ also known as “*R-I Plot*”. (Figura dai grafici degli array). Basically when the values $x_i = \log(G_i \cdot R_i)$ and $y_i = \log_2(G_i/R_i)$ are assessed the Lowess normalization function estimates the $y(x_k)$ and then uses the function, point by point, to calculate a correct $\log_2(\text{ratio})$. In that way the corrected value became:

$$\log_2 = \log_2 \left(M_i \cdot \frac{1}{2^{y(x_i)}} \right) = \log_2 \left(\frac{G_i}{R_i} \cdot \frac{1}{2^{y(x_i)}} \right)$$

so, as see before for the total intensity normalization, the transformation for the intensities is:

$$R'_i = R_i \cdot 2^{y(x_i)} \text{ and } G'_i = G_i$$

after the normalization, a *Variance Regularization* has been done.

3.13.19.3 Variance Regularization

Nevertheless with the total and Lowess normalizations the mean of the $\log_2(\text{ratio})$ has been corrected, some stochastic process can cause a variance of measured logarithmic ratio between the regions of the array or between different arrays. To avoid such situation, a possible approach has been proposed by Yang et al. 2002 and Huber et al. 2002 (Huber, von Heydebreck et al. 2002; Yang, Dudoit et al. 2002). Considering a single array with distinct subgrids, as the array platform used for my study, after a local normalization there is a factor for each subgrid that can be used to scale all the measurements within that subgrid. In this case a good scaling factor is the variance for a particular subgrid divided by the geometric mean of the variances for all the subgrids. So, considering that each array slide has n -grids and considering that each subgrid has Q elements (spots) already adjusted to the mean of the logarithm ratio to zero, the variance of the n^{th} subgrid will be:

$$\sigma_n^2 = \sum_{j=1}^Q [\log_2(M_j)]^2$$

if the number of subgrids in the array is N_{grids} the best scaling factor for the k^{th} subgrid elements is:

$$e_k = \frac{\sigma_k^2}{\left[\prod_{n=1}^{N_{\text{grids}}} \sigma_n^2 \right]^{1/N_{\text{grids}}}}$$

after that all the elements of the k^{th} subgrid are divided for the $e_{k\text{th}}$ value calculated for the same subgrid, equivalent to taking the $e_{k\text{th}}$ cause of the individual intensities in the k^{th} subgrid:

$$\log_2(M'_i) = \frac{\log_2(M_i)}{e_k}; \quad R'_i = [R_i]^{1/e_k} \quad \text{and} \quad G'_i = [G_i]^{1/e_k}$$

The same kind of subgrid regularization can be used for the regularization between different arrays. When the array data are analyzed by an “*R-I Plot*” it can be noticed that the variability in the measured logarithmic ratio increase if the measured intensity values are decreasing. For that the relative error tends to increase for low intensities levels, because the signal is close to the background noise. A solution of that problem is to use a threshold baseline higher than the background signal or screen the signal intensities to use only those that are statistically different from the signal background, which is two standard deviations above the background level.

3.13.19.4 Replicate Filtering

Replication is essential for identifying and reducing the variation in an experiment and the microarray experiment is not an exception. Technical replicates include replicated elements within a single array, multiple independent elements for a particular gene within an array, or replicated hybridizations for a particular sample. One widely used technical replication in two-color array assay is the *Dye-reversal* or *Flip Dye Analysis*, which consists of duplicate labeling and hybridization by swapping the fluorescent dyes used for each RNA sample. So the analysis will be:

$$M_{1i} = \frac{R_{1i}}{G_{1i}} = \frac{A_{1i}}{B_{1i}} \quad \text{and} \quad M_{2i} = \frac{R_{2i}}{G_{2i}} = \frac{B_{2i}}{A_{2i}}$$

where A and B are two samples labeled respectively with red and green first and then the color swapped and labeled with green and red respectively. If we compare the samples we expect that the differences between the normalized samples is zero:

$$\log_2(M_{1i} \cdot M_{2i}) = \log_2\left(\frac{A_{1i}}{B_{1i}} \cdot \frac{B_{2i}}{A_{2i}}\right) = 0$$

Therefore, the values expected to be “good values” are close to zero for the $\log_2(M_{1i} * M_{2i})$ while the “bad values” are far from the zero. Depending the stringency set, the values with a $\log_2(M_{1i} * M_{2i})$ will be within a certain number of standard deviations of the mean.

3.13.19.5 Averaging over the replicates experiments

To reduce the complexity of the data, an average of the replicate measures has been done. If the treated and untreated sample are A and B respectively, and c_i the constant used to correct the array, the averaging between the replicates as been calculated in the following way:

$$\log_2\left(\frac{A_{1i}}{B_{1i}}\right) + c_i = \log_2\left(\frac{A_{2i}}{B_{2i}}\right) - c_i \Rightarrow c_i = \frac{1}{2} \log_2\left(\frac{A_{2i} B_{1i}}{B_{2i} A_{1i}}\right) = \log_2\left[\left(\frac{A_{2i} B_{1i}}{B_{2i} A_{1i}}\right)^{\frac{1}{2}}\right]$$

errors in the replicates experiments are corrected following the error propagation theory (Bevington and Robinson 1991):

in general for a function $f(x)$ the error will be:

$$f(x)_{error} \Rightarrow \sigma_x^2 = \frac{\partial f}{\partial x}$$

because we are working with the array the function will be the logarithmic ratio of the intensities:

$$f = \log_2\left(\frac{A_i}{B_i}\right) = \log_2(A_i) - \log_2(B_i)$$

but for the logarithm properties:

$$\log_2(x) = \frac{\ln x}{\ln 2}$$

for the partial derivatives properties:

$$\frac{\partial}{\partial x}(\ln(x)) = \frac{1}{x}$$

then the error for the average logarithm is calculated as follow:

$$\frac{\partial}{\partial x}(\log_2(x)) = \frac{\partial}{\partial x}\left(\frac{\ln x}{\ln 2}\right) = \left(\frac{1}{\ln 2}\right) \frac{\partial}{\partial x} \ln x = \left(\frac{1}{\ln 2}\right) \frac{1}{x} = \frac{1}{x \ln 2}$$

and

$$\sigma_r^2 = \left(\frac{\partial_r}{\partial_A}\right) \sigma_A^2 + \left(\frac{\partial_r}{\partial_B}\right) \sigma_B^2 = \left(\frac{1}{A \ln 2}\right) \sigma_A^2 + \left(\frac{1}{B \ln 2}\right) \sigma_B^2$$

The estimated error for A and B has been used to estimate the error for the ratio.

3.13.19.6 Significance Analysis of Microarray (S.A.M.)

The *S.A.M.* software (Tusher et al. 2001 implemented as in Chu et al. 2002), distributed under GNU public license and running on Microsoft Excel, pick-out significant genes based on differential expression between sets of samples. The *a priori* hypothesis is that some genes will have significantly different expression mean levels between different samples sets, as for example between different treated samples. A valuable feature of *S.A.M.* is that it gives an estimate of the *False Discovery Rate* (FDR), which is the proportion of genes likely identified by chance being significant. Moreover it

allows looking at the t-test and then let set thresholds for significance (delta parameter) after looking the distribution, rendering the data-miming process more sensitive. In particular for the microarray analysis the function “one-class” has been used, where it is possible to specify the value against which the mean expression of each gene will be tested. The gene expression is considered significant if the mean \log_2 of expression ratios is significantly different from the specified mean.

3.14 Immunohistochemistry

2×10^5 cells were plated the day before the differentiation treatment on polyornithine coated glass coverslips in a 12-well cell culture plate. The cells were treated with medium containing 2 nM activin A and 500 pM HGF for 2 days. After 48h the medium was removed, the cells were washed with PBS 1X and fixed with a 4% paraformaldehyde solution for 10 min. The cells were then permeabilized with a 0.1% TritonX100 solution in PBS at room temperature and then saturated with the same solution completed with 1 M glycine and incubated for 45 min at room temperature. After saturation the cells were incubated overnight with anti-insulin antibody (1:100) at 4°C. Following the overnight incubation the cells were washed and incubated for 30 min at room temperature anti-Rabbit FITC conjugated (1:200). The samples have been treated with DAPI and mounted with MOWIOL and analyzed with a Leica DC 300 microscope.

3.15 PROTEOMIC ANALYSIS

3.15.1 Cell culture and differentiation.

Approximately 2×10^7 AR42J cells were plated in 100 mm cell culture dish in the appropriate medium and treated with 2 nM Activin A and 500 pM of HGF for 2 days

changing medium every other day. After the differentiation period the cells were washed with cold PBS. The cell were recovered by a cell scraper centrifuged and the pellet extracted as described below.

3.15.2 Protein Extraction

The cell pellet from above were washed with 5 ml of PBS and frozen at -80°C . After 1 h the pellet has been re-suspended with 2.5 ml of a PBS solution containing 2% CHAPS and protease inhibitors and sonicated three times for 30 sec. The protein solution was then centrifuged at $14,000 \times g$ for 25 min at 4°C and the solution loaded on a PD10 desalting. The flow-through was discarded and the proteins eluted with 3.5 ml dH_2O collecting the flow-through. The concentration of the eluted protein was determined as described above. The protein samples were then divided in $50 \mu\text{g}$ aliquots and freeze-dried.

3.15.3 First-dimension: isoelectric focusing

Freeze-dried $50\mu\text{g}$ protein aliquot prepared as above were resuspended with $125 \mu\text{l}$ of Rehydration Buffer for 1 h at room on both 13 cm pH 3-10 and pH 4-7 Immobiline DryStrips. The strips were positioned on the “EttanIPGphor” system and run for 12 h at 50V. For the first dimension the following isoelectric focusing program has been used:

Strip pH 3-10

STEP	Voltage	V/h	Mode
Step 1	1000 V	100 V/h	Gradient
Step 2	6000 V	18000 V/h	Gradient
Step 3	8000 V	160000 V/h	Step-n-Hold
Step 4	10 V	100 V/h	Standby

Strip pH 4-7

STEP	Voltage	V/h	Mode
Step 1	500 V	500 V/h	Gradient
Step 2	3500 V	5250 V/h	Gradient
Step 3	8000 V	28000 V/h	Step-n-Hold
Step 4	10 V	100V/h	Standby

3.15.4 Second-dimension: SDS-PAGE

Before the SDS-PAGE, two steps of equilibration have been performed. The first step was the reduction of the disulphide bonds using the Reduction Buffer followed by an alkylation of the thiol groups of the protein using the Alkylation Buffer. The second step was an equilibration step with the Running Buffer. The strips were then placed on the cathode side of an 8 % polyacrilamide running gel and run at 100V for 30 min and at 200V for 4.3 h.

3.15.5 Silver Staining

After the second dimension separation, the gels were washed with milliQ water, fixed for 30 min with a solution containing 45% methanol and 5% acetic acid and then washed with milliQ water for 1 h. Following the washes the gels were treated for 2 min with a solution containing 0.02% sodium thiosulphate and briefly washed with milliQ water. The gels have been then incubated with a 0.1% silver nitrate solution for 30 min and then washed with milliQ water as described before. The gel development has been performed using a 3% sodium bicarbonate/0.03% formaldehyde solution followed by a blocking treatment with a 1% acetic acid solution. The analysis of differentially expressed proteins was done using the ImageMaster system from Amersham Bioscience; the images have been processed and digitalized.

3.15.16 Gel Digestion, Protein Reduction, Alkylation and Tryptic Digestion

Individual differential expressed spot identified as above was excised from the gel and washed for 15 min with milliQ water, followed by a destaining step with 50 μ l of a solution containing 15 mM potassium ferricyanide/50mM sodium thiosulphate for 10 min at room temperature. The gel slices were then washed with milliQ water and then with acetonitrile for 15 min at room temperature. The slices were then incubated for 15 min with 100 nM ammonium bicarbonate solution and washed for 15 min with 300 μ l of a solution of 100 nM ammonium bicarbonate/acetonitrile (1:1) and incubated. Each slice was dehydrated with acetonitrile and dried using a vacuum-drier unit. The proteins in the gel slices were rehydrated and reduced for 1 h at 56°C with a solution of 10 mM DTT in 100 mM NH_4HCO_3 and then alkylated for 30 min in the dark with a solution of 50 mM iodoacetamide in 100 mM NH_4HCO_3 . The samples were then washed twice for 15 min with 100 mM ammonium bicarbonate solution and then with iodoacetamide in 20 mM ammonium bicarbonate (1:1) solution. After that the gel slices were dehydrated with acetonitrile and dried using a vacuum-drier unit. The slices were then rehydrated with a

100 mM iodoacetamide solution containing 0.1 µg/µl trypsin and incubated overnight at 37°C. After trypsin digestion, the protein peptides were collected and the minced gels were subjected to a second extraction using a 60% acetonitrile/1% acetic acid solution for 10 min. The supernatants from the two extractions were combined, vacuum-dried and stored at -80°C until the MS analysis.

3.15.17 Peptide Analysis

For the peptide analysis has been used the LC/MS system (HPLC Surveyor unit connected to the LCQ Deca-Ion Quadrupole). Data analysis and protein identification were performed using the Turbo Sequest Search software connected to the Mascot Search (Matrix Science) Database.

3.16 Preparation of the Rev-Ngn3 viral particles

3.16.1 Cloning the mouse Ngn3 CDS into the pRevTRE vector

The complete mouse Ngn3 CDS sequence has been amplified by PCR using a specific set of primers reported in Table 1 (RevNgn3 primers set) and pBS-Ngn3 as template. Typically the PCR mixture contained:

5X GoTaq Buffer	10 µl (1X final)
25 mM MgCl ₂	7.5 µl (3.75 mM)
10 mM dNTP	2.5 µl (0.5 mM)
10 µM 5'-Primer	2.5 µl (0.5 µM)

10 μ M 3'-Primer	2.5 μ l (0.5 μ M)
Template (100 ng/ μ l)	1.0 μ l
GoTaq	0.5 μ l (2.5 Units)
Sterile dH ₂ O	To 50 μ l

Following 28 cycles of amplification, 5 μ l of PCR product was run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 μ l of dH₂O. The purified PCR fragment was used to perform a TOPO cloning reaction using the TOPO TA cloning kit. In a 200 μ l tube were mixed:

Purified PCR	4.0 μ l
Salt Solution	1.0 μ l
pCR2.1-TOPO vector	1.0 μ l

The reaction was mixed gently and incubated for 5 min at room temperature. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 50 μ g/ml of kanamycin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 2	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Mini-prep	5.0 μ l

<i>BamHI</i>	0.3 µl (6 Units)
<i>HindIII</i>	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before.

The DNA plasmid midi-prep (pCR2.1-TOPO-Ngn3) was analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Midi-prep (100 ng/µl)	1.0 µl
<i>BamHI</i>	0.3 µl (6 Units)
<i>HindIII</i>	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction mix was gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

The Ngn3 CDS from the pCR2.1-TOPO-Ngn3 vector was subcloned into the pRevTRE vector by restriction digestion with *BamHI* and *HindIII* as follow. In two 1.5 ml microcentrifuge tubes were added respectively:

10X NEB Buffer 2	10 µl (1X final)
100X BSA	1.0 µl (1X final)
pRevTRE	20 µl (28 µg final)
<i>BamHI</i>	1.5 µl (30 Units)
<i>HindIII</i>	1.5 µl (30 Units)
Sterile dH ₂ O	To 100 µl

10X NEB Buffer 2	10 μ l (1X final)
100X BSA	1.0 μ l (1X final)
pCR2.1-TOPO-Ngn3	20 μ l (15 μ g final)
<i>Bam</i> HI	1.0 μ l (20 Units)
<i>Hind</i> III	1.0 μ l (20 Units)
Sterile dH ₂ O	To 100 μ l

The reactions were gently mixed and incubated for 1 h at 37°C. 2 μ l of the digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 348 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 20 μ l of dH₂O. The digestion products were run on a 1% agarose gel, analyzed by ethidium bromide staining and gel extracted as described above. The purified DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 ml tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 μ l (1X final)
pRevTRE <i>Bam</i> HI/ <i>Hind</i> III	1.0 μ l (25 ng final)
Ngn3 <i>Bam</i> HI/ <i>Hind</i> III	1.0 μ l (7 ng final)
T4 DNA ligase	0.3 μ l (600 U final)
Sterile dH ₂ O variable	To 20 μ l

The mixture was incubated at 16°C for 16 h. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 100 µg/ml of ampicillin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Mini-prep	5.0 µl
<i>Bam</i> HI	0.3 µl (6 Units)
<i>Hind</i> III	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described above. The purified plasmid DNA (pRevTRE-Ngn3) was quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The DNA plasmid midi-prep was analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Midi-prep (100 ng/µl)	1.0 µl
<i>Bam</i> HI	0.3 µl (6 Units)
<i>Hind</i> III	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction mix was gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

3.16.2 Selection of Stable Virus-Producing PT67 Packaging Cell Line

The RetroPack PT67 Packaging Cell Line was cultivated in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose and 110 mg/L sodium pyruvate and 4 mM L-glutamine, supplemented with 10% (v/v) heat-inactivated Tet System Approved fetal bovine serum. The day before the transfection 8×10^5 cells were plated into three 60 mm cell culture dishes with 4 ml of medium. For the transfection in sterile microcentrifuge tubes 200 μ l of DMEM (containing 4.5 g/L glucose and 110 mg/L sodium pyruvate and 4 mM L-glutamine), 6 μ l of FuGENE 6 were added and incubated for 5 min at room temperature. 2 μ g of pRevTet-On, pRevTRE-Ngn3 or pRevTRE-Luc DNA were added, respectively, and the mixtures incubated at room temperature for 15 min. The FuGENE6/DNA complexes were then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were incubated at 37°C for 2 days. After that the packaging cells transfected with pRevTet-On, pRevTRE-Ngn3 and pRevTRE-Luc plasmids are diluted 1:500 and plated in selection medium containing 300 μ g/ml of G418 (for pRevTet-On transfected cells) and 150 mg/ml of hygromycin (for the pRevTRE-Ngn3 and pRevTRE-Luc transfected cells), respectively. The culture plates were incubated at 37°C for 7–10 days, replenishing the growth medium when needed. At least 24 single colonies were obtained for each vector.

3.16.3 Determining the Viral Titer

This assay is used to determine the titer (pfu/ml) of the viral stocks. 5×10^5 NIH/3T3 cells were plated in two 6-well tissue culture plates overnight at 37°C. The retroviral stocks for RevTet-On and RevTRE-Ngn3, obtained as above were diluted in 1 ml volumes over a 10-fold series from 10^{-5} to 10^{-9} in growth medium containing 4 μ g/ml of polybrene, respectively. Each dilution were then added to a separate well of the 6-well plates and incubated at 37°C for 48 h. 48 h post-infection, cells infected with RevTet-On

viruses were subjected G418 selection (500 µg/ml) and those infected with RevTRE-Ngn3 were subjected hygromycin selection (300 µg/ml) for one week. The titer was determined by counting the number of colonies obtained with the highest viral dilution and multiplying that number by the dilution factor to estimate pfu/ml. The colonies showing a titer of about 1×10^8 pfu/ml were expended. Ten high titer colonies producing RevTet-On viruses (average of 5×10^8 pfu/ml) and 8 high titer (average of 3×10^8 pfu/ml) colonies producing RevTRE-Ngn3 viruses were obtained.

3.16.4 Establishing Stable Regulatory Tet-On-F9 Cell Lines

Approximately 2×10^6 F9 cells were plated in 60 mm cell culture dish in the appropriate medium and infected with 5 MOI of RevTet-On viruses for 24 h in presence of 4 µg/ml of polybrene. The medium was replaced after 24 h with fresh one. The cells were maintained in culture for two days. The infected target cells were diluted 1:500 and plated in selection medium containing 300 µg/ml of G418 for 7 days, changing medium as necessary. An average of 15 healthy colonies were obtained and expanded.

The expended stable colonies were assayed to detect those with the higher inducibility and the lower background.

5×10^5 RevTet-On-F9 cells were plated in 6-well tissue culture plates overnight at 37°C. The cells were then infected with 5 MOI of RevTRE-Luc and incubated at 37°C for 24 h. After 24 h of incubation the medium was changed with fresh one containing 1 µg/ml of doxycycline and incubated for further 48 h.

As control to evaluate the background another set of the same cells were treated as above but with no addition of doxycycline.

The cells were then washed twice with PBS and lysed with 250 µl of 1X Passive Lysis Buffer. The plates were then rocked several times to ensure complete coverage of the cells with lysis buffer. The cells were then scraped and all the liquid transferred to a microcentrifuge tube, vortexed and centrifuged at $12,000 \times g$ for 10 min at 4°C.

For the Luciferase assay, 100 µl of Luciferase assay reagent for each samples were dispensed into luminometer tubes and 20 µl of cell lysates were added. Immediately

the tubes were read in a luminometer programmed to perform a 2-second measurement delay followed by a 10-second measurement read for Luciferase activity.

The fold induction was calculated as follow:

$$\text{Fold-induction} = +\text{Dox RLU}/-\text{Dox RLU}$$

The clones showing the highest fold-induction (highest expression with lowest background) were propagated. About 8 clones (Tet-On-F9) were obtained showing about a 15-fold induction.

3.16.5 Establishing a Double-Stable, Inducible Tet-On/TRE-Ngn3-F9 Cell Line

Approximately 2×10^7 Tet-On-F9 cells from one of the clone obtained above were plated in 100 mm cell culture dish in the appropriate medium and infected with 5 MOI of RevTet-On viruses for 24 h in presence of 4 $\mu\text{g}/\text{ml}$ of polybrene. The medium was replaced after 24 h with fresh one. The cells were maintained in culture for two days. Two days after infection the target cells were diluted 1:500 and plated in selection medium containing 150 $\mu\text{g}/\text{ml}$ of hygromycin for 7 days, changing medium as necessary. An average of 20 healthy colonies (Tet-On/TRE-Ngn3-F9) were obtained and expanded.

3.16.6 Doxycycline-induced expression of Ngn3

Approximately 5×10^5 Tet-On/TRE-Ngn3-F9 cells from ten of the clones obtained above were plated in 6-well tissue culture plates in the appropriate medium and tested for Doxycycline-regulated gene expression by assaying for Ngn3 expression in presence of 0, 2, 4, 6, 8 and 10 $\mu\text{g}/\text{ml}$ of doxycycline for at least 48 h.

The cells were then washed twice with PBS and RNA extracted with 1 ml of Trizol reagent as described above. One μg of total RNA was then DNase treated and reverse transcribed as described above.

Expression of Ngn3 has been evaluated by PCR using a specific set of primers reported in Table 1 (Ngn3 primers set).

3.17 Preparation of An Ngn3-TETRACYCLINE-INDUCIBLE SYSTEM (T-REx System)

3.17.1 Cloning the mouse Ngn3 CDS into the pcDNA5/TO vector

The complete mouse Ngn3 CDS sequence has been amplified by PCR using a specific set of primers reported in Table 1 (T-RexNgn3 primers set) and pBS-Ngn3 as template. Typically the PCR mixture contained:

5X GoTaq Buffer	10 μl (1X final)
25 mM MgCl_2	7.5 μl (3.75 mM)
10 mM dNTP	2.5 μl (0.5 mM)
10 μM 5'-Primer	2.5 μl (0.5 nM)
10 μM 3'-Primer	2.5 μl (0.5 nM)
Template (100 ng/ μl)	1.0 μl
GoTaq	0.5 μl (2.5 Units)
Sterile dH_2O	To 50 μl

Following 28 cycles of amplification, 5 μl of PCR product was run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 μl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium

Acetate pH 5.2 and 0.7 volumes of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 µl of dH₂O. The purified PCR fragment was used to perform a TOPO cloning reaction using the TOPO TA cloning kit. In a 200 µl tube were mixed:

Purified PCR	4.0 µl
Salt Solution	1.0 µl
pCR2.1-TOPO vector	1.0 µl

The reaction was mixed gently and incubated for 5 min at room temperature. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 50 µg/ml of kanamycin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Mini-prep	5.0 µl
<i>Bam</i> HI	0.3 µl (6 Units)
<i>Hind</i> III	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before.

The DNA plasmid midi-prep (pCR2.1-TOPO-Ngn3) was analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Midi-prep (100 ng/µl)	1.0 µl
<i>BamHI</i>	0.3 µl (6 Units)
<i>HindIII</i>	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction mix was gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

The Ngn3 CDS from the pCR2.1-TOPO-Ngn3 vector was subcloned into the pcDNA5/TO vector by restriction digestion with *HindIII* and *BamHI* as follow. In two 1.5 ml microcentrifuge tubes were added respectively:

10X NEB Buffer 2	10 µl (1X final)
100X BSA	1.0 µl (1X final)
pcDNA5/TO	20 µl (30 µg final)
<i>BamHI</i>	1.5 µl (30 Units)
<i>HindIII</i>	1.5 µl (30 Units)
Sterile dH ₂ O	To 100 µl

10X NEB Buffer 2	10 µl (1X final)
100X BSA	1.0 µl (1X final)
pCR2.1-TOPO-Ngn3	20 µl (20 µg final)
<i>BamHI</i>	1.0 µl (20 Units)
<i>HindIII</i>	1.0 µl (20 Units)
Sterile dH ₂ O	To 100 µl

The reactions were gently mixed and incubated for 1 h at 37°C. Two µl of the digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 348 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 20 µl of dH₂O. The digestion products were run on a 1% agarose gel, analyzed by ethidium bromide staining and gel extracted as described above. The purified DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 µl tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 µl (1X final)
pcDNA5/TO <i>BamHI/HindIII</i>	1.0 µl (25 ng final)
Ngn3 <i>BamHI/HindIII</i>	1.0 µl (8 ng final)
T4 DNA ligase	0.3 µl (600 U final)
Sterile dH ₂ O variable	To 20 µl

The mixture was incubated at 16°C for 16 h. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 100 µg/ml of ampicillin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)

Mini-prep	5.0 μ l
<i>Bam</i> HI	0.3 μ l (6 Units)
<i>Hind</i> III	0.3 μ l (6 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described above. The purified plasmid DNA was quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies)

The DNA plasmid midi-prep (pcDNA/5TO-Ngn3) was analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 2	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Midi-prep (100 ng/ μ l)	1.0 μ l
<i>Bam</i> HI	0.3 μ l (6 Units)
<i>Hind</i> III	0.3 μ l (6 Units)
Sterile dH ₂ O	To 10 μ l

The reaction mix was gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

3.17.2 Establishing a Double-Stable, Inducible 6TR/5TO-Ngn3-F9 Cell Line

The F9 cell line was cultivated in the appropriate medium. The day before the transfection 8×10^5 cells were plated into three 60 mm cell culture dishes with 4 ml of medium. For the transfection in sterile microcentrifuge tubes 500 μ l of OptiMEM, 20 μ l of Lipofectamine2000 were added and incubated for 5 min at room temperature. Six μ g

of pcDNA6/TR and 1 µg of pcDNA5/TO-Ngn3 DNA were added and the mixtures incubated at room temperature for 15 min. The Lipofectamine2000/DNA complex was then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were Incubate at 37°C for 2 days. After that, the transfected cells were diluted 1:500 and plated in selection medium containing 25 µg/ml of blasticidin and 150 µg/ml of hygromycin for 7-10 days, changing medium as necessary. At least 15 single colonies (6TR/5TO-Ngn3-F9) were obtained.

3.17.3 Doxycycline-induced expression of Ngn3

Approximately 5×10^5 6TR/5TO-Ngn3-F9 cells from the clones obtained above were plated in 6-well tissue culture plates in the appropriate medium and tested for Doxycycline-regulated gene expression by assaying for Ngn3 expression in presence of 0, 2, 4, 6, 8 and 10 µg/ml of doxycycline for at least 48 h.

The cells were then washed twice with PBS and RNA extracted with 500 µl of Trizol reagent as described above. One µg of total RNA was then DNase treated and reverse transcribed as described above.

Expression of Ngn3 has been evaluated by PCR using a specific set of primers reported in Table 1 (Ngn3 primers set).

3.17.4 Differentiation Protocol

Approximately 1×10^7 6TR/5TO-Ngn3-F9 cells from the clones obtained above were plated in 100 mm tissue culture plates in the appropriate medium containing 0.1 nM Retinoic Acid (RA) and 1 mM dibutyryl-c-AMP (db-cAMP) for 72 h. Differentiation ratio was assessed by detecting GATA4 expression by RT-PCR. An aliquot of about 1×10^6 RA-treated cells were washed twice with PBS and RNA extracted with 500 μ l of Trizol reagent as described above. One μ g of total RNA was then DNase treated and reverse transcribed as described above.

Expression of GATA4 has been evaluated by PCR using a specific set of primers reported in Table 1 (GATA4 primers set).

During the differentiation, cell culture medium was changed every 24h. After 72h of differentiation the cell were split and about 1×10^6 RA-treated cells were seeded into 100 mm bacterial-grade Petri dishes and cultured in RPMI 1640 supplemented with 10% fetal bovine serum containing 2 mM L- glutamine, 100 units/ml penicillin and 100 μ l/ml streptomycin at 37°C in an atmosphere of 5% CO₂.

After cells were transferred from suspension culture to bacterial-grade Petri dishes, they failed to adhere and formed small aggregates. Acquisition of the configuration of discrete EBs started as early as 2 days after transfer to suspension culture. After 2 days of culture as EBs the expression of Ngn3 was started by addition of 8 μ g/ml of doxycycline for further 48 h. After 48 h the medium was changed with fresh one without doxycycline and the EBs maintained in culture for further 48h.

At the end the EBs were harvested by centrifugation at 600 x g for 10 min at room temperature. After harvesting the pelleted cells were resuspended in 1 ml of Trizol reagent and total RNA extracted as described above. One μ g of total RNA was then DNase treated and reverse transcribed as described above. The expression of Ngn3 and of other pancreas specific genes such as NeuroD/BETA2, Pax6, Pax4, Nkx2.2, Pdx1 and Isl1 and for the pancreatic hormones insulin and glucagon, have been evaluated by PCR.

3.18 PREPARATION OF AN Ngn3-RheoSwitch® INDUCIBLE SYSTEM

3.18.1 Cloning the mouse Ngn3-CDS into the pNEBR-X1 Hygro Vector

The complete mouse Ngn3-CDS sequence has been amplified by PCR using specific primer's set reported in Table 1 (Ngn3-pNEBR-X1 Hygro) and pBS-Ngn3 as template. The PCR mixture was made as follow:

5X Phusion HF Buffer	10 µl (1X final)
10 mM dNTPs	1.0 µl (0.5 mM)
10 µM 5'-primer	2.0 µl (0.4 µM)
10 µM 3'-primer	2.0 µl (0.4 µM)
Template (100 ng/µl)	1.0 µl
Phusion Hot Start DNA Pol	0.5 µl (0.5 Units)
Sterile dH ₂ O	To 50 µl

Following 30 cycles of amplification, 5µl of PCR product were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 µl of dH₂O. The purified PCR product and the pNEBR-X1-Hygro were digested as follow:

10X NEB Buffer 3	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
Ngn3 PCR	10 μ l (1 μ g final)
<i>NotI</i>	1.0 μ l (10 Units)
<i>Sall</i>	1.0 μ l (10 Units)
Sterile dH ₂ O	To 20 μ l

10X NEB Buffer 3	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
pNEBR-X1 Hygro	10 μ l (10 μ g final)
<i>NotI</i>	1.0 μ l (10 Units)
<i>Sall</i>	1.0 μ l (10 Units)
Sterile dH ₂ O	To 20 μ l

The reactions were gently mixed and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To both the reaction mixtures, 480 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 μ l of dH₂O. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining and extracted using the QIAquick Gel Extraction Kit as described above. The purified DNAs were quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 μ l tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 μ l (1X final)
pNEBR-X1 Hygro <i>NotI/Sall</i>	1.0 μ l (50 ng final)
Ngn3 <i>NotI/Sall</i>	1.0 μ l (14 ng final)
T4 DNA ligase	0.3 μ l (600 Units)
Sterile dH ₂ O	To 10 μ l

The mixture was incubated at 16°C for 16 h. The reaction was placed on ice and the transformation was performed as described above. Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 10 mg/ml of ampicillin, and growth with vigorous shaking for 12-16 h. The plasmid DNA was then extracted as described. The plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 3	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Mini-prep	5.0 μ l
<i>NotI</i>	0.5 μ l (5 units)
<i>Sall</i>	0.5 μ l (5 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. The positive colonies were submitted for sequencing at the NEB Sequencing Core. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before. The purified DNA (pNEBR-X1-Hygro-Ngn3) was quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.18.2 Cloning the mouse Ngn3-CDS into the pNEBR-X1-Puro Vector

The complete mouse Ngn3-CDS sequence has been amplified by PCR using specific primer's set reported in Table 1 (Ngn3-pNEBR-X1-Puro) and pBS-Ngn3 as template. The PCR mixture was made as follow:

5X Phusion HF Buffer	10 μ l (1X final)
10 mM dNTPs	1.0 μ l (0.5 mM)
10 μ M 5'-primer	2.0 μ l (0.4 μ M)
10 μ M 3'-primer	2.0 μ l (0.4 μ M)
Template (100 ng/ μ l)	1.0 μ l
Phusion Hot Start DNA Pol	0.5 μ l (0.5 Units)
Sterile dH ₂ O	To 50 μ l

Following 30 cycles of amplification, 5 μ l of PCR product were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 μ l of dH₂O. The purified PCR product and the pNEBR-X1-Puro were digested as follow:

10X NEB Buffer 3	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
Ngn3 PCR	10 μ l (1 μ g final)
<i>Sall</i>	1.0 μ l (10 Units)

Sterile dH ₂ O	To 20 µl
10X NEB Buffer 3	2.0 µl (1X final)
100X BSA	0.2 µl (1X final)
pNEBR-X1 Puro	10 µl (10 µg final)
<i>Sall</i>	1.0 µl (10 Units)
Sterile dH ₂ O	To 20 µl

The reactions were gently mixed and incubated at 37°C for 1h. To the digestion products 435 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 µl of dH₂O. After that the following digestion reactions were set up:

10X NEB Buffer 4	2.0 µl (1X final)
100X BSA	0.2 µl (1X final)
Ngn3 <i>Sall</i> /	10 µl (1µg final)
<i>Apal</i>	1.0 µl (10 Units)
Sterile dH ₂ O	To 20 µl

10X NEB Buffer 4	2.0 µl (1X final)
100X BSA	0.2 µl (1X final)
pNEBR-X1 Puro <i>Sall</i> /	10 µl (5 µg final)
<i>Sall</i>	1.0 µl (10 Units)
Sterile dH ₂ O	To 20 µl

The reactions were gently mixed and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To both the reaction mixtures, 480 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 µl of dH₂O. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining and extracted using the QIAquick Gel Extraction Kit as described above. The purified DNAs were quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 µl tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 µl (1X final)
pNEBR-X1 Puro <i>Apal/Sall</i>	1.0 µl (50 ng final)
Ngn3 <i>Apal/Sall</i>	1.0 µl (14 ng final)
T4 DNA ligase	0.3 µl (600 Units)
Sterile dH ₂ O	To 20 µl

The mixture was incubated at 16°C for 16h. The reaction was placed on ice and the transformation was performed as described above. Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 10 mg/ml of ampicillin, and growth with vigorous shaking for 12-16 h. The plasmid DNA was then extracted as described. The plasmid mini-preps were analyzed by restriction analysis. In 200µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)

Mini-prep	5.0 μ l
<i>NotI</i>	0.5 μ l (5 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. The positive colonies were submitted for sequencing at the NEB Sequencing Core. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before. The purified DNA (pNEBR-X1-Puro-Ngn3) was quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.18.3 Removal of an aspecific ATG codon from the pNEBR-X1-Hygro-Ngn3 vector

To remove the ATG codon located at 25bp downstream the TATA-Box, the pNEBR-X1-Hygro-Ngn3 vector was digested as follow:

10X NEB Buffer 2	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
pNEBR-X1-Hygro-Ngn3	6.0 μ l (6 μ g final)
<i>NotI</i>	1.0 μ l (10 Units)
<i>SacI</i>	1.0 μ l (10 Units)
Sterile dH ₂ O	To 20 μ l

The reaction was mixed gently and incubated at 37°C for 1 h. Two μ l from the digestion reaction was run on a 1% agarose gel, analyzed by ethidium bromide staining. The restriction enzymes were heat inactivated at 65°C for 10 min and to the remaining 18 μ l of digestion reaction the following components were added:

1 mM dNTPs Mix	0.2 μ l
Blunt Enzyme Mix	1.0 μ l
Sterile dH ₂ O	To 20 μ l

The mixture was incubated at room temperature for 30 min and immediately after that, the enzymes were inactivated by heating at 70°C for 10 min.

For ligation reaction in a 200 μ l tube were mixed:

Blunted DNA	9.0 μ l
2X Quick Ligation Buffer	10 μ l
Quick T4 DNA ligase	1.0 μ l

The reaction was incubated for 5 min at room temperature. The reaction was placed on ice and the transformation was performed as described above. Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 10 mg/ml of ampicillin, and growth with vigorous shaking for 12-16 h. The plasmid DNA was then extracted as described. The colonies were submitted for sequencing at the NEB Sequencing Core. At least one positive colony containing the expected modification was used to set-up a DNA plasmid midi-prep as described before. The purified DNA (pNEBR-X1-Puro-Ngn3/2) was quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.18.4 Cloning the IRES and dTomato sequences into the pNEBR-X1-Hygro-Ngn3/2 Vector

The IRES sequence has been amplified by PCR using specific primer's set reported in Table 1 (IRES) and pNEBR-X1-Klf4-IRES as template. Typically the PCR mixture contained:

5X Phusion HF Buffer	10 μ l (1X final)
----------------------	-----------------------

10 mM dNTPs	1.0 µl (0.5 mM)
10 µM 5'-primer	2.0 µl (0.4 µM)
10 µM 3'-primer	2.0 µl (0.4 µM)
Template (100 ng/µl)	1.0 µl
Phusion Hot Start DNA Pol	0.5 µl (0.5 Units)
Sterile dH ₂ O	To 50 µl

Following 30 cycles of amplification, 5µl of PCR product were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 µl of dH₂O. The purified PCR product and the pNEBR-X1-Puro-Ngn3/2 were digested as follow:

10X NEB Buffer 3	2.0 µl (1X final)
100X BSA	0.2 µl (1X final)
IRES PCR	10 µl (1µg final)
<i>Sall</i>	1.0 µl (10 Units)
<i>NotI</i>	1.0 µl (10 Units)
Sterile dH ₂ O	To 20 µl

10X NEB Buffer 3	2.0 µl (1X final)
100X BSA	0.2 µl (1X final)
pNEBR-X1-Puro-Ngn3/2	10 µl (10 µg final)
<i>Sall</i>	1.0 µl (10 Units)

<i>NotI</i>	1.0 μ l (10 Units)
Sterile dH ₂ O	To 20 μ l

The reactions were gently mixed and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To both the reaction mixtures, 480 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 μ l of dH₂O. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining and extracted using the QIAquick Gel Extraction Kit as described above. The purified DNAs were quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 μ l tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 μ l (1X final)
pNEBR-X1-Hygro-Ngn3/2 <i>NotI/SalI</i>	1.0 μ l (50 ng final)
IRES <i>NotI/SalI</i>	1.0 μ l (12 ng final)
T4 DNA ligase	.3 μ l (600 Units)
Sterile dH ₂ O	To 20 μ l

The mixture was incubated at 16°C for 16h. The reaction was placed on ice and the transformation was performed as described above. Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 10 mg/ml of ampicillin, and growth with vigorous shaking for 12-16 h. The plasmid DNA was then extracted as described. The plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 3	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Mini-prep	5.0 μ l
<i>NotI</i>	0.5 μ l (5 Units)
<i>BstEII</i>	0.5 μ l (2.5 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. The positive colonies were submitted for sequencing at the NEB Sequencing Core. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before. The purified DNA (pNEBR-X1-Hygro-Ngn3/2/IRES) was quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

The dTomato sequence has been amplified by PCR using specific primer's set reported in Table 1 (dTomato) and dTomato CDS as template. Typically the PCR mixture contained:

5X Phusion HF Buffer	10 μ l (1X final)
10 mM dNTPs	1.0 μ l (0.5 mM)
10 μ M 5'-primer	2.0 μ l (0.4 μ M)
10 μ M 3'-primer	2.0 μ l (0.4 μ M)
Template (100 ng/ μ l)	1.0 μ l
Phusion Hot Start DNA Pol	0.5 μ l (0.5 Units)
Sterile dH ₂ O	To 50 μ l

Following 30 cycles of amplification, 5 μ l of PCR product were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1

minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 μ l of dH₂O. The purified PCR product was used to perform a pNEBR-X1-Hygro-Ngn3/2/IRES cloning. The dTomato sequence and the pNEBR-X1-Hygro-Ngn3/2/IRES were digested as follow:

10X NEB Buffer 3	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
dTomato PCR	10 μ l (1 μ g final)
<i>Sall</i>	1.0 μ l (10 Units)
Sterile dH ₂ O	To 20 μ l
10X NEB Buffer 3	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
pNEBR-X1-Hygro-Ngn3/2/IRES	10 μ l (10 μ g final)
<i>Sall</i>	1.0 μ l (10 Units)
Sterile dH ₂ O	To 20 μ l

The reactions were gently mixed and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. The pNEBR-X1-Hygro-Ngn3/2/IRES *Sall* plasmid has been dephosphorylated by adding 1 μ l of Calf Intestinal Phosphatase (CIP) and incubating the reaction for 1 h at 37°C. To both the reaction mixtures, 480 μ l of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and

dissolved in 10 μ l of dH₂O. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining and extracted using the QIAquick Gel Extraction Kit as described above. The purified DNAs were quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 μ l tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 μ l (1X final)
pNEBR-X1-Hygro-Ngn3/2/IRES <i>Sall</i> /	1.0 μ l (50 ng final)
dTomato <i>Sall</i> /	1.0 μ l (12 ng final)
T4 DNA ligase	0.3 μ l (600 Units)
Sterile dH ₂ O	To 20 μ l

The mixture was incubated at 16°C for 16 h. The reaction was placed on ice and the transformation was performed as described above. Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 10 mg/ml of ampicillin, and growth with vigorous shaking for 12-16 h. The plasmid DNA was then extracted as described. The plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 3	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Mini-prep	5.0 μ l
<i>Sall</i>	0.5 μ l (5 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated at 37°C for 1 h. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. The positive colonies were submitted for sequencing at the NEB Sequencing Core. At least one positive colony containing the expected fragment size was used to set-up a DNA

plasmid midi-prep as described before. The purified DNA (pNEBR-X1-Hygro-Ngn3/2/IRES/dTomato) was quantified by using a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.18.5 Establishing of Inducible RheoSwitch® F9 Cell Line

The F9 cell line was cultivated in the appropriate medium as described above. The day before the transfection 8×10^5 cells were plated into three 60 mm cell culture dishes with 4 ml of medium. For the transfection in sterile microcentrifuge tubes 500 μ l of OptiMEM, 20 μ l of Lipofectamine 2000 were added and incubated for 5 min at room temperature. Six μ g of pNEBR-R1 Regulator plasmid were added and the mixture incubated at room temperature for 15 min. The Lipofectamine 2000/DNA complex was then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were incubate at 37°C for 2 days. After that, the transfected cells were diluted 1:500 and plated in selection medium containing 300 μ g/ml of G418 15-20 days, changing medium as necessary. At least 17 single colonies (F9-R1) were obtained.

3.18.6 Clones Selection

F9-R1 clones were maintained in culture in selection medium containing 300 μ g/ml G418. The day before transfection 4×10^5 cells were plated into three wells of a 6-well cell culture plate with 2 ml of medium. For the transfection in sterile microcentrifuge tubes 250 μ l of OptiMEM and 10 μ l of Lipofectamine2000 were added and incubated for 5 min at room temperature. Two μ g of pNEBR-GLuc and 2 μ g of pCMV- β Gal plasmid were added and the mixture and incubated at room temperature for 15 min. The Lipofectamine2000/DNA complex was then added to the cells in a drop-wise manner and

the cells returned to the incubator. After 6-8 h the transfection reaction was replaced with medium containing 0.5 μ M of RSL-1 and incubated at 37°C for 12 h.

3.18.7 *Gaussia* Luciferase Assay

For the *Gaussia* Luciferase assay 20 μ l of supernatant from the above transfection were collected in a 96-well black plate. 1X *Gaussia* Luciferase assay working solution was prepared by adding 50 μ l of GLuc Substrate to 5 ml of the GLuc Assay Buffer immediately before the assay.

Fifty μ l of 1X GLuc buffer were added to each sample and immediately the plate was read in a luminometer programmed to perform a 2-second measurement delay followed by a 10-second measurement read.

The cells transfected as above were washed twice with PBS and lysed with 250 μ l of 1X Passive Lysis Buffer. The plates were then rocked several times to ensure complete coverage of the cells with lysis buffer. The cells were then scraped and all the liquid transferred to a microcentrifuge tube, vortexed and centrifuged at 12,000 \times g for 10 min at 4°C.

For the assay β -Galactosidase, 20 μ l of cell lysates were aliquoted in a 96-well black plate. 100X Galacton-Plus Reagent were diluted down to 1X in Reaction Buffer Diluent. Seventy μ l of 1X Galacton Plus Reagent were added to each sample and the reaction was incubated at room temperature for 30 sec. After that, 90 μ l of Accelerator II Reagent were added to each sample. Immediately the plate was read in a luminometer programmed to perform a 2-second measurement delay followed by a 10-second measurement read. The ratio of the *Gaussia*/b-galactosidase activity represents the normalized reporter gene expression.

The fold induction was calculated as follow:

$$\text{Fold-induction} = +\text{RSL1-RLU}/-\text{RSL1-RLU}$$

The clones showing the highest fold-induction (highest expression with lowest background) were propagated. About 5 clones (F9-R1) were obtained showing about a 10-fold induction.

3.18.8 Establishing of Inducible RheoSwitch®F9-R1-Ngn3 Cell Line

At least one of the F9 cell line showing the highest fold-induction was cultivated in the appropriate medium supplemented with 200 µg/ml G418. The day before the transfection 8×10^5 cells were plated into three 60 mm cell culture dishes with 4 ml of medium. For the transfection in sterile microcentrifuge tubes 500 µl of OptiMEM, 20 µl of Lipofectamine2000 were added and incubated for 5 min at room temperature. Six µg of pNEBR-X1-Hygro-Ngn3/2 plasmid were added and the mixture incubated at room temperature for 15 min. The Lipofectamine2000/DNA complex was then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were incubate at 37°C for 2 days. After that, the transfected cells were diluted 1:500 and plated in selection medium containing 300 µg/ml of G418 and 150 µg/ml of hygromycin for 15-20 days, changing medium as necessary. At least 7 single colonies (F9-R1/Ngn3) were obtained.

3.18.9 Transfection and Activation of Stable Cell Line NIH-3T3-47 (general protocol)

The NIH-3T3-47 cell line was cultivated in the appropriate medium supplemented with 800 µg/ml of G418. The day before the transfection the cells were plated into 12 well cell culture plate with 0.5 ml of medium. For the transfection, in sterile microcentrifuge tubes 125 µl of DMEM, 2 µl of TransPassD2 were mixed and incubated for 15 minutes at room temperature. At the same time 0.5 µg of pNEBR-X1-Ngn3 Hygro

(or pNEBR-GLuc) and 0.5 μg of control plasmid pCMV- β -Galactosidase were mixed onto sterile microcentrifuge tubes and incubated at room temperature for 15 minutes. The two mixtures were then mixed and incubated for additional 30 min at room temperature. The TransPassD2/DNA complex was then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were incubate at 37°C for 6 h. After that, 250 μl of medium containing 0.5 μM RSL1 activator were added to the transfected cells and mandained in culture for 12 h before assay for protein expression.

3.18.10 Western-Blot Analysis

The F9-R1 and NIH/3T3-47 were washed twice with PBS and lysed using 100 μl of Passive Lysis Buffer. The cell extracts were then transferred into microcentrifuge tubes and centrifuged for 5 min at 14,000 $\times g$ at 4°C. The supernatants were transferred to new tubes and used for protein quantification as described above.

The protein samples were added with 1/10 of Sample Buffer and boiled for 5 min and 10 μg of proteins were separated onto 10-20% Tris-Glycine gel. The gels were run at 120V (constant voltage) until the bromophenol blue dye is just off in presence of the 1X Electrode Buffer.

After run, the gels were recovered and soaked for 15 min at room temperature in transfer buffer along with 2 sponges, 4 sheets (7.5 cm X 10 cm) of 3MM Whatman paper, and 1 PVDF membrane (7.5 cm X 10 cm).

The transfer sandwiches were then set-up adding, in the order, from the negative side of the sandwich: one sponge, two 3MM Whatman papers, the gel, the PVDF membrane, two more 3MM Whatman papers, the last sponge. The sandwiches are then put into the tank apparatus filled with the tank buffer and blotted for 1 hr at 100V (constant voltage). After transfer the PVDF membranes were incubated in Blotto solution for 1 hr. The membranes were then incubated with rabbit anti-HA-tag (1:1000) and goat anti- β -Actin (1:5000), respectively, in Blotto solution overnight at 4°C with gentle shaking.

The day after the membranes were washed for 10 min three times with TBST and then incubated with goat anti-rabbit-IgG HRP-conjugated (1:1000) and with rabbit anti-goat-IgG HRP-conjugated (1:1000), respectively, in Blotto solution for 1 hr at room temperature with gentle shaking. The membranes were then washed for 15 min four times with TBST. The membranes were then developed for 1 min with Western Lightning Chemiluminescence Reagent and autoradiographed with 5 min exposure.

3.19 Preparation of the AdNgn3-EGFP AND AdNeuroD/BETA2 viral particles

3.19.1 Cloning the mouse Ngn3 CDS into the pShuttle-IRES-hrEGFP2 vector

The complete mouse Ngn3 and the complete hamster NeuroD/BETA2 CDS sequences have been amplified by PCR using specific set of primers reported in Table 1 (AdNgn3 and AdNeuroD primers sets) and pBS-Ngn3 and pcDNA-NeuroD/BETA2 as templates, respectively.

Following 28 cycles of amplification, 5 μ l of PCR products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixtures 445 μ l of TE buffer were added and the solutions extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phases were collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solutions were then centrifuged at 10,000 x g for 20 min. The pellets further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellets were air dried and dissolved in 10 μ l of dH₂O. The purified PCR fragments were used to perform TOPO cloning reactions using the TOPO TA cloning kit. In two separate 200 μ l tube were mixed:

Purified PCRs	4.0 μ l
Salt Solution	1.0 μ l
pCR2.1-TOPO vector	1.0 μ l

The reactions were mixed gently and incubated for 5 min at room temperature. The reactions were placed on ice and the transformations were performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 50 μ g/ml of kanamycin, and grown with vigorous shaking for 12–16 h. The plasmid DNAs were then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 3	1.0 μ l
100X BSA	0.1 μ l
Mini-prep	5.0 μ l
<i>ScaI</i>	0.3 μ l
<i>XhoI</i>	0.3 μ l
Sterile dH ₂ O	To 10 μ l

The reactions was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before.

The DNA plasmid midi-preps (pCR2.1-TOPO-Ngn3 and pCR2.1-TOPO-NeuroD) were analyzed by restriction analysis. In two 200 μ l tubes were mixed:

10 X NEB Buffer 3	1.0 μ l
100X BSA	0.1 μ l

Midi-prep (100 ng/μl)	1.0 μl
<i>ScaI</i>	0.3 μl
<i>XhoI</i>	0.3 μl
Sterile dH ₂ O	To 10 μl

The reaction mixtures were gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

The Ng3 CDS and the NeuroD/BETA2 from the pCR2.1-TOPO-Ngn3 and the pCR2.1-TOPO-NeuroD vectors were subcloned into the pShuttle-IRES-hrEGFP2 by restriction digestion with *ScaI* and *XhoI* as follow. In three 1.5 ml microcentrifuge tubes were added respectively:

10X NEB Buffer 3	10 μl (1X final)
100X BSA	1.0 μl (1X final)
pShuttle-IRES-hrEGFP2	20 μl (15 μg final)
<i>ScaI</i>	1.5 μl (15 Units)
<i>XhoI</i>	1.0 μl (20 Units)
Sterile dH ₂ O	To 100 μl

10X NEB Buffer 3	10 μl (1X final)
100X BSA	1.0 μl (1X final)
pCR2.1-TOPO-Ngn3	20 μl (15 μg final)
<i>ScaI</i>	1.5 μl (15 Units)
<i>XhoI</i>	1.0 μl (20 Units)
Sterile dH ₂ O	To 100 μl

10X NEB Buffer 3	10 µl (1X final)
100X BSA	1.0 µl (1X final)
pCR2.1-TOPO-NeuroD	20 µl (15 µg final)
<i>Scal</i>	1.5 µl (15 Units)
<i>XhoI</i>	1.0 µl (20 Units)
Sterile dH ₂ O	To 100 µl

The reactions were gently mixed and incubated for 1 h at 37°C. Two µl of the digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the digestions mixture 348 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solutions were then centrifuged at 10,000 x g for 20 min. The pellets further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellets were air dried and dissolved in 20 µl of dH₂O. The digestion products were run on a 1% agarose gel, analyzed by ethidium bromide staining and gel extracted as described above. The purified DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In two 200 µl tube were mixed, respectively:

10X T4 DNA ligase Buffer	2.0 µl (1X final)
pShuttle-IRES-hrEGFP2 <i>Scal/XhoI</i>	1.0 µl (25 ng final)
Ngn3 (or NeuroD) <i>Scal/XhoI</i>	1.0 µl (5 ng final)
T4 DNA ligase	0.3 µl (600 U final)
Sterile dH ₂ O variable	To 20 µl

10X T4 DNA ligase Buffer	2.0 µl (1X final)
pShuttle-IRES-hrEGFP2 <i>Scal/XhoI</i>	1.0 µl (25 ng final)

NeuroD <i>Scal/XhoI</i>	1.0 μ l (5 ng final)
T4 DNA ligase	0.3 μ l (600 U final)
Sterile dH ₂ O variable	To 20 μ l

The mixtures were incubated at 16°C for 16 h. The reactions were placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 50 μ g/ml of kanamycin, and grown with vigorous shaking for 12–16 h. The plasmid DNAs were then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 3	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Mini-prep	5.0 μ l
<i>Scal</i>	0.3 μ l (3 Units)
<i>XhoI</i>	0.3 μ l (6 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described above. The purified plasmid DNAs (pShuttle-Ngn3-IRES-hrEGFP2 and pShuttle-NeuroD-IRES-hrEGFP2) were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The DNA plasmid midi-preps were analyzed by restriction analysis. In two 200 μ l tubes were mixed:

10 X NEB Buffer 3	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)

Midi-preps	1.0 μ l (100 ng)
<i>ScaI</i>	0.3 μ l (3 Units)
<i>XhoI</i>	0.3 μ l (6 Units)
Sterile dH ₂ O	To 10 μ l

The reaction mixtures were gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

Both the pShuttle-Ngn3-IRES-hrEGFP2 and pShuttle-NeuroD-IRES-hrEGFP2 vectors were then linearized. In two 200 μ l tubes were mixed respectively:

10 X NEB Buffer 4	5.0 μ l (1X final)
100X BSA	0.5 μ l (1X final)
pShuttle-Ngn3-IRES-hrEGFP2	10 μ l (10 μ g final)
<i>PmeI</i>	1 μ l (10 Units)
Sterile dH ₂ O	To 50 μ l

10 X NEB Buffer 4	5.0 μ l (1X final)
100X BSA	0.5 μ l (1X final)
pShuttle-NeuroD-IRES-hrEGFP2	10 μ l (10 μ g final)
<i>PmeI</i>	1 μ l (10 Units)
Sterile dH ₂ O	To 50 μ l

The reactions mixtures were gently and incubated for 1 h at 37°C. Five μ l of the digestion product was run on a 1% agarose gel and analyzed by ethidium bromide staining. The mixtures were then added with 5 Units of Alkaline Calf Intestinal Phosphatase and incubated for 1 h at 37°C. To the digestion mixtures 445 μ l of TE buffer were added and the solutions extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phases were collected, added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solutions were then centrifuged at 10,000 x g for 20 min. The

pellets further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellets were air dried and dissolved in 20 μ l of dH₂O. The digestion products were run on a 1% agarose gel, analyzed by ethidium bromide staining and gel extracted as described above.

The purified *PmeI* linearized plasmid DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies) and diluted to a final concentration of 1 μ g/ μ l.

3.19.2 Transformation of the BJ5183Ad Cells to Produce Recombinant AdPlasmid

The BJ5183Ad cells were transformed with the linearized pShuttle-Ngn3-IRES-hrEGFP2 and pShuttle-NeuroD-IRES-hrEGFP2 vectors. A recombination event that takes place in the bacterial cells results in the production of recombinant AdEasy plasmid DNA. The transformation was carried out by electroporation of the competent BJ5183Ad cells. Just prior the transformation the BJ5183Ad cells are removed from -80°C storage and thawed on ice. 1 μ l (1 μ g) of linearized, dephosphorylated pShuttle-Ngn3-IRES-hrEGFP2 and pShuttle-NeuroD-IRES-hrEGFP2 vectors were added to the chilled aliquot of BJ5183Ad cells. The cells were mixed by tapping the tube gently and kept on ice. The electroporator was set to the appropriate setting (200 Ω , 2.5 kV, 25 μ F) and the DNA/cells mixtures transferred in chilled electroporation cuvettes. The samples were then pulsed and immediately added with 1 ml of LB broth. The cell suspensions were transferred to sterile 14-ml BD Falcon polypropylene round-bottom tubes and incubated at 37°C for 1 hour while shaking at 225–250 rpm. The entire volume (50 μ l, 100 μ l, 250 μ l, and 600 μ l of cell suspension, respectively) of recovered cells was plated onto 4 LB agar plates containing 50 μ g/ml of kanamycin and incubated overnight at 37°C .

Single colonies from the transformed bacteria plates obtained as above were inoculated into 5 ml of LB medium containing 50 μ g/ml of kanamycin, and grown with vigorous shaking for 12–16 h. 2 ml of the bacterial cells were subjected to Mini-Prep protocol.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 1	2.0 μ l (1X final)
100X BSA	0.2 μ l (1X final)
Mini-preps	10 μ l
<i>PacI</i>	0.5 μ l (5 Units)
Sterile dH ₂ O	To 20 μ l

The reaction was mixed gently and incubated for 1 h at 37°C. The entire digestion products were run on a 0.8% agarose gel and analyzed by ethidium bromide staining. At least five positive colonies were selected for the subsequent amplification. The five selected purified recombinant Ad plasmids were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

For bacteria transformation, to 100 μ l aliquot of XL10-Gold ultracompetent cells, thawed on ice in a sterile 14-ml BD Falcon polypropylene round-bottom tube, 4 μ l of β -mercaptoethanol were added and the mixture incubated on ice for 10 min. The cells were then added 50 ng of recombinant Ad plasmids obtained as above and incubated on ice for 30 min. The tubes were heat-pulsed in a 42°C water bath for 30 seconds and placed on ice for 2 min. After addition of 900 μ l of prewarmed (42°C) NZY+ into the mixture the cells were incubated at 37°C for 60 min with vigorous shake (250 rpm). Five μ l, 25 μ l, and 100 μ l of each transformation reaction were plated on LB agar plates containing 50 μ g/ml of kanamycin and incubated overnight at 37°C. At least five colonies from transformations were picked and transferred into 10 ml LB broth containing 50 μ g/ml of kanamycin and incubated overnight at 37°C while shaking at 225-250 rpm. Five ml of the above overnight cultures were inoculated in 500 ml of LB broth containing 50 μ g/ml of kanamycin and incubated overnight at 37°C while shaking at 225-250 rpm and subjected to a plasmid Maxi-Prep protocol. The purified plasmid DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The AdNgn3 and AdNeuroD/BETA2 recombinant plasmids were then linearized. In 200 μ l tubes were mixed, respectively:

10X NEB Buffer 1	5.0 μ l (1X final)
100X BSA	0.5 μ l (1X final)
AdNgn3	20 μ l (25 μ g)
<i>PacI</i>	2.5 μ l (25 Units)
Sterile dH ₂ O	To 50 μ l
10X NEB Buffer 1	5.0 μ l (1X final)
100X BSA	0.5 μ l (1X final)
AdNeuroD/BETA2	20 μ l (25 μ g)
<i>PacI</i>	2.5 μ l (25 Units)
Sterile dH ₂ O	To 50 μ l

The reaction mixtures were gently and incubated for 1 h at 37°C. One μ l of the digestion products was run on a 0.8% agarose gel and analyzed by ethidium bromide staining to confirm the desired restriction pattern.

To the digestion mixtures 449 μ l of TE buffer were added and the solutions extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phases were collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solutions were then centrifuged at 10,000 x g for 20 min. The pellets further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellets were air dried and dissolved in 50 μ l of dH₂O.

3.19.3 Preparation of Primary Adenovirus Stock

The 293Ad cell line was cultivated in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose and 110 mg/L sodium pyruvate and 4 mM L-glutamine, supplemented with 10% (v/v) heat-inactivated fetal bovine serum. The day before the transfection, 8×10^5 cells were plated into two 60 mm cell culture dishes with

4 ml of medium. For the transfection in two sterile microcentrifuge tubes 200 μ l of DMEM (containing 4.5 g/L glucose and 110 mg/L sodium pyruvate and 4 mM L-glutamine), 6 μ l of FuGENE 6 were added and incubated for 5 min at room temperature. Two μ g of *PacI* digested AdNgn3 and AdNeuroD/BETA2 DNAs were added, respectively, and the mixtures incubated at room temperature for 15 min. The FuGENE6/DNA complexes were then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were incubated at 37°C for 7–10 days, replenishing the growth medium when needed. To prepare the primary viral stocks the growth media were removed from the adenovirus-producing 293Ad plates and the cells washed once with PBS. To each plate 500 μ l of PBS were added and the cells collected by scraping. The cell suspensions were then transferred into a sterile microcentrifuge tube and subjected to four round of freeze/thaw by alternating the tubes between a dry ice-methanol bath and a 37°C water bath, vortexing briefly after each thaw. The cell debris were collected by centrifugation at 12,000 x g for 10 min and the supernatants (primary virus stocks) were transferred in new microcentrifuge tubes and stored at -80°C.

3.19.4 Amplification of the Primary Viral Stock

Amplification of the primary viral stocks has been obtained by infection of 293Ad cells. 3.5×10^6 293Ad were plated into a 100 mm culture dish in 5 ml of complete medium containing 200 μ l of primary viral stocks. The cells were incubated for 2 h and then supplied with additional growth medium. The cells were then maintained in culture until appearance of the first signs of cytopathic effect (CPE) (cells will round up and detach from the plates, and the nucleus will occupy a major part of the cell). At that point the Adenoviral particles were collected by four round of freeze/thaw as described above.

3.19.5 Plaque Assay

The plaque assay is used to determine the titer (pfu/ml) of the viral stocks. 5×10^5 293Ad cells were plated in two 6-well tissue culture plates overnight at 37°C. The adenoviral stocks obtained as above were diluted in 1 ml volumes over a 10-fold series from 10^{-5} to 10^{-9} in growth medium. Each dilution were then added to a separate well of the 6-well plates and incubated at 37°C for 2 h. The wells were overlaid with agarose as follow; a sterile solution of 5% agarose in dH₂O was prepared by autoclaving and left to equilibrate at 65°C. 30 ml of growth medium equilibrated to 37°C were added and 3 ml of this agarose/growth medium mix was pipetted along the side of the well previously emptied of medium. The plates were then incubated at 37°C for about 10 days or until the appearance of hrEGFP expressing colonies. The titer was determined by counting the number of hrEGFP expressing colonies obtained with the highest viral dilution and multiplying that number by the dilution factor to estimate pfu/ml.

3.20 CHROMATIN IMMUNOPRECIPITATION CLONING (ChiP-cloning)

3.20.1 Cell culture and infection.

Approximately 2×10^7 AR42J and INS-1 cells were plated in 100 mm cell culture dishes in the appropriate medium and infected with 20 MOI of AdNgn3-EGFP and Ad-EGFP, respectively, for 4 hr and then the medium was changed with fresh one. The cells were maintained in culture for two days.

3.20.2 *In Vivo* Crosslinking and Lysis

For the ChiP-cloning experiments a modification of the EZ-ChiP kit has been used. To the cell plates prepared as above 540 μ l of fresh 18.5% formaldehyde (final concentration is 1%) prepared dissolving 925 mg of paraformaldehyde into 4.8 ml of dH₂O in presence of 35 μ l of 1N KOH were added. The plates were gently swirled and incubated at room temperature for 10 min. After that 1 ml of 10X Glycine to each dish was added to quench unreacted formaldehyde and the plates incubated at room temperature for 5 min. The plates were placed on ice and the medium aspirated. The cells were washed twice with 10 ml of cold PBS. 1 ml cold PBS containing Protease Inhibitor Cocktail was added to the dish. The cells were then scraped into a microfuge tube and spun at 700 x g at 4°C for 2-5 min to pellet cells and the supernatant discarded. 1 ml of SDS Lysis Buffer containing Protease Inhibitor Cocktail was then added, the cell pellet resuspended and aliquoted between 300-400 μ l per microcentrifuge tube. Cells in SDS Lysis Buffer at a cell concentration of 2×10^7 per ml were sheared with 4-5 sets of 10-second pulses on wet ice using a 60 SONIC Dismembrator (Fisher Scientific) equipped with a 2 mm tip and set to 30% of maximum power. The sheared cells were then spun down at a minimum of 10,000 x g at 4°C for 10 min to remove insoluble material and the supernatant collected in fresh microcentrifuge tubes in 100 μ l aliquots. Each 100 μ l aliquot contains $\sim 2 \times 10^6$ cell equivalents of chromatin.

3.20.3 Immunoprecipitation (IP) of Crosslinked Protein/DNA

To at least 2 x 100 μ l aliquots (experiment + negative control) 900 μ l of Dilution Buffer containing Protease Inhibitor Cocktail were added. Sixty μ l of Protein G magnetic beads for each IP were then added and the tubes incubated for 1 hour at 4°C with rotation. The cleared supernatants were separated from the Protein G magnetic beads using a magnetic separation rack. Ten μ l (1%) of the supernatant were collected as Input

and to the remaining supernatant fractions 10 μg of rabbit anti-Ngn3 and 10 μg of normal rabbit IgG were added to the “experiment” and “negative control” tubes, respectively. The tubes were incubated overnight at 4°C with rotation. To recover the antibody/antigen/DNA complexes 60 μl of Protein G magnetic beads were added and the tubes incubated for 1 hour at 4°C with rotation. The magnetic beads were recovered using a magnetic separation rack as above and the beads were washed as follows by resuspending the beads in 1 ml each of the cold buffers in the order listed below and incubating for 5 min on a rotating platform followed by recovery of the supernatant using a magnetic separation rack:

- a. Low Salt Immune Complex Wash Buffer, **one wash.**
- b. High Salt Immune Complex Wash Buffer, **one wash.**
- c. LiCl Immune Complex Wash Buffer, **one wash.**
- d. TE Buffer, **two washes.**

3.20.4 Elution of Protein/DNA Complexes

To each tubes (IPs and inputs) obtained as described above 100 μl of Elution Buffer were added mixed by flicking and incubated at room temperature for 15 min. The magnetic beads were pelleted using the magnetic separation rack and the supernatant collected. The steps were repeated with an extra 100 μl of Elution Buffer and the eluates combined.

3.20.5 Reverse Crosslinks of Protein/DNA Complexes to Free DNA

To all tubes (IPs and Inputs) were added 8 μl 5 M NaCl and incubated at 65°C overnight to reverse the DNA-Protein crosslinks. To all tubes, 1 μl of RNase A were

added and incubated for 30 min at 37°C. Four μl 0.5 M EDTA, 8 μl 1M Tris-HCl and 1 μl Proteinase K were added and incubated at 45°C for 2 h.

3.20.6 DNA Purification Using Spin Columns

One ml of Bind Reagent A was added to each 200 μl DNA sample tube (IPs and Inputs), and the sample/Bind Reagent A. mixture transferred to the Spin Filter in Collection Tube and centrifuged for 30 sec at a minimum of 10,000 x g. Five hundred μl of the Wash Reagent B were then added to the Spin Filter in Collection Tube and centrifuged for 30 seconds at a minimum of 10,000 x g. The Spin Filter was put into a clean Collection Tube, added with 50 μl of Elution Buffer C and centrifuged for 30 sec at a minimum of 10,000 x g.

3.20.7 Blunting reaction, ligation and bacteria transformation.

For blunting reaction in a sterile 200 μl tube were mixed:

Purified DNA	19 μl
10X Blunting Buffer	2.5 μl
1 mM dNTP Mix	2.5 μl
Blunting Enzyme Mix	1.0 μl
Sterile dH ₂ O variable	To 25 μl

The mixtures were incubated at room temperature for 30 min. Immediately after that the enzymes were inactivated by heating at 70°C for 10 min.

For ligation reaction a sterile 200 μl tube were mixed:

Blunted DNA	10 µl
10X T4 DNA ligase Buffer	2.0 µl
pCR-BLUNT vector	1.0 µl
T4 DNA ligase	1.0 µl
Sterile dH ₂ O variable	To 20 µl

The mixtures were incubated at 16°C for 16 h. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 50 µg/ml of kanamycin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The purified plasmid DNA was quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.20.8 Bioinformatics analysis.

About 400 ng of each mini-prep DNA samples were sequenced using the M13 forward (-20) (5'-GTAAAACGACGGCCAG-3') and the M13 reverse (5'-CAGGAAACAGCTATGAC-3') primers at the Brigham and Women's DNA Core Sequencing Facility (Boston, MA, USA).

The sequences not belonging to the vector were used to perform BLAST analyses using the *blastn* that compares a nucleotide query sequence against a nucleotide sequence database. The rat genome database (<http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=10116>) has been analyzed using the default filters setting. The output results were analyzed for sequences that span the 5'term in the proximity of transcription factors and that contain at least one DNA binding motif for the bHLH family of transcription factors (E-box, CANNTG).

3.20.9 Confirmation of the results

The sequences showing the presence of E-box(es) and localized at the 5'-term of transcription factors were confirmed by independent classical ChiP experiments. Approximately 2×10^7 AR42J cells were plated in three 100 mm cell culture dishes in the appropriate medium and infected with 20 MOI of AdNgn3-EGFP, 20 MOI of AdNeuroD and 20 MOI of AdEGFP, respectively, for 4 hr and then the medium was changed with fresh one. The cells were maintained in culture for two days. The DNAs and protein complexes were crosslinked and immunoprecipitated using anti-Ngn3 and anti-NeuroD/BETA2 antibodies, respectively as described above. The purified DNA was analyzed by PCR using clone specific primer(s).

Following 35 cycles of amplification the PCR products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

3.21 PREPARATION OF THE OVO11-PROM1-LUCIFERASE REPORTER SYSTEM AND TRANSACTIVATION EXPERIMENTS

3.21.1 Cloning of the Rat 1.1 kb OVO11 Promoter into the pGL2-Basic vector

The 1.1 kb sequence encompassing the 5'-flanking region from -841 to +230 of the rat *ovoll* gene and containing the E-box 1 has been amplified by PCR using a specific set of primers reported in Table 1 (OVO11Prom1 primers set) and rat genomic DNA as template.

Following 30 cycles of amplification, 5 μ l of PCR product was run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 445 μ l of

TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volumes of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 μ l of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 10 μ l of dH₂O. The purified PCR fragment was used to perform a TOPO cloning reaction using the TOPO TA cloning kit. In a 200 μ l tube were mixed:

Purified PCR	4.0 μ l
Salt Solution	1.0 μ l
pCR2.1-TOPO vector	1.0 μ l

The reaction was mixed gently and incubated for 5 min at room temperature. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 50 μ g/ml of kanamycin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 2	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Mini-prep	5.0 μ l
<i>Bam</i> HI	0.3 μ l (6 Units)
<i>Sac</i> I	0.3 μ l (6 Units)
Sterile dH ₂ O	To 10 μ l

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At

least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described before.

The DNA plasmid midi-prep (pCR2.1-TOPO-OVO11Prom1) was analyzed by restriction analysis. In 200 μ l tubes were mixed:

10X NEB Buffer 2	1.0 μ l (1X final)
100X BSA	0.1 μ l (1X final)
Midi-prep (100 ng/ μ l)	1.0 μ l
<i>Bam</i> HI	0.3 μ l (6 Units)
<i>Sac</i> I	0.3 μ l (6 Units)
Sterile dH ₂ O	To 10 μ l

The reaction mix was gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

The OVO11Prom1 from the pCR2.1-TOPO-OVO11Prom1 vector was subcloned into the pGL2-Basic vector by restriction digestion with *Bam*HI and *Sac*I as follow. In two 1.5 ml microcentrifuge tubes were added respectively:

10X NEB Buffer 2	10 μ l (1X final)
100X BSA	1.0 μ l (1X final)
pGL2	20 μ l (30 μ g final)
<i>Bam</i> HI	1.5 μ l (30 Units)
<i>Sac</i> I	1.5 μ l (30 Units)
Sterile dH ₂ O	To 100 μ l

10X NEB Buffer 2	10 μ l (1X final)
100X BSA	1.0 μ l (1X final)
pCR2.1-TOPO-OVO11Prom1	20 μ l (20 μ g final)
<i>Bam</i> HI	1.0 μ l (20 Units)
<i>Sac</i> I	1.0 μ l (20 Units)

Sterile dH₂O

To 100 µl

The reactions were gently mixed and incubated for 1 h at 37°C. Two µl of the digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. To the PCR mixture 348 µl of TE buffer were added and the solution extracted with 1 volume of phenol:chloroform:isoamyl alcohol (24:25:1). After centrifugation at 10,000 x g for 1 minute the upper aqueous phase was collected added with 1/10 volume of 3 M Sodium Acetate pH 5.2 and 0.7 volume of isopropyl alcohol. The solution was then centrifuged at 10,000 x g for 20 min. The pellet further washed with 500 µl of 70% ethanol and centrifuged once more at 10,000 x g for 20 min. The pellet was air dried and dissolved in 20 µl of dH₂O. The digestion products were run on a 1% agarose gel, analyzed by ethidium bromide staining and gel extracted as described above. The purified DNAs were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

The ligation reaction was set up at a 3:1 insert:vector ratio. In a 200 µl tube were mixed:

10X NEB T4 DNA ligase Buffer	2.0 µl (1X final)
pGL2 <i>BamHI/SacI</i>	1.0 µl (25 ng final)
OV011Prom1 <i>BamHI/SacI</i>	1.0 µl (15 ng final)
T4 DNA ligase	0.3 µl (600 U final)
Sterile dH ₂ O	To 20 µl

The mixture was incubated at 16°C for 16hr. The reaction was placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 100 µg/ml of ampicillin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described.

The DNA plasmid mini-preps were analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Mini-prep	5.0 µl
<i>BamHI</i>	0.3 µl (6 Units)
<i>SacI</i>	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction was mixed gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining. At least one positive colony containing the expected fragment size was used to set-up a DNA plasmid midi-prep as described above. The purified plasmid DNA was quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies)

The DNA plasmid midi-prep (pGL2-OVO11Prom1) was analyzed by restriction analysis. In 200 µl tubes were mixed:

10X NEB Buffer 2	1.0 µl (1X final)
100X BSA	0.1 µl (1X final)
Midi-prep (100 ng/µl)	1.0 µl
<i>BamHI</i>	0.3 µl (6 Units)
<i>SacI</i>	0.3 µl (6 Units)
Sterile dH ₂ O	To 10 µl

The reaction mix was gently and incubated for 1 h at 37°C. The digestion products were run on a 1% agarose gel and analyzed by ethidium bromide staining.

3.21.2 Site-Directed Mutagenesis

To mutate the E-box 1 in the OVO11Prom1 the Phusion Site-Directed Mutagenesis Kit has been used. This kit uses the highly processive Phusion™ Hot Start

High-Fidelity DNA Polymerase for exponential PCR amplification of dsDNA plasmid to be mutated. The mutagenesis protocol comprises only three steps:

1. PCR amplification of target plasmid with two phosphorylated primers.
2. Circularization of mutated PCR products.
3. Transformation to *E. coli*.

The point mutation for E-Box 1 has been created by designing a mismatch in the mutagenic primer (Table 1, EboxMut primer set). The desired mutation has been collocated in the middle of the primer with 10-15 perfectly matched nucleotides on each side. As Template the pGL2-OVO11Prom1 has been used. The amount of plasmid template in the mutagenesis reaction has been titrated as down as possible to prevent background transformants after ligation and transformation. Typically the PCR mixture contained:

5X Phusion HF Buffer	10 μ l (1X final)
10 mM dNTP	1.0 μ l (0.5 mM)
10 μ M 5'-Primer	2.5 μ l (0.5 μ M)
10 μ M 3'-Primer	2.5 μ l (0.5 μ M)
Template (50 pg/ μ l)	1.0 μ l
Phusion Hot Start DNA Pol	0.5 μ l (0.5 Units)
Sterile dH ₂ O	To 50 μ l

After PCR the product was circularized with Quick T4 DNA Ligase. For ligation reaction in a 200 μ l tube were mixed:

PCR	5.0 μ l
2X Quick Ligation Buffer	5.0 μ l
Quick T4 DNA ligase	0.2 μ l

The reaction was incubated at room temperature for 5 min and then placed on ice and the transformation was performed as described above.

Single colonies from the transformed bacteria plates obtained as above were inoculated into 3 ml of LB medium containing 100 µg/ml of ampicillin, and grown with vigorous shaking for 12–16 h. The plasmid DNA was then extracted as described and sequenced using a specific set of primers (Table 1, pGL2SEQ). At least two positive colonies containing the expected mutated fragment were used to set-up a DNA plasmid midi-prep as described before. The purified plasmids DNA (pGL2-OVO11Prom1Mut) were quantified by use of a ND-1000 Spectrophotometer (NanoDrop Technologies).

3.21.3 Transactivation Experiments

The day before the transfection 5×10^5 AR42J or INS-1 cells were plated into 6-well cell culture plates with 2 ml of medium. For the transfection, in sterile microcentrifuge tubes 200 µl of DMEM (containing 4.5 g/L glucose and 1mM sodium pyruvate, 25 mM HEPES, 2 mM L-glutamine, 100 units/ml penicillin, 100 µg/ml streptomycin for AR42J) and in RPMI 1640 (containing 4.5 g/L glucose and 1 mM sodium pyruvate, 2 mM L-glutamine, 10 mM HEPES, 50 µM β-mercaptohetanol 100 units/ml penicillin, 100 µg/ml streptomycin for INS-1 cells), 3 µl of FuGENE 6 were added and incubated for 5 min at room temperature. Two µg total of the appropriate mixture of vectors (containing 1µg of pGL2-OVO11Prom1 or pGL2-OVO11Prom1Mut, 100 ng of pcDNA5/TO-Ngn3, 100 ng of pcDNA5/TO-NeuroD/BETA2, 100 ng of pCMV-MyoD, 100 ng of pcDNA-E47, 100 ng of pRL-TK and pcDNA to 2 µg) were then added.

The mixtures were incubated at room temperature for 15 min. The FuGENE6/DNA complexes were then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were Incubate at 37°C for 2 days.

The cells were then washed twice with PBS and lysed with 250 µl of 1X Passive Lysis Buffer. The plates were then rocked several times to ensure complete coverage of

the cells with lysis buffer. The cells were then scraped and all the liquid transferred to a microcentrifuge tube, vortexed and centrifuged at $12,000 \times g$ for 10 min at 4°C .

For the Luciferase assay, 100 μl of Luciferase Assay Reagent II for each samples were dispensed into luminometer tubes and 20 ml of cell lysates were added. Immediately the tubes were read in a luminometer programmed to perform a 2-second measurement delay followed by a 10-second measurement read for Firefly Luciferase (reporter plasmid, pGL2-OVO11Prom1) activity. Immediately after, 100 μl of 1X Stop & Glo Reagent were added and the tubes briefly to mixed. The tubes were read again with 2-second measurement delay followed by a 10-second measurement read for *Renilla* Luciferase activity (internal control vector, pRL-TK). The ratio of the firefly:*Renilla* Luciferase activity represents the normalized reporter gene expression.

3.21.4 Immunohistochemistry

The adult rat pancreas and testis tissue were obtained from 4-week old male Sprague-Dawley rats. Tissues were fixed in 4% (w/v) paraformaldehyde dissolved in 0.1M Sorensen's buffer at room temperature for 2 h and incubated in cold PBS until use. The fixed tissues were then processed to make wax block and sectioned for 5 μm -thickness slides for staining.

For immunohistochemistry, tissue sections were first rehydrated by sequential incubation in:

Xylene I	7 min
Xylene II	7 min
100% Ethanol I	5 min
100% Ethanol II	3 min
95% Ethanol I	3 min
95% Ethanol II	3 min
70% Ethanol	10 min
dH ₂ O	5 min

The slides were then incubated twice in PBS containing 1% lamb serum for 5 min. The tissues were then permeabilized with 0.3% (v/v) Triton X-100 in PBS for 20 min. The antigen retrieval step was obtained by the incubation of slides in 10 mM citrate buffer for 20 min in microwave. To avoid high staining background and non-specific signals, endogenous peroxidase activity was quenched by incubating slides in 3% H₂O₂ for 30 min and the tissues were blocked in blocking serum (Vecstain Elite ABC), for 30 min at room temperature. Afterwards, tissues were incubated overnight at room temperature with rabbit anti-OVO11 antibody at 1:750 dilutions. As negative control, the OVO11 antibody was replaced with blocking serum. On the following day, the tissues were washed twice with PBS for 10 min each; the biotinylated secondary antibodies were then added onto the slides for 30 min and washed twice with PBS for 10 min each. The slides were incubated with ABC reagent for 30 min and developed in solution containing 0.5 mg/ml 3, 3'-Diaminobenzidine (DAB) and 0.015% (v/v) H₂O₂ until desired signal was observed. For counter staining for cell nuclei, the slides were incubated in filtered Harris's Hematoxylin solution for 10 seconds and sequentially washed in 0.15 M Sodium Borate solutions for 3 times. After the final wash, slides were incubated in sodium Borate solution for at least for 3 min. The slides were then dehydrated as follow:

70% Ethanol	5 min
-------------	-------

95% Ethanol I	1 min
95% Ethanol II	1 min
100% Ethanol I	1 min
100% Ethanol II	1 min
Xylenes I	2 min
Xylenes II	2 min

The slides were lastly mounted in Xylene-based mounting medium. The samples were examined under Olympus BH-2 microscope. The DIC images were photographed using Sony 3 CCD camera with LG-3 PCI module. Final figures were processed with Adobe Photoshop 7.0.

3.21.5 Western-Blot Analysis

The day before the transfection 5×10^5 AR42J or INS-1 cells were plated into 6-well cell culture plates with 2 ml of the appropriate medium. For the transfection in sterile microcentrifuge tubes 200 μ l of DMEM (containing 4.5 g/L glucose and 1mM sodium pyruvate, 25 mM HEPES, 2 mM L-glutamine, 100 units/ml penicillin, 100 μ g/ml streptomycin for AR42J) and in RPMI 1640 (containing 4.5 g/L glucose and 1 mM sodium pyruvate, 2 mM L-glutamine, 10 mM HEPES, 50 μ M β -mercaptohetanol 100 units/ml penicillin, 100 μ g/ml streptomycin for INS-1 cells), 6 μ l of FuGENE 6 were added and incubated for 5 min at room temperature. 2 μ g total of the appropriate mixture of vectors (containing 1 μ g of pcDNA5/TO-Ngn3, 1 μ g of pcDNA5/TO-NeuroD/BETA2, 1 μ g of pCMV-MyoD, 1 μ g of pcDNA-E47 and pcDNA5/TO to 2 μ g when necessary) were then added.

The mixtures were incubated at room temperature for 15 min. The FuGENE6/DNA complexes were then added to the cells in a drop-wise manner and the cells returned to the incubator. The culture plates were Incubate at 37°C for 2 days.

The cells were then washed twice with PBS and lysed with 250 μ l of CytoBuster Protein Extraction Reagent. The cell extracts were then transferred into microcentrifuge

tubes and centrifuged for 5 min at $16,000 \times g$ at 4°C . The supernatants were transferred to new tubes and used for protein quantification. For the protein quantification the BioRad Protein Assay reagent was used. In a 96-well plate 5 and 10 μl of cell extracts and 10 μl of solutions containing 0.1, 0.2, 0.3, 0.4 and 0.5 $\mu\text{g}/\mu\text{l}$ of BSA were pipetted. To the well 200 μl of 1X dye reagent were added, the plates incubated at room temperature for 5 min and the absorbance read at 595 nm in a microplate reader.

Five μg of proteins were separated onto two 15% SDS-Acrylamide/Bis-Acrylamide gel prepared as follow:

Resolving Gel.

30% Acrylamide/Bis Solutions, 37.5:1 mixture (30%T, 2.67% C)	5.0 ml
1.5 M Tris-HCl, pH 8.8	2.5 ml
10% APS (freshly prepared)	100 μl
10% SDS	100 μl
TEMED	10 μl
dH ₂ O	To 10 ml

Stacking Gel

30% Acrylamide/Bis Solutions, 37.5:1 mixture (30%T, 2.67% C)	840 μl
0.5 M Tris-HCl, pH 6.8	630 μl
10% APS (freshly prepared)	50 μl
10% SDS	50 μl
TEMED	5.0 μl
dH ₂ O	To 5 ml

The protein samples were added with 1/10 of Sample Buffer and boiled for 5 min. The gels were run gels at 200V (constant voltage) until the bromophenol blue dye is just off in presence of the 1X Electrode Buffer. After run the protein are transferred onto a nitrocellulose membrane using the Mini-Trans Blot Cell system (Bio-Rad)

After run, the gels were recovered and soaked for 15 min at room temperature in transfer buffer along with 2 sponges, 4 sheets (7.5 cm X 10 cm) of 3MM Whatman paper, and 1 Nitrocellulose membrane (7.5 cm X 10 cm).

The transfer sandwiches were then set-up adding, in the order, from the negative side of the sandwich: one sponge, two 3MM Whatman papers, the gel, the nitrocellulose membrane, two more 3MM Whatman papers, the last sponge. The sandwiches are then put into the tank apparatus filled with the tank buffer and blotted for 1 hr at 100V (constant voltage). After transfer the nitrocellulose membranes were incubated in Blotto solution for 1 hr. The membranes were then incubated with rabbit anti-OVO11 (1:2000) and goat anti- β -Actin (1:5000), respectively, in Blotto solution overnight at 4°C with gentle shaking.

The day after the membranes were washed for 10 min three times with TBST and then incubated with goat anti-rabbit-IgG HRP-conjugated (1:10000) and with rabbit anti-goat-IgG HRP-conjugated (1:10000), respectively, in Blotto solution for 1 hr at room temperature with gentle shaking. The membranes were then washed for 15 min four times with TBST. The membranes were then developed for 1 min with Western Lightning Chemiluminescence Reagent and autoradiographed with 5 min exposure.

3.22 IDENTIFICATION OF THE β -CELL SPECIFIC INSULIN GENE COATIVATOR A2.2

3.22.1 Nuclear Protein Extract from HIT T-15 cells

Nuclear extracts of HIT T-15 have been prepared using a large scale nuclear extract protocol as follow. About 1×10^{12} HIT T-15 cells were harvested from cell culture medium by centrifugation at room temperature for 10 min at 1000 x g. The cell pellet was then resuspended in five volumes of cold PBS and centrifuged again for 10 min at 1000 x g at room temperature. The cells were resuspended in five packed cell

pellet volumes of buffer A and allowed to stand for 10 min. The cells were collected by centrifugation at 1000 x g for 10 min at 4°C, suspended in two packed cell pellet volumes of buffer A (referred to the cell pellet volume before the addition of buffer A) and homogenized by 10 strokes of a glass Dounce homogenizer. The cell lysates were centrifuged for 10 min at 1000 x g and the pellet was resuspended in two pellet volumes of buffer A and centrifuged again for 20 min at 25,000 x g. The resulting pellet was resuspended in 3 ml every 10^9 cells of buffer B and homogenized with a Dounce homogenizer. The resulting suspension was first stirred for 30 min at room temperature and then centrifuged at 25,000 x g for 30 min at 4°C. The resulting supernatant was dialyzed against 50 volumes of buffer C for 5 h. The protein were quantified as described above.

The dialyzed nuclear extract was centrifuged at 25,000 x g for 20 min at 4°C and the supernatant aliquoted and stored at -80°C.

3.22.2 Electrophoretic Mobility Shift Assay (EMSA)

3.22.2.1 Oligonucleotides Labeling

The sequences of the oligonucleotides used as probes and competitors are reported in Table 1. The annealing of the probes and competitors were obtained dissolving each oligos (direct and complementary) in TE Buffer to a final concentration of 10 μ M. Each strand were then mixed together in equal molar amounts and heated at 94°C for 10 min and gradually cooled.

The double-stranded probes obtained as above were radiolabeled with [α - 32 P]dCTP and the Klenow fragment (3' \rightarrow 5') exo⁻ of DNA polymerase I. Briefly, in a 200 μ l tube were added:

NEB Buffer 2	2.0 μ l
dsProbe	4.0 μ l

dNTPs 33 μ M	1.0 μ l
$[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ (3000 Ci/mmole)	5.0 μ l
Klenow fragment (3' \rightarrow 5') exo^-	1.0 μ l
Sterile dH_2O	To 20 μ l

The reaction has been incubated at room temperature for 20 min. Radiolabeled double-strand oligonucleotides probes were purified by using G-50 Sephadex Columns TE (mini Quick Spin Columns) as described by the manufacturer.

3.22.2.2 Gel Mobility Shift Assay

For the gel mobility shift assay 3-5 μ g of nuclear extracts from HIT T-15 cell line prepared as described above were incubated in 1X DNA binding buffer with 1 μ g of poly(dI-dC) on ice for 15 min. One hundred fmole of the radiolabeled probe were then added to the mixture and the extracts further incubated at room temperature for additional 15 min. The competition experiments were performed by simultaneous addition to the binding reaction of an excess of unlabeled competitors along with the radiolabeled probes. The mixtures were separated onto a 6% Acrylamide/Bis-Acrylamide gel prepared as follow:

30% Acrylamide/Bis Solutions, 37.5:1 mixture (30%T, 2.67% C)	10 ml
5X TBE	5.0 ml
10% APS (freshly prepared)	100 μ l
TEMED	37 μ l
dH_2O	To 50 ml

The gel was run at 120V for 4 h at 4°C. At the end of run the gel was cast out and transferred on a Whatman sheets, covered with cellophane and dried for ~ 40 min at 80°C under vacuum. The gel(s) is then exposed with Biomax MR Kodak films at -80°C overnight. The day after the film is developed.

For the super-shift assay prior running the gel as described above the mixtures were pre-incubated for 30 min with 200 ng of the appropriate antibody (anti-C/EBP β and anti-C/EBP α).

3.22.3 Purification of A2.2 Coactivator

For the purification of the A2.2 MafA-transcription coactivator, 25 ml of nuclear extract obtained as described above, were used. The purification of A2.2 has been done using HiTrap Heparin columns and streptavidin-agarose columns. Briefly the HiTrap Heparin column was first equilibrated with 10 volumes of Binding Buffer, and then the nuclear extract was loaded on it. The column was then washed with 5-10 volumes of Binding Buffer and the fractions eluted with 10 ml of Elution Buffer by increasing salt concentration from 1 to 2 M. The fractions were tested for ability to form complex with A2.2 oligonucleotide as described above. Those showing gel-shift activity were pooled and further purified using an affinity chromatography on a streptavidin-agarose column.

The pooled protein fractions obtained from the first purification diluted to 15 ml with streptavidin-agarose Binding Buffer and incubated with biotinylated double strand A2.2 sequence (5'-CCAGGCAAGTGTTTGGAAACTGCAGCTTCAGCCCCTCTG-3') for 1 h on a rotating wheel at room temperature.

For the purification 3 ml of streptavidin-agarose resin were spun for 30 sec in at 1000 x g and resuspended with 15 ml of streptavidin-agarose Binding Buffer containing 500 μ g/ml BSA and 200 μ g of carrier DNA. The tube containing the resin was then gently mixed on a rotating wheel for 1 h at room temperature to block nonspecific protein and nucleic acid binding sites present on the streptavidin-agarose resin. The resin was then spun and added with the protein/biotinylated-oligonucleotides complexes, washed twice with streptavidin-agarose Binding Buffer. The proteins were eluted with Elution A2.2-Buffer containing increasing concentration of salt and collected into fractions.

3.22.4 Colloidal Gold Protein Staining

To confirm the presence of proteins in the eluted fractions a colloidal gold staining was performed. Briefly, after SDS-PAGE the proteins were transferred onto a nitrocellulose membrane, as describe above. The membrane was incubated twice with 100 ml of TTBS for 20 min. After the incubation, 100 ml of dH₂O were added and the membrane washed for 2 min. The washing was repeated twice. 50 ml of Colloidal Total Protein Stain were added and incubated overnight. Following the gold staining the membrane was washed twice with 100 ml of dH₂O for 1 min.

3.22.5 Western-Blot Analysis

About 5 µg of proteins obtained from the affinity chromatography on an streptavidin-agarose column were separated onto two 12% SDS-Acrylamide/Bis-Acrylamide gel prepared as follow:

Resolving Gel.

30% Acrylamide/Bis Solutions, 37.5:1 mixture (30%T, 2.67% C)	4.0 ml
1.5 M Tris-HCl, pH 8.8	2.5 ml
10% APS (freshly prepared)	100 µl
10% SDS	100 µl
TEMED	10 µl
dH ₂ O	To 10 ml

Stacking Gel

30% Acrylamide/Bis Solutions, 37.5:1 mixture (30%T, 2.67% C)	840 μ l
0.5 M Tris-HCl, pH 6.8	2.5 ml
10% APS (freshly prepared)	50 μ l
10% SDS	50 μ l
TEMED	5 μ l
dH ₂ O	To 5 ml

The protein samples were added with 1/10 of Sample Buffer and boiled for 5 min. The gels were run at 200V (constant voltage) until the bromophenol blue dye is just off in presence of the 1X Electrode Buffer. After run the protein are transferred onto a nitrocellulose membrane using the Mini-Trans Blot Cell system (Bio-Rad)

After run, the gels were recovered and soaked for 15 min at room temperature in transfer buffer along with 2 sponges, 4 sheets (7.5 cm X 10 cm) of 3MM Whatman paper, and 1 Nitrocellulose membrane (7.5 cm X 10 cm).

The transfer sandwiches were then set-up adding, in the order, from the negative side of the sandwich: one sponge, two 3MM Whatman papers, the gel, the nitrocellulose membrane, two more 3MM Whatman papers, the last sponge. The sandwiches are then put into the tank apparatus filled with the tank buffer and blotted for 1 hr at 100V (constant voltage). The membranes were then incubated with mouse anti-MafF/G/K (1:1000) and mouse anti-MafA (1:1000), respectively, in Blotto solution overnight at 4°C with gentle shaking.

The day after the membranes were washed for 10 min three times with TBST and then incubated with rabbit anti-mouse-IgG HRP-conjugated (1:10000) in Blotto solution for 1 hr at room temperature with gentle shaking. The membranes were then washed for 15 min four times with TBST. The membranes were then developed for 1 min with Western Lightning Chemiluminescence Reagent and autoradiographed with 5 min exposure.

3.22.6 TCA Protein Concentration

In a 1.5 ml microcentrifuge tube 1 volume of 100% (w/v) of trichloroacetic acid was added to 4 volumes of eluted fractions and incubated for 10 min at 4°C. The samples were then centrifuged at 14,000 x g at 4°C for 5 min and the pellet washed with 200 µl of cold acetone and spun again as before. The cold-acetone washing has been done twice. The pellet was let dry for 5 min at 95°C before resuspend it with Protein Loading Buffer.

3.22.7 Mass Spectrometry Analysis

Mass spectrometry analysis was performed at the Joslin Proteomic Core by a microcapillary Reverse-Phase HPLC nanoelectrospray tandem mass spectrometry (µLC/MS/MS).

The results were analyzed blasting the sequences using SEQUEST database software. The peptide and the matching proteins coming out from the tandem mass spectrometry analysis and blast were blasted against different databases to confirm the first blast results and to select candidates with nuclear localization. The databases used are:

MASCOT (http://www.matrixscience.com/search_form_select.html)

OMSSA (<http://pubchem.ncbi.nlm.nih.gov/omssa/>)

NCBI-BLST (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>)

EBI Protein WU-Blast2 (<http://www.ebi.ac.uk/blast2/>)

Softberry (<http://www.softberry.com/berry.phtml>)

Prowl Tools (<http://prowl.rockefeller.edu/>)

BLSTM-LOC (<http://139.91.72.10/blstm/blstm.html>)

Database Used For Sequence Analysis

DBD: Transcription factor prediction database (<http://dbd.mrc-lmb.cam.ac.uk/DBD/index.cgi?Home>)

MatInspector (<http://www.genomatix.de/products/MatInspector/>)

TRANSFAC (<http://www.gene-regulation.com/pub/databases.html>)

Chapter 4

RESULTS

4 RESULTS

During development the formation of β -cells is controlled by the activation and extinction of a large number of genes. Transcription factors orchestrate the intricate pathways of differentiation by regulating the rate of transcription of an array of genes. Genetic studies have begun to unravel the complex cascade of factors that control the proliferation and differentiation of cells in the development of the pancreas. Several homeodomain and bHLH transcription factors have been reported to play important roles in regulating the differentiation of pancreatic endocrine cell. However the general mechanism leading to the pancreas formation it is still unclear. The first approach, used in the present Thesis work, to address the problem has been studying the genetic pathway leading the differentiation from undifferentiated cells to the β -cell phenotype.

Therefore, the aims of the first part of the work have been:

- 1) *To mimic the differentiation program during endocrine pancreas differentiation.*
- 2) *To identify new genes involved in this process.*

Among the transcription factors involved in pancreas differentiation, Ngn3 is one of the essential expressed early in the differentiation of the islet cells and it is considered to be a marker of endocrine progenitor cell during development of the pancreas. Moreover, the timing of expression of this transcription factor corresponds to the specific developmental stage of the formation of the definitive endoderm. The subsequent patterning of endoderm leads to the formation of many of the major organs including pancreas, liver, lungs, thyroid and intestine.

According to that, the present research project aimed to address also the following points:

- 3) *Is the transient expression of Ngn3 a prerequisite to have the proper molecular context to achieve β -cells differentiation in vitro?*
- 4) *Is the endoderm patterning a prerequisite to have the proper cellular context to achieve β -cells differentiation?*

5) *Which are the direct downstream targets of Ngn3?*

The experimental strategy developed here helped answering these questions or at least helped to set-up strategy that needs to be refined.

4.1 Mimicking the differentiation program of the pancreas and identify new genes involved in endocrine development

To address the first two aims a simple cell model of differentiation and a high-throughput experimental approach have been chosen. As a model of differentiation the AR42J cell line was selected for their ability to differentiate into insulin-producing cells when treated with growing factors. In particular, the AR42J clone named B13 posses the highest ratio (up to 90%) of differentiation into insulin producing cells upon treatment with Act-A and HGF. In order to analyze the gene and protein expression profile driving the differentiation program during endocrine pancreas development, both microarray platform and two-dimensional analyses have been chosen as high-throughput assays. The goal of microarray and two-dimensional gel experiments is to survey the patterns of expression of hundred or thousands of genes and proteins in a single experiment.

4.1.1 AR42J-B13 Differentiation

Although the AR42J-B13 cell line is, at the moment, the best cell model to mimic as much as possible the physiological differentiation of β -cells it shows some phenotypic instability, such as spontaneous differentiation in absence of growth factors stimuli. Therefore, it has been first necessary to stabilize the cells by treating them with 1 nM dexamethasone for 2 days. In that way the cells were “reprogrammed” back to their acinar phenotype. The cells were then maintained in culture, as described in the “Materials and Methods” Chapter, for 2 days, then diluted and plated in 12-well plates to isolate single clonal colonies. The AR42J-B13 isolated clones were then assayed before

and after the differentiation to confirm the efficiency of the growth factors treatment. Briefly 2×10^6 AR42J-B13 cells were split and plated the day before the treatment in a 75 cm^2 cell culture flask. After 24 h, 2 nM Act-A and 500 pM HGF were added to the culture medium and the cells were maintained in culture for 48 h. The degree of the differentiation treatment has been determined evaluating the change of phenotype (Fig. 1 A) and analyzing the gene pattern expression.

When treated with Act-A coupled with HGF the AR42J-B13 cells change morphology extending three or more neurite-like processes. More detailed informations about the degree of differentiation were obtained by the analysis of the transcription factors activated during β -cell development and know to be important for insulin gene transcription. For that purpose total RNA, from treated and untreated cells, was isolated using Trizol reagent. The RNA has been reverse transcribed and pancreas specific markers Isl-1, Nkx2.2, NeuroD/Beta2, Pax4, Ngn3, Insulin have been amplified by PCR (Fig. 1B).

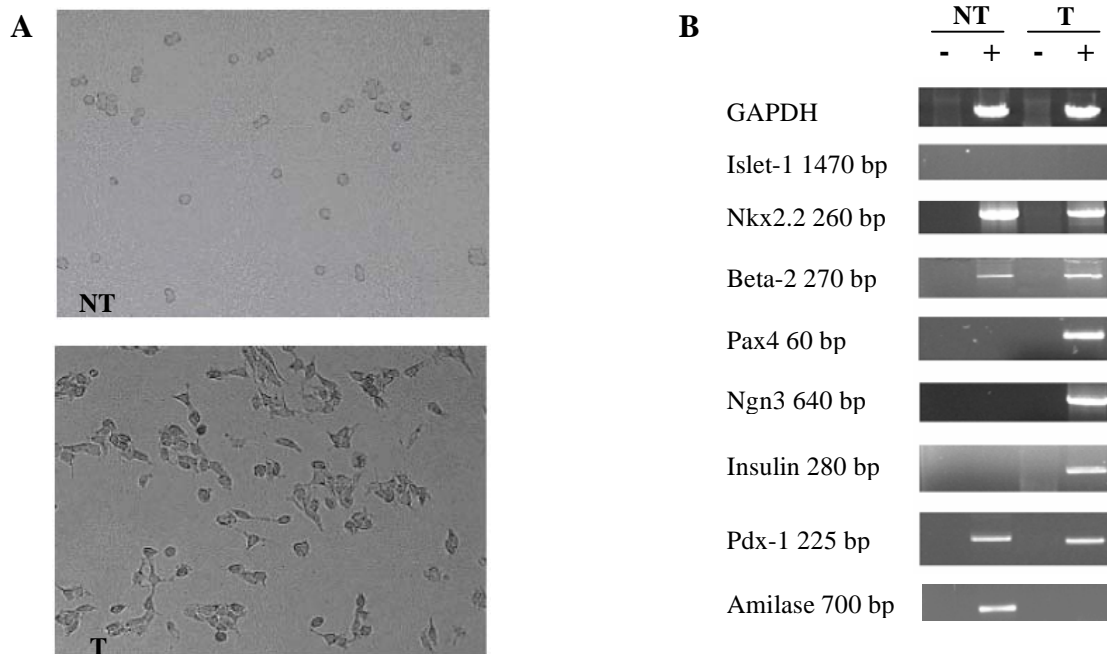


Fig. 1: **A)** The selected AR42J-B13 clone before treatment (NT) and after 48 h treatment with 500 pM HGF and 2 nM Act-A (T). **B)** RT-PCR analysis of the markers for the cell line AR42J-B13 before treatment (NT) and after differentiating treatment. 30 cycles PCR amplification.

As shown in Figure 2, Pdx1, Nkx2.2 and NeuroD/BETA2 were detected in naive AR42J-B13 cells and did not show any significant changes during the differentiation process. On the other hand the expression of Pax4, Ngn3, and insulin genes was markedly increased after the differentiation treatment. Conversely the expression of amylase, a marker of the undifferentiated phenotype, disappears after the treatment with Act-A and HGF. The expression of insulin has been confirmed also at protein level through immunohistochemistry. 1×10^5 cells have been split and cultivated on a 12- well cell culture plate in which were placed polyornithine-treated coverslip and treated with 2 nM Act-A and 500 pM HGF for 48 h. The cells were then fixed and treated with rabbit anti-insulin antibody and then with anti-rabbit-IgG FITC-conjugated antibody as described in the “Materials and Methods”. The results are reported in Figure 2. Similarly to the RT-PCR results, the insulin has been detected only in the Act-A/HGF-treated cells while any signal has been detected in the wild-type undifferentiated cells (Fig. 3). The positive signal for insulin is homogeneous with almost 100% of the treated cells appearing positive indicating that the step of clone selection has been efficient. The rat insulinoma RINm5F cell line has been used as a positive control for the immunohistochemistry (Fig. 4).

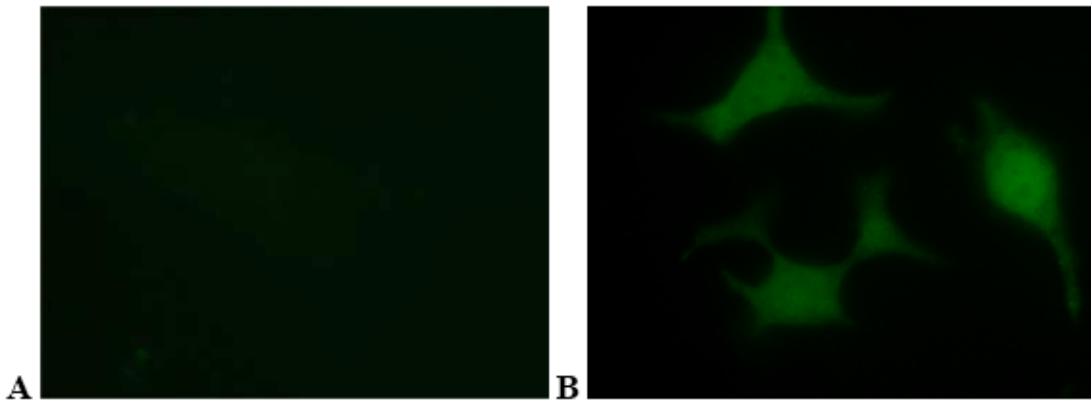


Fig. 2: AR42J-B13 cells treated with 500pM HGF and 2nM Act-A **A)** differentiated AR42J-B13 cells stained with α -Insulin antibody without α -rabbit-FITC-conjugated secondary antibody **B)** differentiated AR42J-B13 cells stained with either α -Insulin antibody and α -rabbit-FITC-conjugated secondary antibody

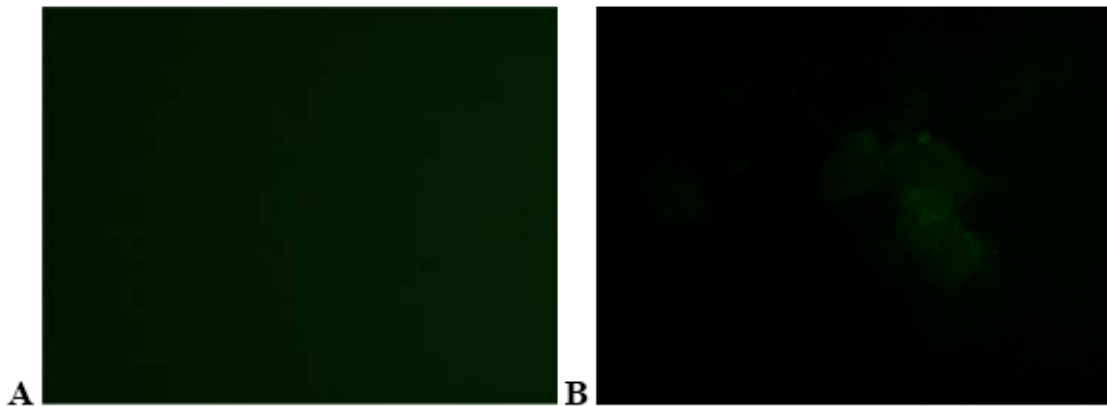


Fig. 3: Wild type AR42J-B13 cells **A)** undifferentiated AR42J-B13 cells stained with α -Insulin antibody without α -rabbit-FITC-conjugated secondary antibody **B)** undifferentiated AR42J-B13 cells stained with either α -Insulin antibody and α -rabbit-FITC-conjugated secondary antibody

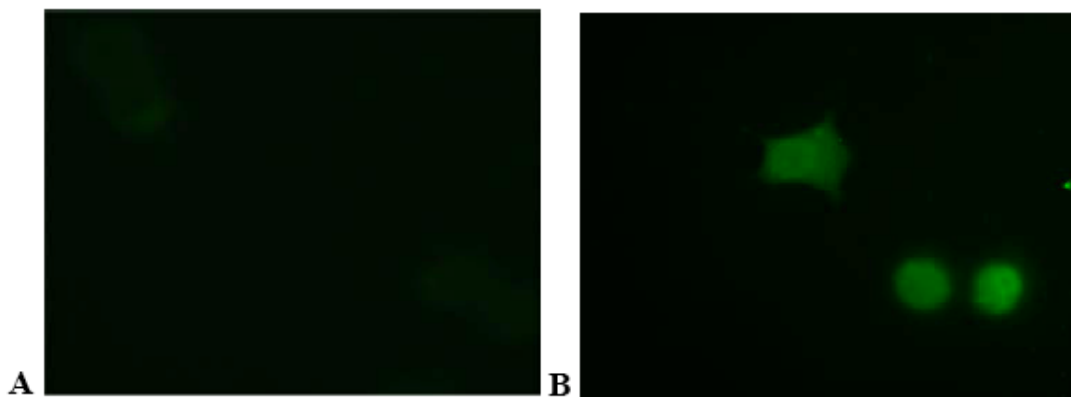


Fig. 4: RIN cells, positive control **A)** RIN cells stained with α -Insulin antibody without α -rabbit-FITC-conjugated secondary antibody **B)** RIN cells stained with either α -Insulin antibody and α -rabbit-FITC-conjugated secondary antibody

To address the second point of the aims we relied on two high throughput methodological approaches that are the microarray and 2-D gel proteomic analyses. A high throughput experimental approach is necessary to obtain the higher degree of information in term of genes and proteins expressed per experiment.

4.1.2 Microarray analysis

There are two main platforms of microarray technologies: cDNA and oligonucleotide arrays. cDNA microarrays are made with long double-stranded DNA molecules generated by enzymatic reactions such as PCR (Schena, Shalon et al. 1995). The cDNA arrays show the advantage that tissues- and development-specific arrays can be prepared offering a snap-shot of a well-defined phenotypic condition.

According to that, the cDNA array technology has been chosen as the most informative tool to identify genes expressed during a differentiative event such as the pancreas development. One of the platforms available for the pancreas is PanChip 6.0 from the Beta Cell Biology Consortium. The Mouse PancChip 6.0 contains 13,059 mouse cDNAs chosen for their expression in various stages of pancreatic development, many of which are not found on commercially available arrays. Total RNA isolated from uninduced and induced AR42J-B13 cells has been reverse transcribed and labeled using mono-reactive Cy3 and Cy5 dyes and hybridized to the arrays using the experimental strategy, as described in the Chapter 2. It is worth to mention that the results obtained from the microarray analyses were not straightforward and needed some adjustments to eliminate spurious data that could impair the identification of those genes that show a statistically significantly differential expression. In particular, an initial statistical analysis has been performed using the TIGR-MIDAS software and the results have been confirmed by R/Bioconductor (mArray package) analysis. The data normalization is shown in Fig. 5. After the normalization of the microarray on the different slides, the resulting signals were used altogether to perform a t-test to find genes differentially expressed. For that purpose the software S.A.M. has been used. As described in the previous Chapter, the S.A.M. software allows for the identification of differentially

expressed genes through a well established statistical analysis. The results obtained using either R/Bioconductor coupled with S.A.M./Excel and TIGR-MIDAS coupled with TIGR-MeV, are shown in figure 6 and in table 1. The genes considered differential expressed are those with a fold of change higher then the calculated delta value in the S.A.M. analysis and a p-value below 0.05%.

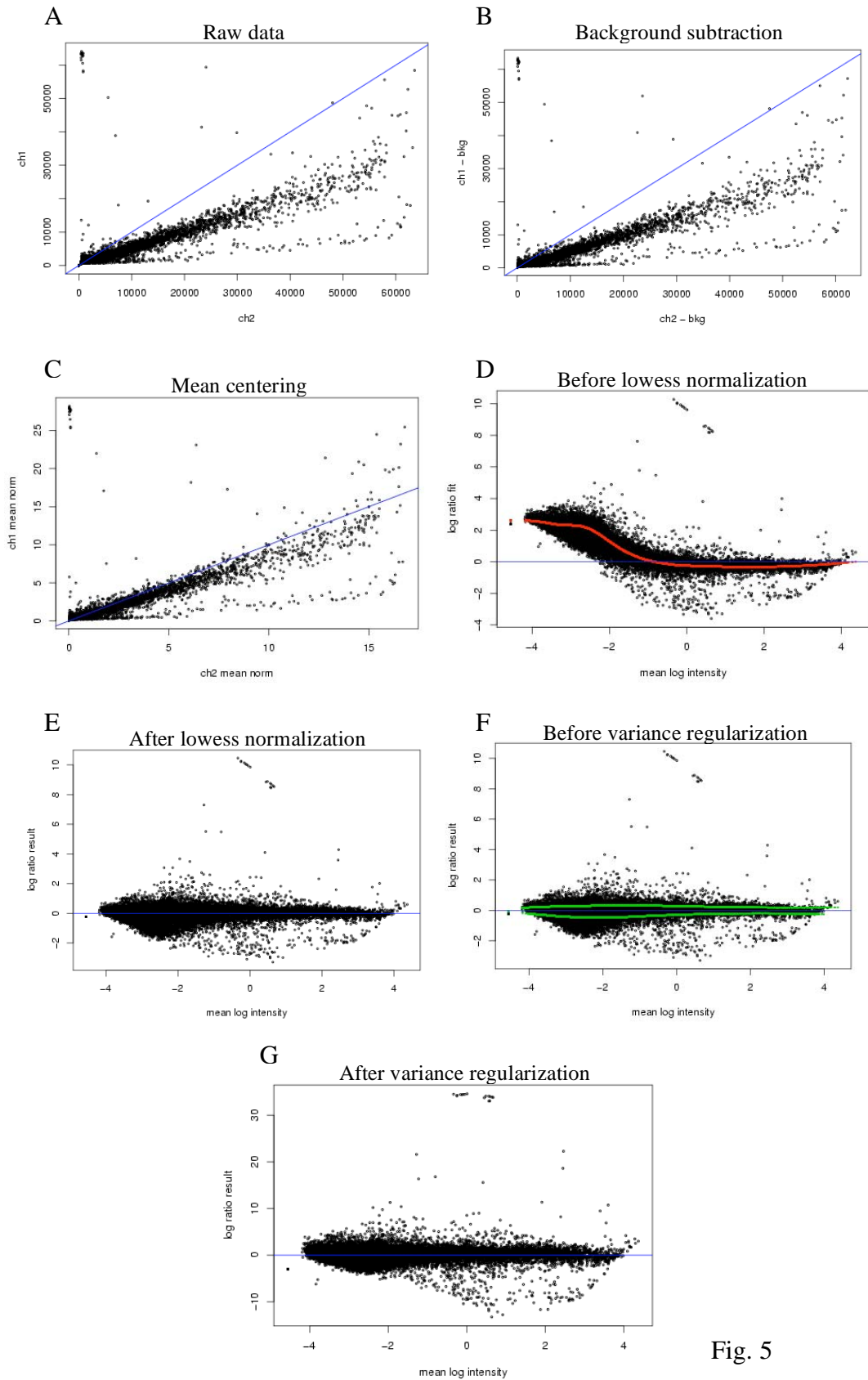


Fig. 5

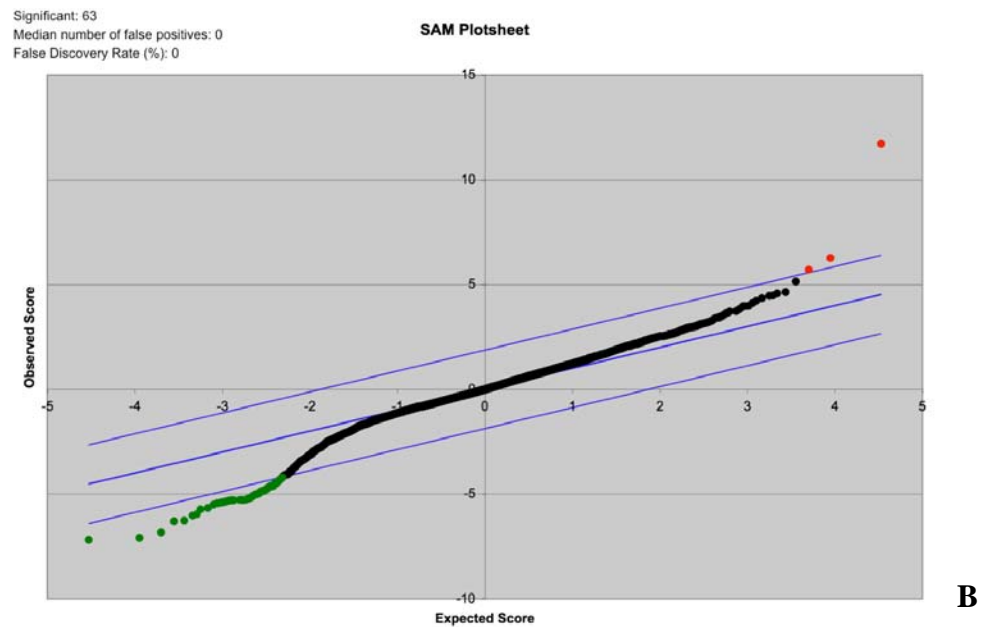
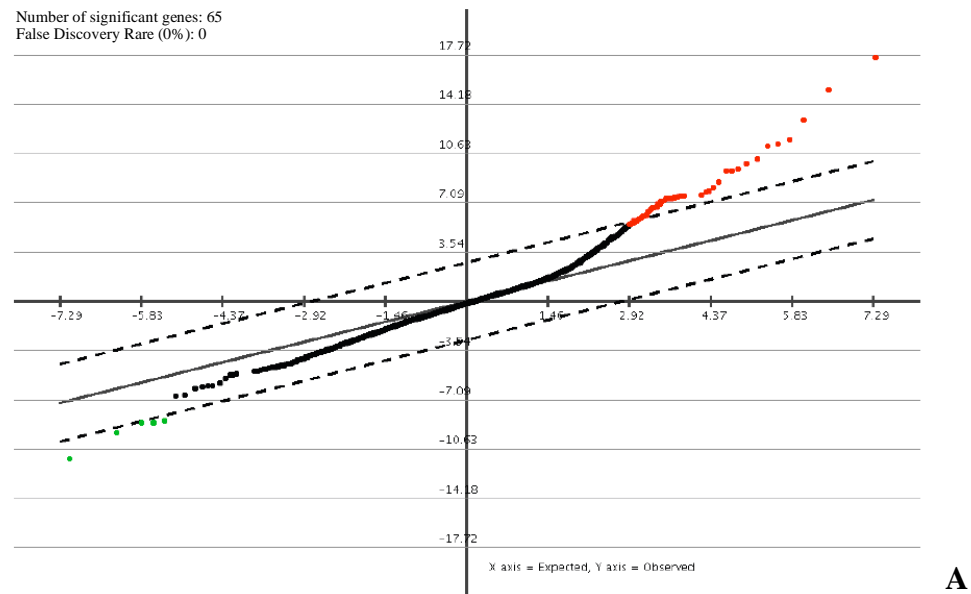


Fig. 6: S.A.M. analysis. **A)** S.A.M. analysis performed with TIGR-Mev after TIGR-MIDAS data normalization. In green the under-expressed genes while the over-expressed gene are in red. **B)** S.A.M. analysis performed with S.A.M./Excel after mArray/Bioconductor data normalization. In green the over-expressed gene while the under-expressed gene are in red.

Table 1:
Genes
Over-expressed

Gene ID	Gene Name	Score(d)	Numerator(r)	Denominator(s+s0)	Q-value(%)	localfdr(%)
1.1.7.1.5.12	2210010C04Rik	11,71558653	5,734673126	0,489490911	0	13,52871871
1.1.11.1.5.12	Tfpi	6,273836489	6,697059654	1,067458431	0	13,49564242
1.1.6.2.12.2	Htf9c	5,716593616	5,921854081	1,03590608	0	13,60791406

Under-expressed
Genes

Gene ID	Gene Name	Score(d)	Numerator(r)	Denominator(s+s0)	Q-value(%)	localfdr(%)
1.1.10.1.7.1	Amy2	-7,175277349	-6,666484715	0,929090876	0	0,836338435
1.1.5.1.11.8	Unc50	-7,086863662	-4,675675701	0,659766566	0	0,773088116
1.1.12.1.3.4	Rab28	-6,825898549	-4,532779252	0,664056053	0	0,587670445
1.1.5.1.11.4	Amy2	-6,295205328	-5,934221232	0,942657296	0	0,227095892
1.1.7.2.15.7	Hnrpa3	-6,268638924	-3,685059183	0,587856348	0	0,210358263
1.1.10.4.3.8	Gdpc5	-6,020324453	-3,106378449	0,5159819	0	0,069154865
1.1.11.4.11.11	Amy2	-5,964818354	-5,815819082	0,975020317	0	0,042589752
1.1.5.2.11.6	Amy2	-5,722898743	-5,604478011	0,979307561	0	0
1.1.2.3.5.10	1110001C20Rik	-5,657370628	-5,885075604	1,040249259	0	0
1.1.2.4.3.12	Pfdn4	-5,507331417	-7,616100573	1,382902171	0	0
1.1.5.4.1.18	4432411E13Rik	-5,443818577	-8,430187332	1,548579772	0	0
1.1.10.1.9.12	Amy2	-5,40488546	-5,760839216	1,06585778	0	0
1.1.9.1.5.10	Amy2	-5,286223179	-7,195571173	1,361193224	0	0,065294148
1.1.1.4.2.4	2410002O22Rik	-5,283346274	-3,224254581	0,610267511	0	0,067531934
1.1.6.2.7.6	Amy2	-5,271653467	-6,921820043	1,313026375	0	0,076893834
1.1.1.3.10.8	Amy2	-5,268407244	-6,677313616	1,267425487	0	0,079569539
1.1.7.3.10.16	5730472N09Rik	-5,220898654	-7,817055958	1,497262536	0	0,122655536
1.1.8.2.11.9	Amy2	-5,163916528	-9,115353675	1,765201592	0	0,184533259
1.1.12.4.11.6	Amy2	-5,06649949	-8,778007643	1,732558675	0	0,318470635
1.1.9.2.6.9	Amy2	-5,006615582	-8,053073651	1,608486515	0	0,420130311
1.1.8.3.6.7	Amy2	-5,001896757	-8,208091349	1,640995756	0	0,428817253
1.1.12.1.5.9	1110001C20Rik	-4,972762494	-7,308868989	1,469780428	0	0,484741995
1.1.9.2.7.7	Bub3	-4,956341953	-1,884672101	0,380254655	0	0,518041798
1.1.10.3.2.4	1810009A16Rik	-4,889339242	-7,19032699	1,470613233	0	0,667921141
1.1.9.1.11.11	UNKNOWN	-4,883129892	-7,912198472	1,620312924	0	0,68299528
1.1.3.2.11.1	Amy2	-4,855960242	-7,44457794	1,533080497	0	0,751422703
1.1.4.4.6.17	Amy2	-4,8556092	-8,044826306	1,656810912	0	0,752333534
1.1.4.2.11.3	Myo1b	-4,854935844	-7,223829028	1,487935013	0	0,754082586
1.1.8.3.7.3	Amy2	-4,816446089	-6,667386087	1,384295799	0	0,858361792
1.1.8.3.11.4	Amy2	-4,778805674	-8,772614573	1,83573369	0	0,968801534
1.1.9.4.1.6	Smoc2	-4,739694294	-4,432088109	0,935099995	0	1,092868576
1.1.6.3.11.9	Sfrs6	-4,695354249	-7,533217704	1,604398157	0	1,245578687
1.1.8.3.11.9	Rab8b	-4,672288908	-6,247306248	1,337097592	0	1,330297376
1.1.11.2.3.15	Amy2	-4,626235782	-5,354510546	1,157422751	0	1,510727212
1.1.6.1.6.16	Amy2	-4,62102627	-6,429889891	1,391441969	0	1,532109162
1.1.11.4.11.6	Anapc1	-4,620766021	-3,370346259	0,72939124	0	1,533182603
1.1.4.1.4.15	Amy2	-4,614397015	-7,386321777	1,600712239	0	1,559609779
1.1.3.3.7.9	Amy2	-4,60955454	-6,948712007	1,507458464	0	1,579905672
1.1.5.1.11.10	Amy2	-4,535517105	-7,272616689	1,603481262	0	1,912673616
1.1.3.1.11.2	Amy2	-4,523265198	-6,872874015	1,519449715	0	1,971931443
1.1.11.2.16.15	Hmqn1	-4,494280491	-2,927840419	0,65145921	0	2,117035405
1.1.6.4.16.8	Ar	-4,487229808	-2,356216718	0,525093837	0	2,153394156
1.1.9.2.1.18	2900070E19Rik	-4,439731404	-4,335612501	0,976548378	0	2,409472447
1.1.11.1.3.13	Rps16	-4,439660158	-1,780024293	0,40093706	0	2,4098714
1.1.12.3.5.10	Amy2	-4,390913946	-9,433323584	2,148373596	0	2,693593607
1.1.8.3.15.2	Eif2s2	-4,377287208	-2,425929704	0,554208483	0	2,77682799
1.1.6.1.2.18	Hmqb2	-4,305038736	-5,265659021	1,223138593	0	3,247988913
1.1.10.3.9.12	Ppm1b	-4,295759802	-4,27649024	0,995514283	0	3,31224673
1.1.11.1.10.15	Amy2	-4,276572021	-7,601115931	1,777385227	0	3,447910523
1.1.9.2.13.12	5230400G24Rik	-4,235868946	-2,435588707	0,574991516	0	3,748376256
1.1.6.4.6.4	Amy2	-4,204518373	-9,31540381	2,215569771	0	3,991857966

4.1.3 Proteomic Analysis

In a first set of experiments protein samples from treated and untreated AR42J-B13 cells were separated according their isoelectric points using a wide range of pH (strips of pH 3-10) followed by the second dimension (a classical SDS- PAGE) on a 10% polyacrilamide gel. The completion of these first sets of experiment gave the possibility to gain a broad overview of total protein distribution between undifferentiated and differentiate condition. The results are reported in Figure 7.

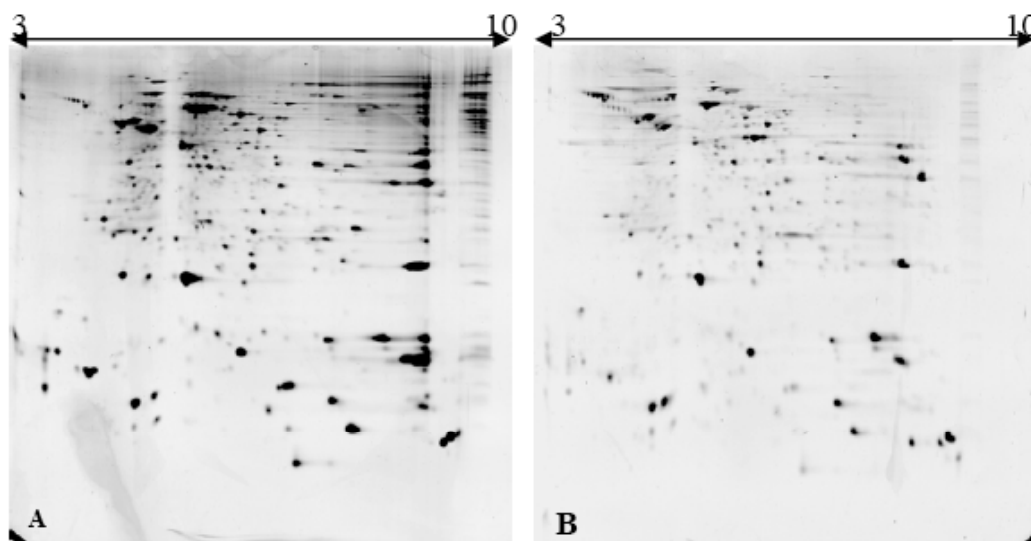


Fig. 7: 2D analysis. For the first dimension 110 μ g of protein extract were separated using a 13 cm strip pH 3-10. For the second dimension the proteins were separated on a 10% SDS-PAGE gel. **A)** sample from untreated AR42J-B13 and **B)** sample from AR42J-B13 treated with 500 pM HGF and 2 nM Act-A.

As evident the most interesting pattern of differentially expressed proteins accumulated on the top-left part of the gel that corresponds to the acid region of the strips and to the high molecular weight region for the SDS-PAGE. In the attempt to increase the efficiency of isoelectric separation for these proteins the pH range has been narrowed from pH 3-10 to pH 4-7. This pH gradient is covered with steps of 1 unit of pH allowing

a more detailed resolution of the protein isoelectrofocusing. At the same time the percentage of acrylamide has been lowered from 10% to 8%, in order to obtain a more efficient protein separation.

Following the 2-D separation and silver staining, the gel were converted into digital images using a scanner and analyzed using the ImageMasterTM software. To quantify the changes, the intensity levels of each spot were normalized by expressing the intensity of each protein spot in a gel as a proportion of the total protein intensity detected for the entire gel (relative volume, % vol).

After normalization, the gels from individual experiments were compared to detect those spots differentially expressed among different assays (Fig. 8) and used for mass spectrometry analysis.

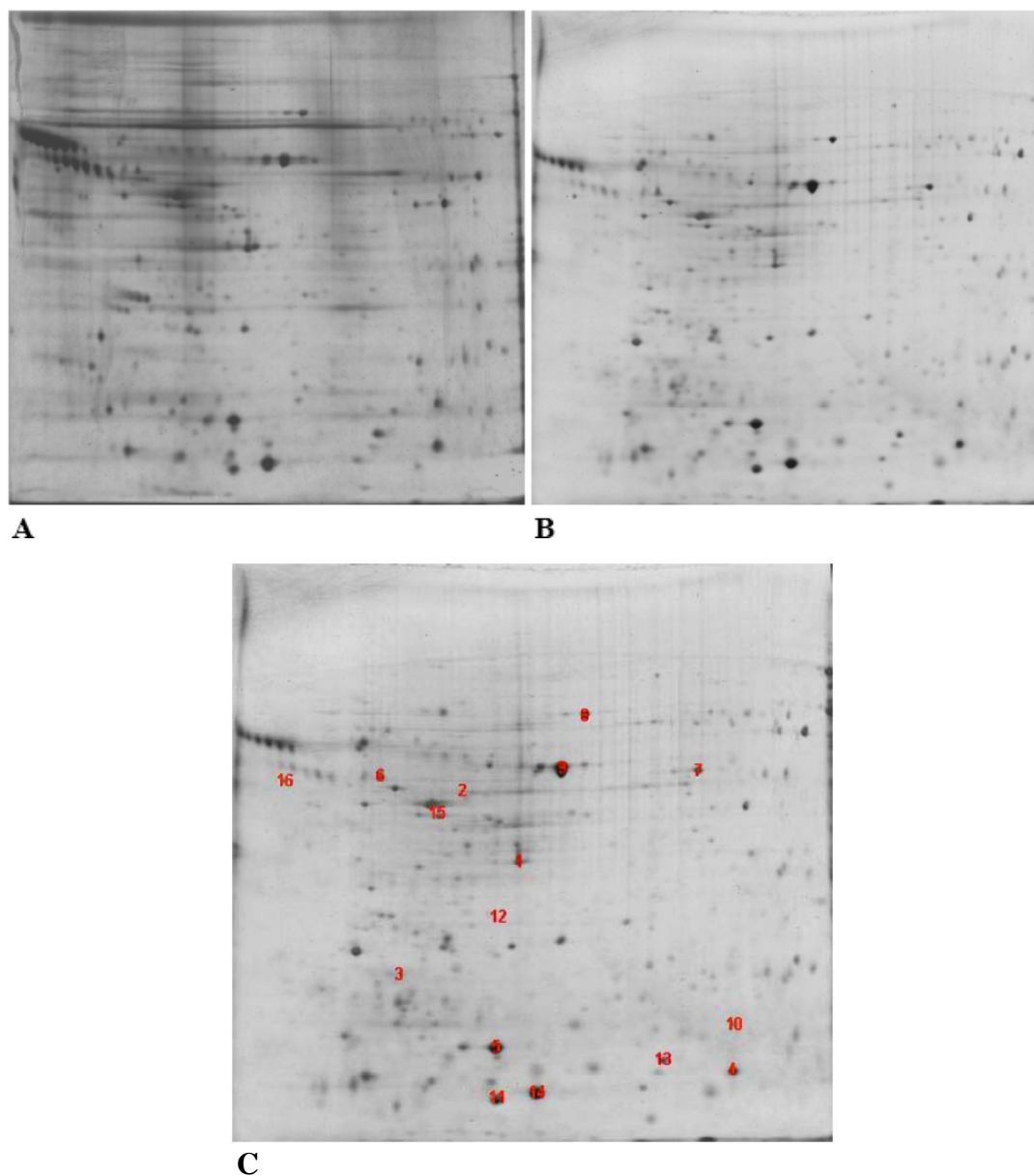


Fig. 8: The protein samples from undifferentiated **A)** and differentiated **B)** AR42J-B13 were separated using a pH range between 4-7 for the first dimension and an 8% acrylamide gel for the second dimension. **C)** After comparison and normalization, the differentially expressed spots were identified and submitted to the proteomic analysis.

For the MS protein identification, the differentially expressed spots were excised from gel, destained and subjected to reduction and alkylation. Once extracted, the peptides were loaded on LC-MS. The first separation has been performed using a C8 capillary column on a liquid-chromatography unit connected to a mass spectrophotometer. The output data (mass spectra and tandem mass spectra) have been used for the protein identification. Eighteen differentially over-expressed proteins were identified (Table 2) after Act-A/HGF treatment. Among the identified proteins from mass analysis the β -actin all the proteins are related with cell differentiation and cellular trafficking.

Table 2

	Identified proteins	Function/localization
1	β -Actin	Cytoskeleton
2	NSF	Vesicular trafficking/cytoplasm
3	VGf8a	Energetic homeostasis/vesicle secretion
4	PARK7	Oxidative stress/cytoplasm
5	UCHL1	Ubiquitin monomers formation/cytoplasm
6	PDIA1	Modification of disulphide bonds/cytoplasm
7	PDIA3	Modification of disulphide bonds/cytoplasm
8	GRP75	Oxidative stress/cytoplasm
9	CH60	oxidative stress/cytoplasm
10	ERp29	Protein folding/endoplasmic reticulum
11	PRDX-2	Redox system/cytoplasm
12	PRDX-6	Redox system/cytoplasm
13	ENOA	Plasminogen activator/cytoplasm
14	PEBP	Cell signaling/cytoplasm-
15	GLE1	tRNA transport/nucleus-cytoplasm
16	Calreticulin	Protein folding/endoplasmic reticulum

4.2 Mimicking the expression of Ngn3 during the differentiation program

To address the points 3 and 4 of the aims, the F9 cell line has been used as a model for a controlled differentiation process. These cells can differentiate into parietal endoderm in the presence of retinoic acid (RA) and dibutyryl cyclic AMP. To attain at the same time both a high transduction efficiency and a customized activation of gene expression the RevTet-On transduction system have been used. The RevTet-On couples the efficiency of *retroviral gene delivery* with that of the Tetracycline-inducible gene expression system.

The cells have been maintained in culture as described in “Materials and Methods”. The day before infection the cells have been split and plated in 100 mm culture dishes at a density of about 60%. The days after the cells were infected with about 5 MOI of RevTet-On viruses and then selected for neomycin resistance for at least seven days. Positive clones were expanded and selected to identify those showing the higher inducibility and lower background as described in “Material and Methods”. At least 8 colonies showing >15-fold induction were obtained. One of these stable rtTA expressing clones has been then infected with 5 MOI of RevTRE-Ngn3 viruses and then selected for hygromycin resistance. Positive clones have been selected and expanded. Twelve clones have been obtained. Stable clones have been then selected for doxycycline-induced Ngn3 expression by assaying for Ngn3 expression after treatment with 2, 4, 6, 8 and 10 µg/ml of doxycycline for at least 48 h. Expression of Ngn3 has been evaluated by RT-PCR (Fig 9). The picture shows the result obtained after the dose-curve experiments. Unfortunately there are very slight or any differences between the negative control (no Dox) and the induced ones. These results have been observed for all the clones selected.

According to that it has been decided to use the T-REx™ System. This system include an inducible expression plasmid (pcDNA5/TO) for expression of gene of interest under the control of the strong human CMV promoter and two tetracycline operator 2 (TetO2) sites and a regulatory plasmid, pcDNA6/TR, which encodes the Tet repressor (*TetR*) under the control of the human CMV promoter. The mouse Ngn3 gene has been cloned in the response vector pcDNA5/TO as described in “Materials and Methods”.

Again the F9 cell line has been used as a model for a controlled differentiation. The cells have been maintained in culture as described in “Materials and Methods”. The day before transfection the cells have been split and plated in 60 mm culture dishes at a density of about 60%. Gene transfer has been carried out using LipofecAMINE 2000 reagent. The pcDNA5/TO-Ngn3 has been cotransfected with the regulatory plasmid, pcDNA6/TR with a ratio 6:1 (w/w) pcDNA6/TR:pcDNA5/TO-Ngn3 for a total of 7 µg of DNA. After transfection, the target cells have been selected for hygromycin and blasticidin resistance, respectively. Positive clones have been selected and expanded. Twenty four clones have been obtained. Stable clones have been then selected for doxycycline-induced Ngn3 expression by assaying for Ngn3 expression after treatment with 2, 4, 6, 8 and 10 µg/ml of doxycycline for at least 48 h. Interestingly, as reported in Fig. 10, there is an increase of level of expression of the Ngn3 messenger with the increase of the doxycycline concentration with low basal expression in not induced sample. Unfortunately, the differences of induction between 2 and 10 µg/ml are not so dramatic.

To establish the time needed to obtain the maximum expression of Ngn3 a time course experiment has been performed. As reported in Fig. 11 the maximum expression can be obtained after 48 h of treatment.

Once established a cell line in which the expression of Ngn3 can be turned-on and off experiments of differentiation/induction have been set-up as outlined in Fig. 12. Briefly, the F9 cells were plated in 100 mm culture dishes at a density of about 60% and treated with 0.1 nM RA and 1 mM db-cAMP for 72 h and then transferred in a Petri dish to allow the formation of Embryod Bodies (EB) for 48 hr. From this point onward the EBs have been treated with 8 µg/ml of doxycycline for further 48 h. After 48 h from the withdrawal expression of Ngn3 and of other pancreas specific genes (NeuroD/BETA2, Pax6, Pax4, Nkx2.2, Pdx1 and Isl1 and for the pancreatic hormones insulin, glucagon) have been evaluated by RT-PCR. The results are reported in Fig. 12.

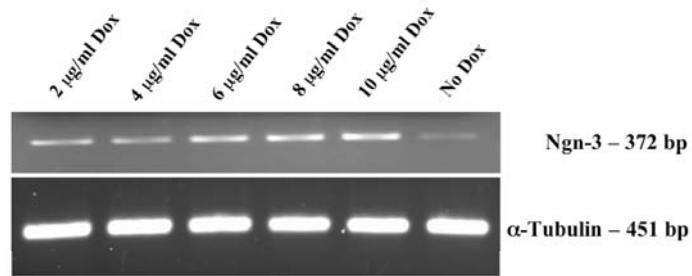


Fig. 9: Dose curve of the RevTet-On system. The stable clone Tet-On/TRE-Ngn3-F9 was induced for 48 h with different concentration of doxycycline before to test the Ngn3 expression by RT-PCR. As control for the retrotranscription α -Tubulin was used. 30 cycles PCR amplification.

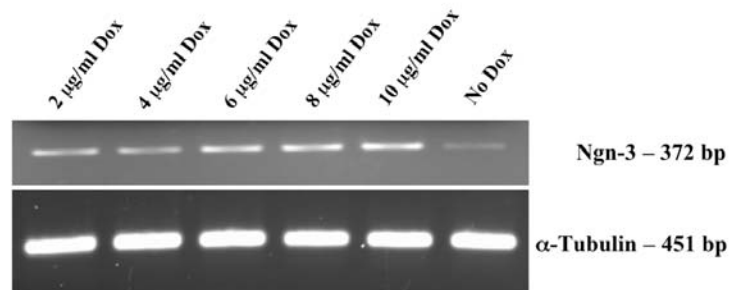


Fig. 10: Dose curve of the T-REx™ System. The stable clone 6TR/5TO-Ngn3-F9 was induced for 48 h with different concentration of doxycycline before to test the Ngn3 expression by RT-PCR. As control for the retrotranscription α -Tubulin was used. 30 cycles PCR amplification.

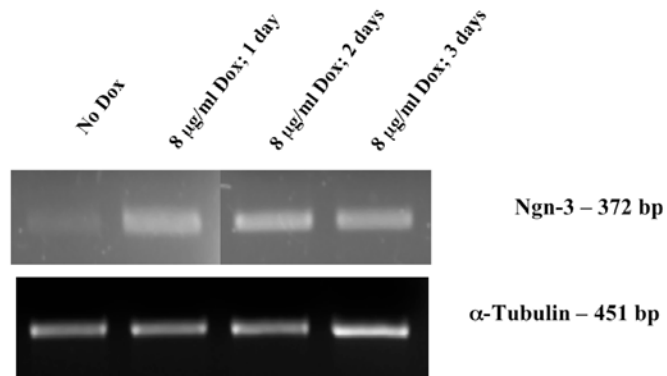


Fig. 11: Time course assay. The stable clone 6TR/5TO-Ngn3-F9 was induced for 24, 48, 72 h with 8 µg/ml of doxycycline before RNA extraction and RT-PCR analysis. As control for the retrotranscription α -Tubulin was used. 30 cycles PCR amplification.

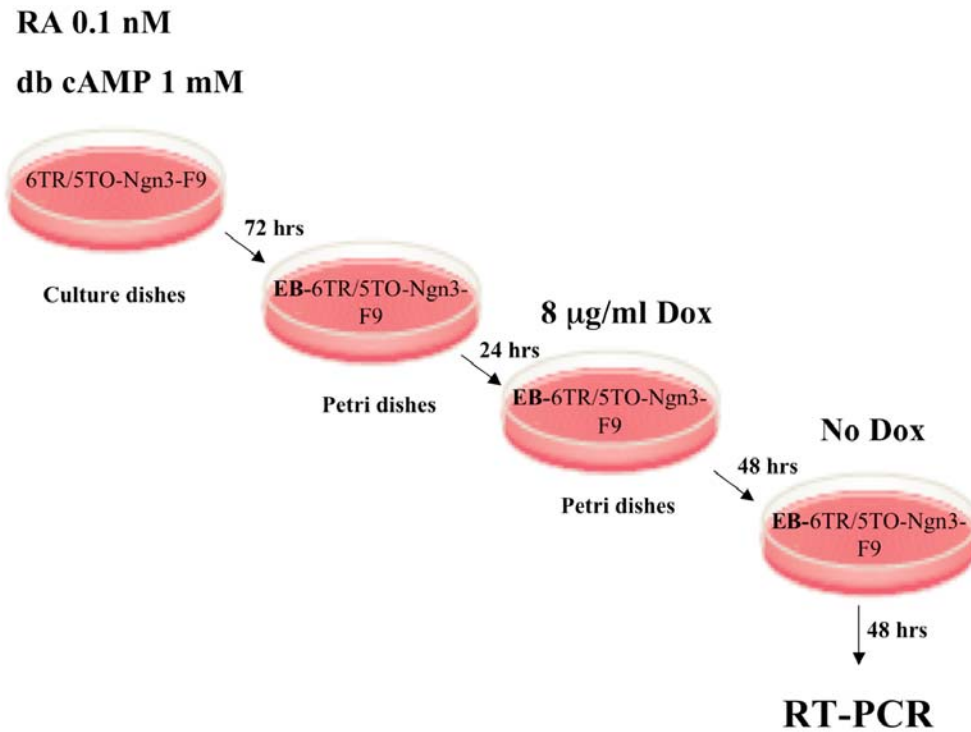


Fig. 12: differentiation protocol for the 6TR/5TO-Ngn3-F9 double stable clone. Briefly the cells were plated on gelatin coated dishes and kept in culture with 0.1 nM RA and 1 mM db-cAMP for 72 h. After that period the cells were kept in culture on Petri dishes for the embryo-bodies formation. After 24 h the medium was changed with fresh medium containing 8 µg/ml Dox and kept in culture for 48 before the Dox was removed from the medium. After 48 h the embryo-bodies were assayed for the expression of beta-cell-development markers. As control, 6TR/5TO-Ngn3-F9 double stable clone was treated with the same protocol except the RA and db-cAMP treatment was skipped.

Ngn3 was found to be expressed in doxycycline-treated EBs but at low level also in control doxycycline-not treated EBs F9/RA. A significant level of expression of the pancreas cell markers NeuroD/BETA2, Pax6 and Isl1 was detected only in the endoderm-derived EBs (see Fig. 12). Slight evidence for the transcription of mRNAs of these factors was found in the control.

Pdx1, Nkx2.2 and Pax4 were no detected in both doxycycline-treated EBs and not-treated EBs (data not shown). The RT-PCR for insulin and glucagon was negative, indicating that the transfected cells were unable to express pancreatic hormones (data not shown).

The presence of a low, but significant (in term of transcription factor), expression of Ngn3 also in the doxycycline-not treated control impaired the results.

The leak of gene expression, even in absence of inducer, of the tetracycline-based systems used induced us to pursue to use new inducible system(s). One of the newest systems available is the RheoSwitch[®] system, from New England Biolabs.

The RheoSwitch System permits maximum control of gene expression in mammalian cells. Analogous to the operation of a rheostat, the RheoSwitch technology allows induction and adjustable control of gene expression. The induction is achieved through the interaction between a ligand RSL1 (a synthetic diacyldrazine) and a chimeric bipartite nuclear receptor expressed by the Rheo-Receptor plasmid pNEBR-R1.

The RheoSwitch technology offers several advantages over other inducible systems. Its precise control gives negligible levels of basal expression in the absence of inducer and greater than 10,000 fold inductions when RSL1 ligand is present because, unlike other systems, which rely on steroids or other drugs, the synthetic ligand RSL1 does not exhibit cross talk with endogenous transcription factors.

As first step stable F9 clones expressing the Rheo-Receptor plasmid have been created. The F9 cells were kept in culture as described in the “Materials and Methods” Chapter. The day before the transfection the cells were split and plated in a 60 mm cell culture plate at a 60% density. The day after, the cells were transfected with the pNEBR-R1 plasmid and selected for G418 resistance for at least 10 days. Seventeen single colonies (F9-R1) were obtained.

Positive clones were expanded and selected to identify those showing higher inducibility. To select the clones a Luciferase assay has been performed. Briefly, the day before the transfection each clone was split to 4 wells of a 12-well cell culture plate. The day after, the cells were transiently co-transfected with pNEBR-GLuc and with pCMV- β -Galactosidase. Height h after the RSL1 activator was added to the media at a final concentration of 0.5 μ M and the cells maintained in culture for further 12 h. The results are shown in Fig. 13. The relative Luciferase intensities have been normalized against the protein concentration as well as against the β -Galactosidase activity. About 8 clones showing the higher induction level with the lower background have been selected and amplified.

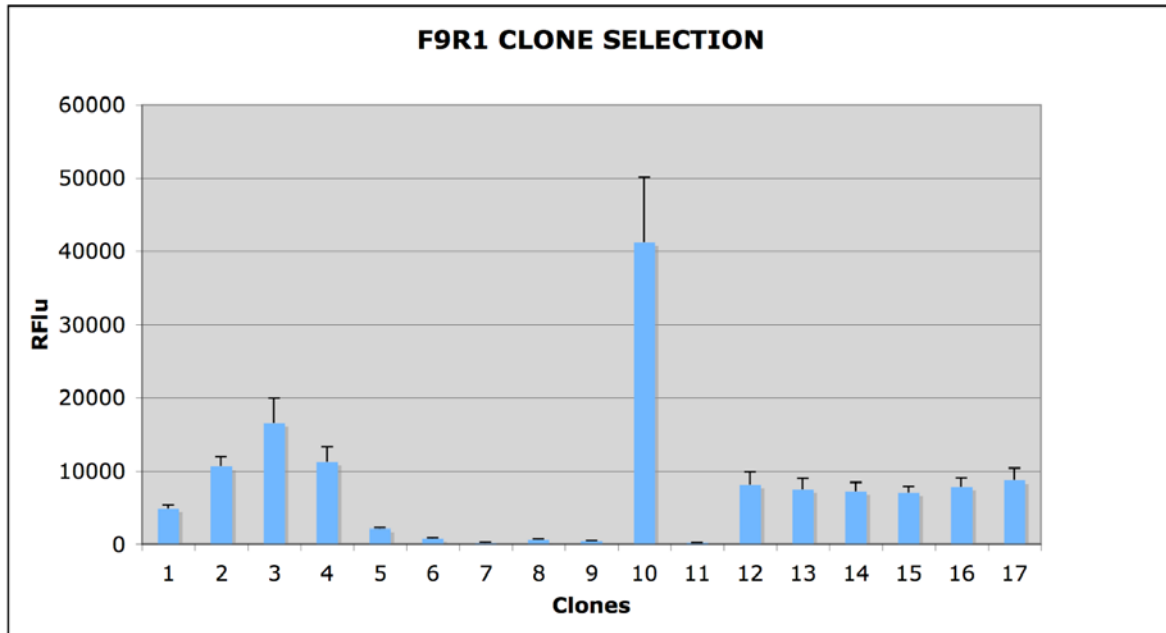


Fig.13: Stables F9-R1 clones selected based on the relative luciferase activity. Each clone has been tested twice in four replicates to be statistically significant. Standard deviation has been calculated.

Once established the regulatory F9 cell line the target gene (Ngn3) has been cloned in the appropriate expression vector. In particular, the Ngn3 CDS has been cloned into the pNEBR-X1-Hygro plasmid as described in the “Materials and Methods”. The efficiency of expression of the recombinant protein has been evaluated by RT-PCR and western-blot analyses after transfection and induction in the NIH-3T3-47 cell line, expressing the Rheo-Receptor. Twenty four h before transfection the cells were split in a 12-well cell culture plate at a density of about 50%. The day after the cells were transfected with 0.75 μg of pNEBR-X1-Hygro-Ngn3, 0.75 μg of the internal transfection control pCMV- β -Galactosidase and 2 μl of TransPassD2. Eight h after transfection, the activator RSL1 was added to the medium at a final concentration of 0.5 μM and the cells maintained in culture for further 24 h. Total RNA and proteins have been extracted and used to perform RT-PCR and western-blot analyses, respectively. The results are reported in Figure 14.

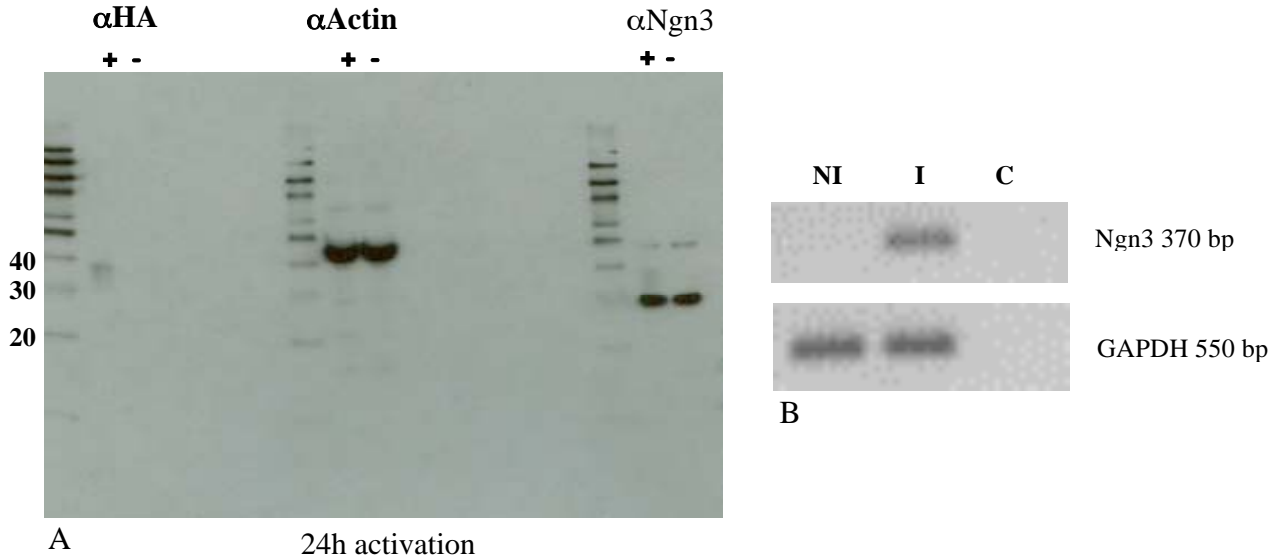


Fig. 14: Transfection of NIH 3T3-47 with pNEBR-X1-Ngn3 Hygro. **A)** After 24 h of activation with RSL1 Ngn3 was detected as very faint band on western blot analysis using antibody against HAtag. Antibody against Ngn3 shown very low specificity. **B)** RT-PCR after 24 h activation with RSL1. 30 cycles of PCR amplification.

As evident at none of the incubation times examined it has been possible to detect a significant expression of Ngn3 both at transcription and translation level regardless the antibody used (anti-HA-tag or anti-Ngn3). The pNEBR-X1-Hygro backbone is structurally organized as reported in Figure 15. Interestingly it shows the Hygromycin resistance CDS with an opposite orientation with respect that of a CDS cloned into the MCS. This opposite orientation of the Hygromycin-resistance CDS could interfere with the expression of the gene of interest (Dr. George Tzertzinis, New England Biolabs, personal communication).

In the attempt to overcome this problem the CDS for Ngn3 has been then cloned in the pNEBR-X1-Puro vector that shows the Puromycin resistance CDS in the same orientation of the gene of interest (Fig. 15).

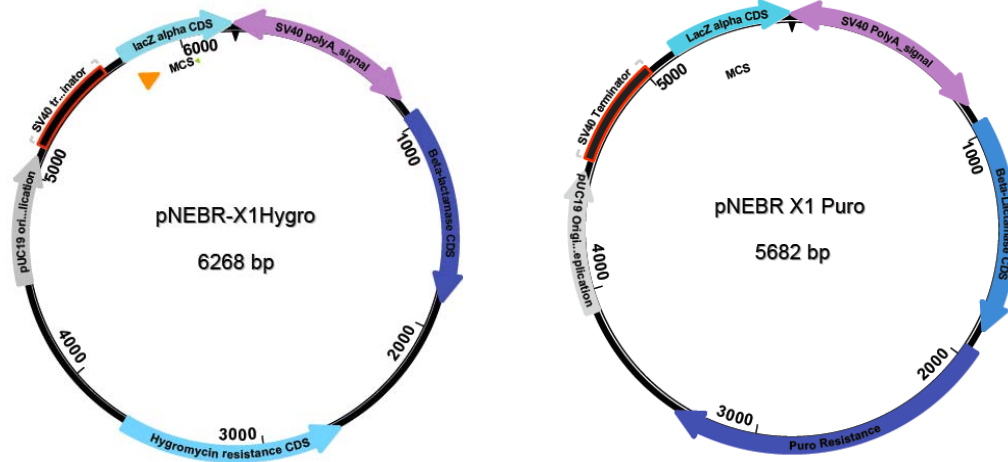


Fig. 15: Backbone of the expression plasmids pNEBR-X1 Hygro and pNEBR-X1 Puro. The expression of the recombinant protein was once again evaluated either by

RT-PCR and western-blot analyses after transfection and induction in the NIH-3T3-47 cells were performed as described above. The expression of Ngn3 is significantly improved as showed by RT-PCR and western-blot analyses (Fig. 16) even though the levels obtained were still below the expectation. A further sequencing analysis performed on the sequence of the Ngn3 did not revealed any mutation or frame-shift that could impair the correct expression of the recombinant protein.

A deeper bioinformatics analysis of the boundary region immediately upstream the MCS of the vector pNEBR-X1-Hygro-Ngn3 revealed the presence of an ATG codon between the promoter region and the starting transcription ATG of the Ngn3-CDS. During the translation process this “proximal” ATG can compete with the legitimate ATG for the transcription machinery producing a transcript with a frame shift that result in no protein expression at all or in expression of a truncated form. The spurious ATG has been removed from the pNEBR-X1-Hygro-Ngn3 vector by restriction digestion, blunting and recircularization of the vector as described in “Materials and Methods” obtaining the pNEBR-X1-Hygro-Ngn3/2 vector. The new construct was used for a time course induction to test the ability and efficiency of the system to be turned off has been evaluated.

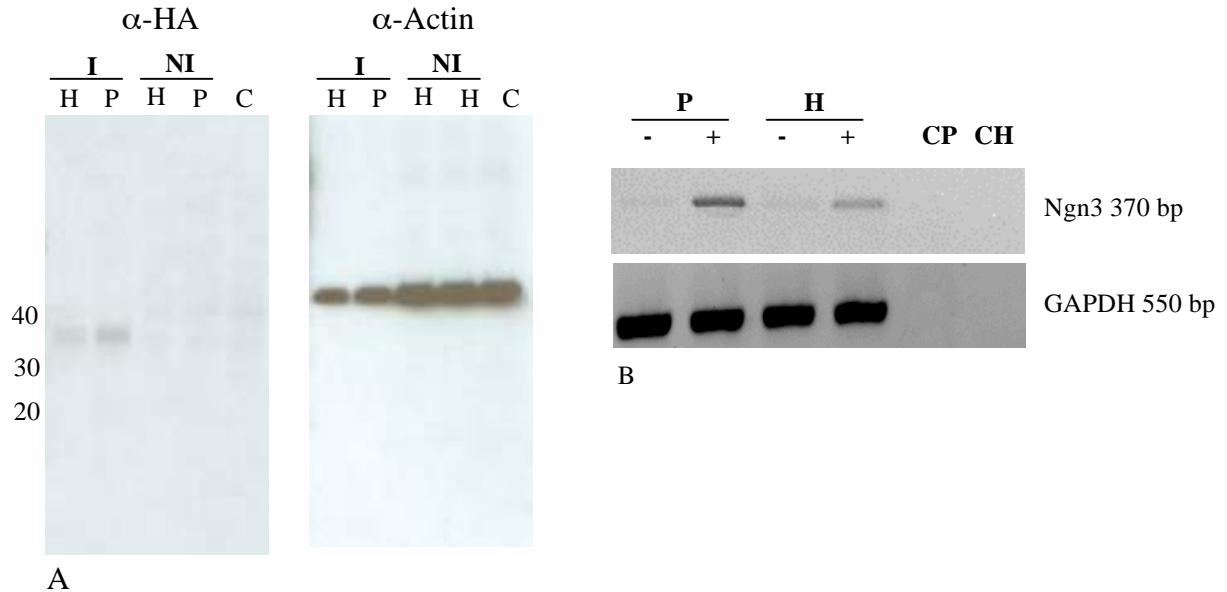


Fig. 16: pNEBR X1 Puro vs pNEBR X1 Hygro. NIH 3T3-47, transfected with Puro or Hygro vector; 12 h induction with RSL1 A) Western blot analysis: H (pNEBR X1 Hygro); P (pNEBR X1 Puro), C (no transfection); I (induction with 500 nM RSL1); NI (no induction). B) RT-PCR analysis: P (pNEBR X1 Puro); H (pNEBR X1 Hygro); C (PCR control, Water). 30 cycles of PCR amplification for GAPDH and non-induced samples, 25 cycles of PCR amplification for the induced samples.

Again, NIH-3T3-47 cells were kept in culture as described and 24 h before the transfection split into three 12-well cell culture plates at a density around 50%. The day after, the cells were transfected as described in the “Materials and Methods”. The three set of samples were activated with 0.5 μ M RSL1 for 12, 18 and 24 h. After that time the first set of samples were subjected to RNA and protein extraction while the others were kept in culture for 12, 18 and 24 h respectively without RSL1 before RNA and protein extraction. The result of the time-course assay is shown in Figure 17. The results show that the protein expression can be finely turned on and off in as much less than 24 h.

Once obtained an efficient tunable expression vector for Ngn3 the next step has been evaluate its activity in the target cells. At least one of the highly inducible F9-R1 clones selected as described above has been transiently transfected with the pNEBR-X1-Hygro-Ngn3/2 as described in “Materials and Methods”. The F9-R1 cells were maintained in culture as described with 300 μ g/ml of G418 and split in a 12-well cell culture plate the day before the transfection at a density around 50%. The day after, the

cells were co-transfected with same amount of pNEBR-X1-Hygro-Ngn3/2 and pCMV- β -Galactosidase. The cells were incubated with the transfection mixture for 8 h and after that period the medium was replaced with regular medium containing 0.5, 1, 1.5, 2 μ M of RSL1. After 12 h of activation the cells were harvested for protein extraction while control samples were kept in culture for further 24 h without activator. The extracts were normalized against protein concentration and β -Galactosidase activity. The same normalized quantities of total protein extract were run on a 12% SDS-PAGE and blotted onto a PVDF membrane for immunodetection as described in the “Materials and Methods”. The results are showed in Figure 18. Interestingly, a significant induction is observed already at very low concentrations of RSL1 with a dose-dependent increment up to 0.5 μ M when no further differences are observed. Again, no protein expression has been detected after 24 h of RSL1 withdrawal.

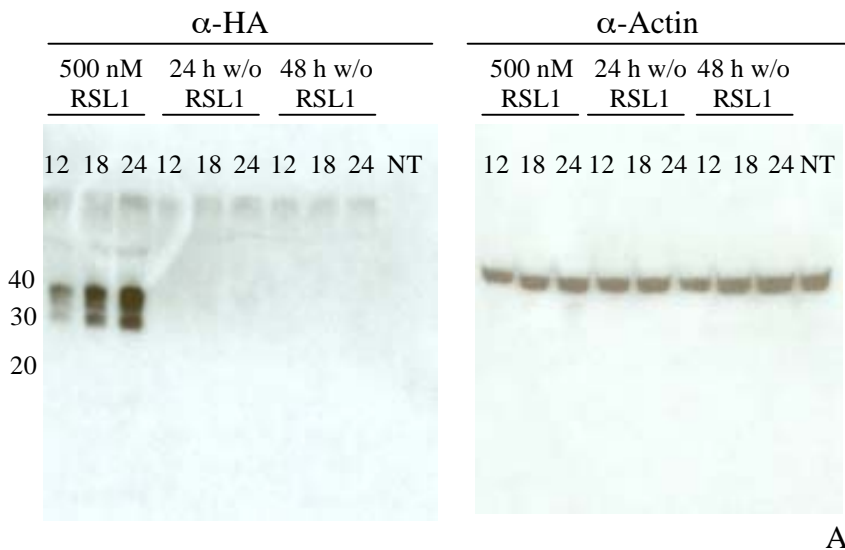
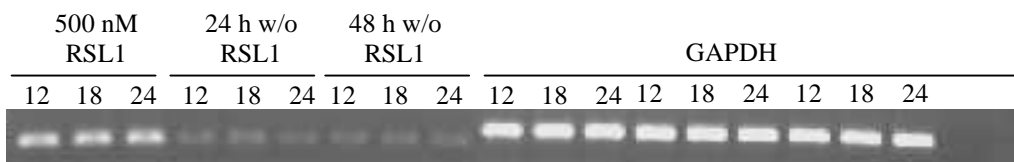


Fig. 17: Time course RheoSwitch on/off. NIH 3T3-47 were transfected with pNEBR-X1-Ngn3-Hygro. Protein and RNA were extracted after 12, 18, 24 h of activation with 0.5 μ M (on), after 24 and 48 h without RSL-1 (off). **A)** western blot assay on total protein extracts using α -HA and α -Actin antibodies. **B)** RT-PCR analysis, 30 cycles of amplification.



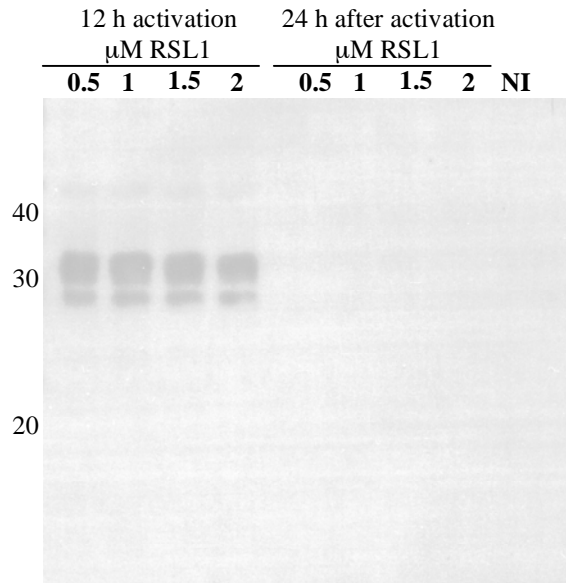


Fig. 18: Dose curve of RSL1 activation. F9-R1 stable clone was transiently transfected with PNEBR-X1-Ngn3-Hygro vector. The cells were induced with 0.5, 1, 1.5, 2 $\mu\text{M RSL1}$ for 12 h before protein extraction. Control samples were kept in culture for additional 24 h without RSL1 before protein extraction. NI: no-induced sample.

Once obtained an efficient tunable expression vector for Ngn3 the next step has been the establishment of a stable F9-R1-inducible cell line. At least one of the highly inducible F9-R1 clones selected as described above has been transfected with the pNEBR-X1-Hygro-Ngn3/2. The F9-R1 cells were kept in culture as described in the “Materials and Methods” Chapter. The day before the transfection the cells were split and plated in a 60 mm cell culture plate at a 60% density. The day after, the cells were transfected with the pNEBR-X1-Hygro-Ngn3/2 vector and selected for hygromycin resistance for at least 10 days. At the moment double-stable F9-R1-Ngn3 are under antibiotic selection.

To improve the RheoSwitch system, a fluorescent reporter gene (dTomato) has been introduced along with an IRES sequence at the 3'-termini of the pNEBR-X1-Hygro-Ngn3/2 vector. This will allow both an immediate detection of the Ngn3 expression during differentiation experiments and an easy purification of pure cell population Ngn3-positive.

To address the fifth point of the aims we relied on the use of both different cells model and methodological approaches. As cell models AR42J and INS-1 have been used. As methodology the ChiP-cloning approach has been used. As an *efficient approach* for

expressing Ngn3, an adenoviral transduction approach that is the fastest and simplest way to obtain specific gene expression has been used.

AR42J and INS-1 cell line has been maintained in culture as described in “Materials and Methods”. The day before infection the cells have been split and plated in a 2 x 100 mm culture dish at a density of about 75%. The days after the cells were infected with about 20 MOI of AdNgn3-EGFP and Ad-EGFP, respectively, for 4 h and then the medium was changed with fresh one. The cells were maintained in culture for two days. The progress and efficiency of infection have been monitored by hrGFP fluorescence. With an MOI of 20, fluorescence appeared one day post-infection and was uniformly distributed to all cells plated with an efficiency of infection near to 100% (Fig. 19).

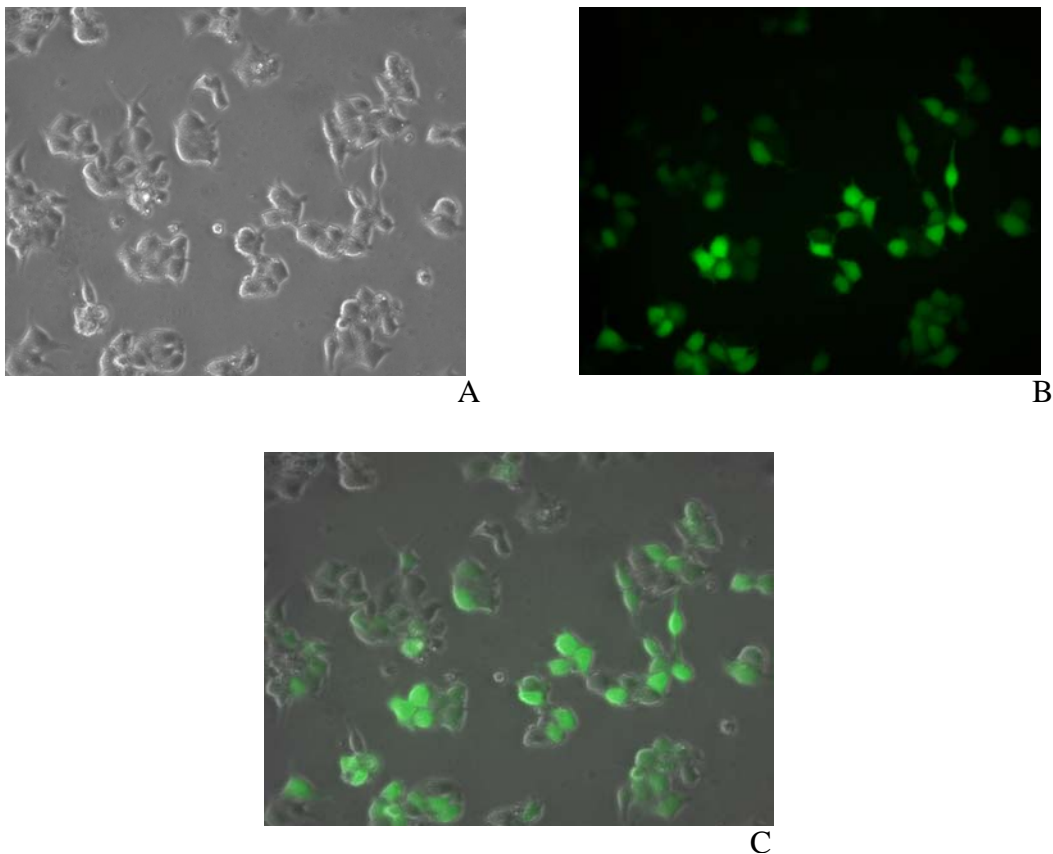


Fig. 19: Control of Ad-Ngn3 infection efficiency. The infection was carried out for 4 h and the cells checked for E-GFP expression 24 h after infection. **A)** AR42J white field. **B)** AR42J E-GFP positive cells. **C)** Merge.

After 48 h post infection no signs of cytotoxicity were observed. At that point the cells were subjected to the ChiP protocol as described in “Materials and Methods” and the DNAs were blunted and ligated into the pCR-BLUNT vector. The ligation reactions were plated onto selective plated and incubated overnight at 37°C. The outputs obtained were genomic libraries with an average of 80 colonies from the AdNgn3-EGFP-treated cells vs. an average of 20 colonies from the control cells treated with the Ad-EGFP alone. Since the pCR-BLUNT vector contains the lethal *E. coli* *ccdB* gene fused to the C-terminus of LacZ α (Bernard et al., 1994) all colonies must contain a cloned DNA fragment allowing for the disruption of the expression of the LacZ α -*ccdB* gene fusion permitting growth of only positive recombinants upon transformation. Cells that contain non-recombinant vector are killed when the transformation mixture is plated. The plasmid DNA from all colonies of both the AdNgn3-EGFP- and Ad-EGFP-infected plates were extracted and sequenced with a specific set of primers (M13 Forward and M13 Reverse) that span the region of the vector containing the inserted DNA fragment. The average size distribution of clone insert from each library was around 840 bp. All the sequences not belonging to the vector, included between two *EcoRI* restriction sites on the pCR-BLUNT vector, were used to perform BLAST analyses using the blastn that compares a nucleotide query sequence against a nucleotide sequence database. The rat genome database Blastn (www.ncbi.nlm.nih.gov/blast/Blast.cgi) has been analyzed using the default filters setting. The output results were analyzed for sequences that span the 5' termini in the proximity of transcription factors and that contain at least one DNA binding motif for the bHLH family of transcription factors (E-box, CANNTG). Among the 80 colonies obtained from the AdNgn3-EGFP-treated cells revealed that about 50% of clones have at least one E-box within the sequences of their insert. On the contrary less than 3% of the colonies deriving from the Ad-EGFP-treated cells showed the presence of E-box within the sequences of their insert.

Among the identified sequences containing E-boxes, between 4 to 5 colonies belonged to the 5' region of three known transcription factors, corresponding to the following proteins:

- 1) The basic Helix-Loop-Helix neurogenic factor 4 (NeuroD4) (n= 4 colonies).

- 2) The SRY (sex determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal) (*Sox9*) (n= 5 colonies).
- 3) The zinc-finger type transcription factor OVO homolog-like 1 (*OVO11*) (n= 5 colonies).

Among the 3% of sequences containing E-box deriving from the Ad-EGFP control, none belonged to transcription factors.

The E-box identified for the *neuroD4* gene is localized about at 18000 bp at the 5' side of the gene. The E-box identified for the *sox9* gene is localized about at 89200 bp at the 5' side of the gene whereas that one identified for the *ovoll* gene maps about at 1000 bp at the 5' side.

The identification of the zinc-finger type transcription factor OVO homolog-like 1 (*OVO11*) with a significant frequency (n=5 colonies) appeared interesting and potentially novel since no data are available regarding an involvement of this transcription factor in pancreas development and differentiation.

To confirm the results obtained by ChiP-cloning approaches, classical ChiP experiments have been carried out as described in “Materials and Methods” but the purified DNA was analyzed by PCR using a specific set of primers (Table 1 Materials and Methods) designed to span the region identified by ChiP-cloning. As positive control a PCR analysis has been performed on the same pool of DNA using a set of primers (Table 1 Materials and MethodsX) encompassing functional E-boxes from the 5' flanking regions of the rat *neuroD/BETA2* gene (data not shown). As showed in Figure 20 the *ovoll* gene promoter is coimmunoprecipitated in the Ngn3-transduced cells using an anti-Ngn3 antibody but not by addition of IgG. Similarly, NeuroD/BETA2 is specifically co-precipitated by the anti-Ngn3 antibody (Fig. 20). *OVO11* and NeuroD/BETA2 are not precipitated in the negative control cells transduced with Ad-GFP (data not showed).

So far neither the expression pattern of *OVO11* nor its function in the pancreas has been addressed directly. Therefore an extensive analysis of the expression of *ovoll* gene in pancreas-related cell line (AR42J and INS-1) and in islets has been carried out by RT-PCR, western blot analyses and immunohistochemistry.

As for RT-PCR, as shown in Fig. 21, a significant signal is present in all the cells analyzed along with rat testis as positive control.

As for western blot the immunoblot analyses using anti-OVO11 antibodies detected specific signals in all the samples analyzed. Interestingly an unexpected pattern of molecular weight distribution has been detected for AR42J (~79 kD), compared with INS-1 cells (~41 kD), islets (~41 kD) and rat testis protein extract used as positive control (~41 kD) (Fig. 22).

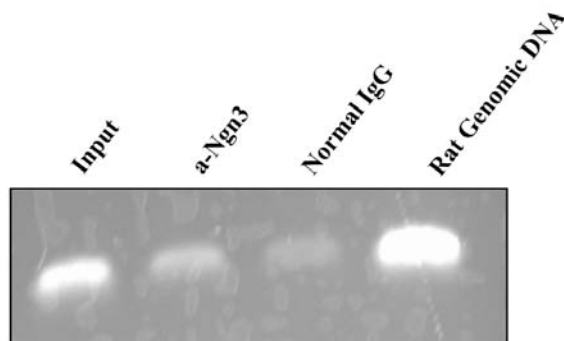


Fig. 20: PCR analysis of ChIP. **Input** (Input fraction); **a-Ngn3** (fraction precipitated using a antibody against Ngn3); **Normal IgG** (fraction precipitated using a generic antibody); **Rat Genomic DNA** (positive control). 30 cycles PCR amplification.

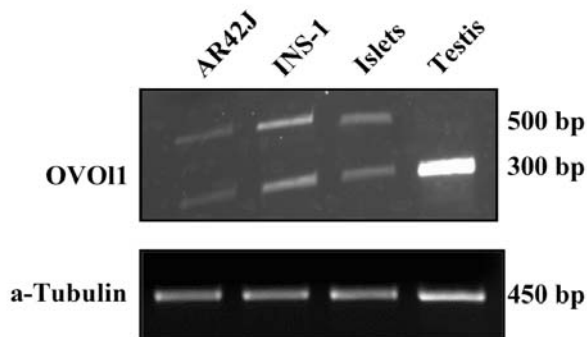


Fig. 21: RT-PCR analysis for the presence of OVO11 protein in pancreatic and testis cell lines. 30 cycles of PCR amplification.

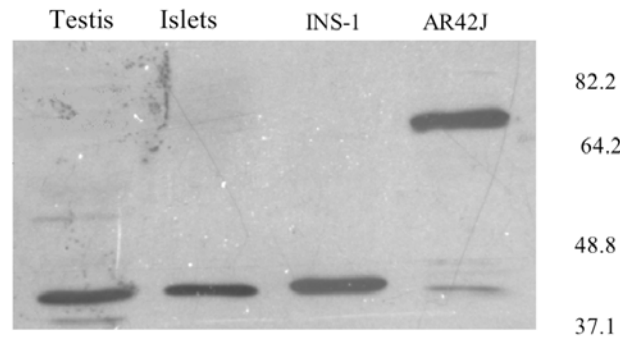


Fig. 22: Western blot analysis using an antibody against Ovov1. For the analysis protein extract from pancreatic cell lines (INS-1, AR42J) as well as from isolated islet were used. As positive control protein extract from rat testis was used.

The subcellular localization of OVOL1 protein has been analyzed by immunohistochemistry on paraffin-embedded sections of rat pancreas. In rat pancreas, OVOL1 is expressed strongly in luminal cells of common pancreatic ducts (CPD, Figure 23A), cells in the evagination of CPD, small ductal and islet cells (Figure 23B & 23C). The expression of OVOL1 is weaker in acini than in ductal or islet cells and, in higher resolution, it is found that the OVOL1 is expressed heterogeneously in acini (Figure 23D, positive-red arrow; negative-yellow arrow).

OVOL1 in rat pancreas

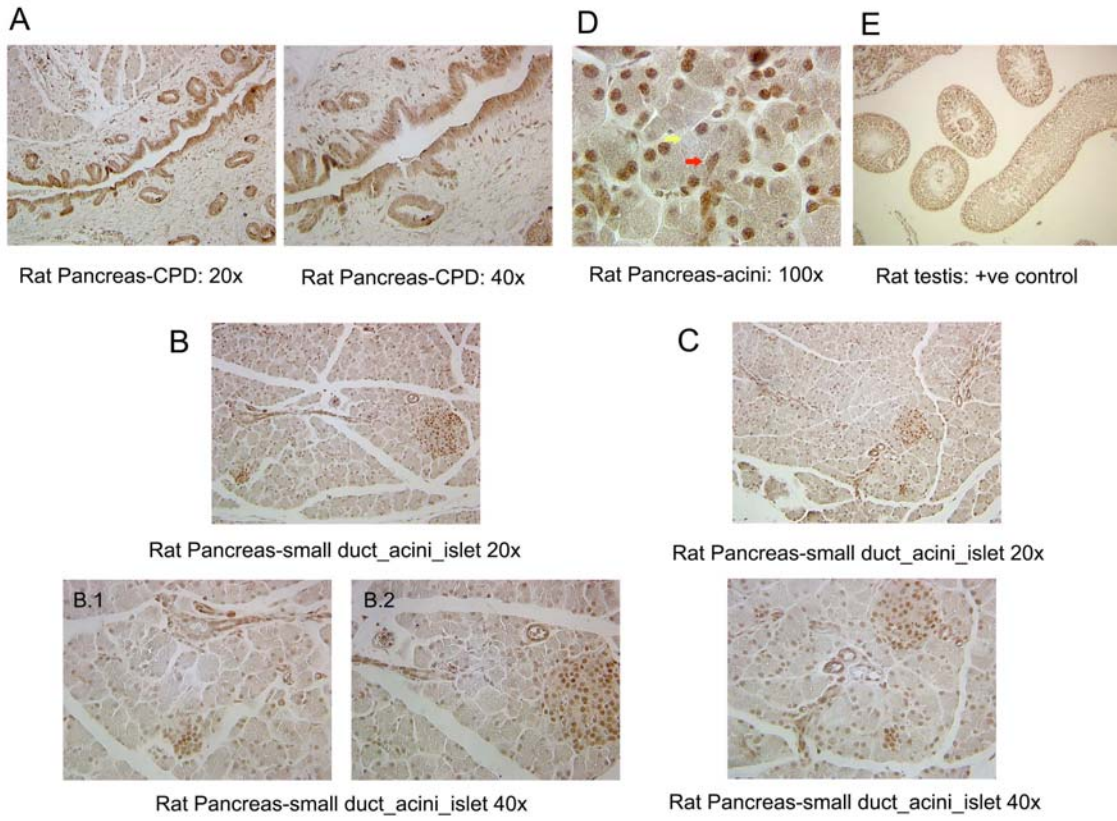


Fig. 23: Cellular localization of the protein Ovov1 in the pancreas. **A)** Luminal cells of the common pancreatic duct. **B)** & **C)** Small duct and islet cells (alternative sections). **D)** Pancreatic acini **E)** Rat testis, positive control

According to the results so far obtained the 5 kb 5' flanking region of the rat *ovoll* gene, retrieved from the available database (www.ncbi.nlm.nih.gov/blast/Blast.cgi) has been analyzed for the presence of E-boxes and other putative transcription factors DNA recognition domains by using MatInspector software. The presence of an E-box at -697 bp (E-box 1) at the 5' region corresponding to that identified by sequence analysis of the ChiP-cloning sample has been confirmed along with the identification of three more putative E-boxes at -1177 bp (E-box 2), -1209 bp (E-box 3) and -1339 bp (E-box 4) at the 5' side, respectively (Fig 24).

```

GAAAGAGTCGCCGAGGCGTGTCCACGTTCTGAGGACAGAGCGTACCAGGGACCACGCGTGGCGTTCCTCAACTACGCCGAAGGTGCGCGGTGT
CCGCCGAGCGGGTGCAAGTCAGGTGCACCGGGATCGGAGAAGCGCTAAGCGAGCGCTTTTCAGCTAGACTGGCGTCCCGCGGGTGTCTCGCGA
AGTCTTCAAGTGCAGCTGGACGGACGGTTCAAGCTTTATTGG

+ 1
CTCGCGGCGGTGTGTGTGGTGGCTCGGGCGGGGGTGGGGTAGGGGTGTTGATCTGATCTAATTAGCTTGGAGTGGCCGAGAGAGCGATGGC
GCATGCGTTGGGAAATAACGGTGACAACCCACCTATTTGTTACTGTGGAACCGGTTTCCATTCGCTGCGGGTGAGGTGGCCTCGCGCCA
GGGCGGGGTGGCCCGACGCGGGGCGAGGGTTGGGGGGCGTGTCTGTGTCACACACCAGAATCACAGGCCGCCCGTACTCCCTGCGCCG
GCGTGTCTCACGTTTGCAAACACTAGAGTTTCGGAGGGAGCGGGCACGGGAAGGGAGGGGCAACGTTATTTGCACCGGCGCTCGCACGCTC
TGTGGGCGATGGCACAAAACACTGGCCCAACAGGCCCGGGCCAGCCCGGCGCCCCGGGCGCCGGGAGAGGCACAAAGCCGCGAGCGGGAAC
CTGCGGTCCGGGCCGAACCCAGCAGCCAGGTATCCCCAAGGCTGGGTGAGCAATCTTGGTTCAGTCGAGAGGTATTTGCTGGGCACCCG
ATTCTGCGGAAAGGCACGCGAGGCGAGAGGAGCGGTGGACCGGAGGCTCTGATGCGCTGTGAATCAAAGCTCGAGCCTGCCTCCGG
TTGCGAGGGCGTTTAAACTGCCTAAGTCCAACCTAATC

E-Box 1
GAGGCTAGCACCTTGGCATATGGGTGGACGTATTGCGCAAGACTCTTCTGGTCAAAAACAGCCGGATTTCAGGTTCTGAAAATCTGGAAG
ACTCTAGGATGAGGACCCTAGCTGAGTGTGATCTGATGCGGTCCTCTTCTGCGGTTGCAACCCGTGTCTCCCTGGCTGGGAGTTTTCAGCAA
GCAGCTCCCATTCACCTGCTGTGTTCTTTCAGCAAGCCTAGTTCGGGCAACGCCCCACCCCCTGAACTGCCGAGAGTGTTCGCTGGCC
TTCTGTGAGCCACTTCACTATGACACTTTGGCCGTGGCTGGTTTCTTTTACTGGCTGACCTATTAAGGCAGGGATGGGACCTGCAGTTA
AAGGCTACTAATATTTTATGATTCCAGACTTTCAGTACACCTGTCTCACATAGTGTAAATGTTTGCTGGAACAGACTGTTGAGGAAA

E-Box 2                               E-Box 3
TTAGTTTGATAAATGAATTAACACAGTTCAAAGTCAAGGAGTCACAGGTGGATACGGTTCATGCTTGTGATCCCAGCACGTGGGAAG
CTGAGGTAGGAAAACCGAGTTAGAGGCCAGCCTATGTTTTCATCGTGAATTCCTGTCTCAAACAATAAAGAATTAAGAAAGGCTGTC

E-Box 4
GGATATGGCTCAGGAGGTATTTGTCACAAGTGGGAAGACCTGAGTTCAACTCCCAGGACCCACAGAGTGGAGGGAGAAAAGTATCCCA
CAACTTGCTCTGACCCGTGGCACACCATGGAACGGTGTGCTGTGT

```

Fig. 24: the 5' -flanking region of the *ovoll* gene. Bolded characters identify the E-Boxes binding elements recognized by transcription factor belonging to the bHLH family.

To determine whether Ngn3 is capable of interact with the *ovoll* promoter, the 1.1 kb sequence encompassing the 5' flanking region from -841 to +230 of the *ovoll* gene and containing the E-box 1 has been PCR amplified using a specific set of primers (Table 1) and cloned into the pCR2.1-TOPO vector as described in “Materials and

Methods”. The *KpnI-SacI* fragment of the pCR2.1-TOPO vector, containing the *ovoll* promoter region, has been then subcloned into the Luciferase reporter vector pGL2-basic as described in “Materials and Methods” to obtain the pGL2-OVO11-Prom1 vector. Mouse Ngn3 has been cloned into pcDNA vector as described in “Materials and Methods” and cotransfected into AR42J cell line along with the pGL2-OVO11-Prom1 vector. As shown in Figure 25 transient expression of Ngn3 induces a significant inhibition of 1.1 kb *ovoll* promoter reporter construct close to 20 % (Fig 25).

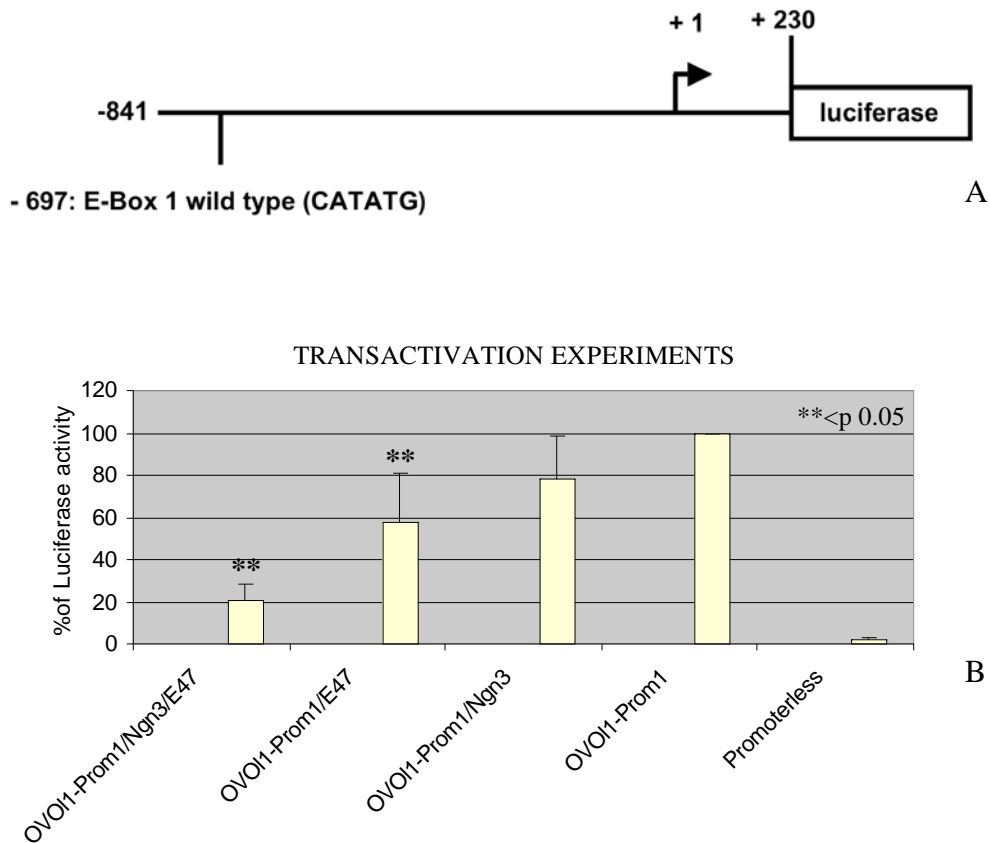


Fig. 25: Transactivation assay. A) Schematic description of pGL2-OVO11-Prom1 reporter construct containing the 1.1 kb 5'-flanking region of the *ovoll* gene. B) Transactivation experiments on AR42J cells cotransfected using different combinations of the plasmids pGL2-OVO11-Prom1, pcDNA-Ngn3 and pcDNA-E47. The results from luciferase assay were performed on three replicates transfection experiments for the statistical significance. Standard deviation has been calculated.

The same cells were also transfected with the same amount of both pcDNA-Ngn3 and pcDNA-E47 vectors alone or in combination; pcDNA-E47 cotransfected along with pcDNA-Ngn3 further increases the inhibition ability of Ngn3 in these cells to 80% (Fig 25) while when transfected alone, it had a slightly inhibitory effect on the *ovoll* promoter (about 40%; Fig 25).

To confirm the importance of the E-box 1, a 1.1 kb promoter sequence containing a mutated form (CATATG → GATAGC) of the E-box 1 has been created. Wild-type and mutated constructs were cotransfected into AR42J cells with expression vectors for Ngn3 and E47. The results indicated that the mutation caused the loss of the Ngn3-mediated inhibition of *ovoll* promoter reporter construct (Fig. 26).

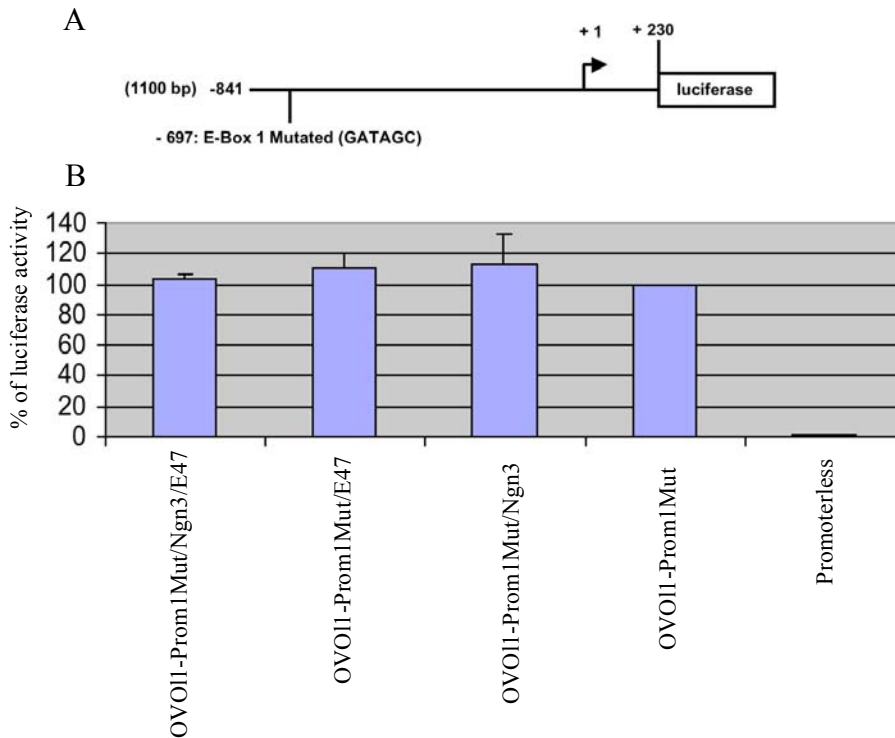


Fig. 26: A) Schematic description of the mutated OVO11-reporter construct. B) Transactivation experiments using the mutated construct pGL2-OVO11-Prom1Mut in AR42J cells were cotransfected with different combinations of : pcDNA-Ngn3, pcDNA-E47, pGL2-OVO11-Prom1Mut. The luciferase activity has been calculated on three different replicate experiments for the statistical significance. Standard deviation has been calculated.

The inhibitory effect of Ngn3 on the *ovoll* gene expression has been quantitatively evaluated by real-time PCR and at level of protein expression, as well, using western blot analyses. AR42J and INS-1 cell lines have been transiently transfected with 1 µg of pcDNA-Ngn3 and 1 µg of pcDNA-E47.

One μg of RNA has been DNase treated and reverse transcribed. Sixty ng of cDNA has been used to perform real-time experiments as described in “Materials and Methods”. The results are reported in Fig 28. Interestingly the inhibitory effect of the Ngn3 has been confirmed for all the cell type analyzed and quantitatively corresponds to a reduction of about 50% of the expression of *ovoll* gene respect to that of not transfected cells or in those transfected with the empty vector (pcDNA) alone.

As for the protein level expression, after extraction the protein content has been determined and 10 μg of total proteins have been used to run a 15% SDS-PAGE. Again, the inhibitory effect of Ngn3 overexpression has been detected (Fig. 27). A quantitative analysis of the bands revealed a reduction, respect the bands obtained from not transfect cells or transfected with the empty vector alone, of about 45% in good agreement with those obtained from quantitative PCR experiments.

Next, the affinity of the pancreatic bHLH protein NeuroD/BETA2 and of the non-pancreatic bHLH protein MyoD alone or in combination with their heterodimeric partner E47 for the E-box 1 of the *ovoll* gene promoter has been tested. The results are reported in Fig. 28. Heterodimers of E47 with NeuroD/BETA2 significantly inhibits the activity of the *ovoll* gene promoter. On the contrary such effect was not shown by the heterodimers of E47 with MyoD (Fig. 29).

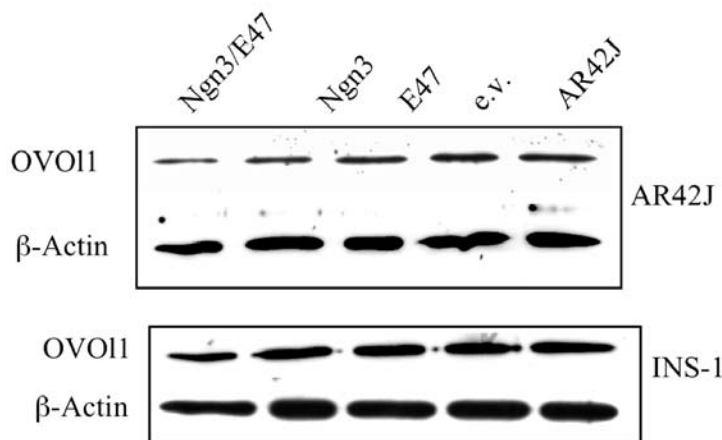


Fig. 27: Inhibitory effect of Ngn3 on *Ovoll* expression. Western blot analysis on protein extracts from the pancreatic cell lines AR42J and INS-1. The cell lines were transfected with different constructs (pcDNA-Ngn3, pcDNA-E47, empty vector pcDNA) alone or in combination. β -actin has been used as control for the quantitative analysis.

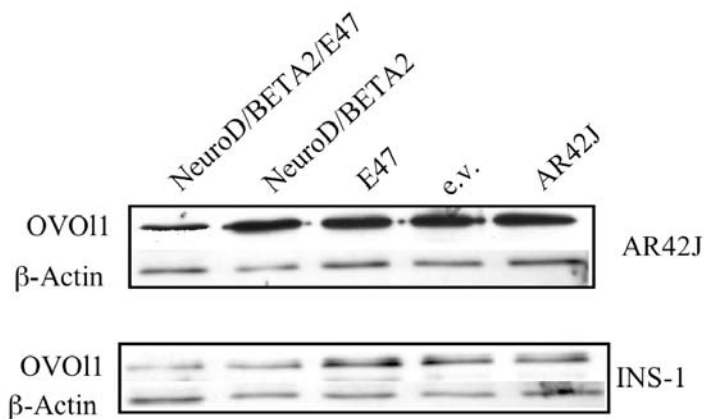


Fig. 28: Inhibitory effect of NeuroD on Ovol1. Western blot analysis on protein extracts from the pancreatic cell lines AR42J and INS-1. The cell lines were transfected with different constructs (pcDNA-NeuroD/BeTA2, pcDNA-E47, e.v. empty vector pcDNA) alone or in combination. β -actin has been used as control for the quantitative analysis

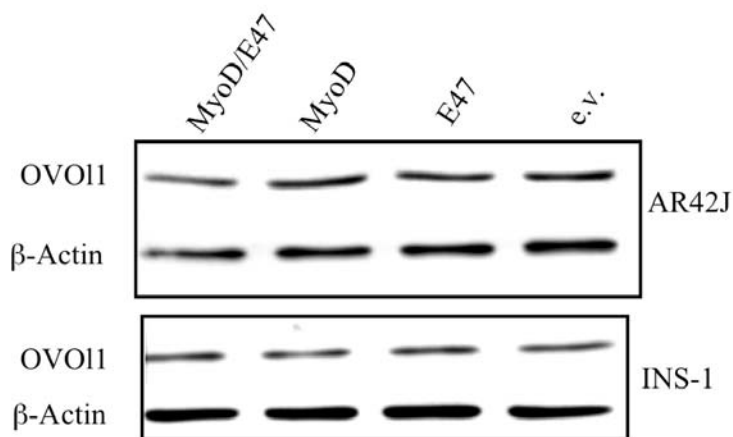


Fig. 29: Inhibitory effect of NeuroD on Ovol1. Western blot analysis on protein extracts from the pancreatic cell lines AR42J and INS-1. The cell lines were transfected with different constructs (pCMV-MyoD, pcDNA-E47, e.v. empty vector pcDNA) alone or in combination. β -actin has been used as control for the quantitative analysis

Once again the effect of bHLH proteins NeuroD/BETA2 and MyoD were analyzed by real-time PCR and western blot analyses. AR42J and INS-1 cell lines have been transiently transfected with 1 μ g of pcDNA-NeuroD/BETA2 or pCMV-MyoD alone or with and 1 μ g of pcDNA-E47. RNA and proteins have been extracted.

One μ g of RNA has been DNase treated and reverse transcribed. Sixty ng of cDNA has been used to perform real-time experiments as described in “Materials and Methods”. As reported in Fig 30 the inhibitory effect of the NeuroD/BETA2 has been confirmed for all the cell type analyzed and quantitatively corresponds to a reduction of about 50% of the expression of *ovoll* gene respect to that of not transfected cells or in

those transfected with the empty vector (pcDNA) alone. No effects have been detected using MyoD protein both alone and in combination with E47 partner (Fig 30).

As for the protein level expression, again, the inhibitory effect of NeuroD/BETA2 overexpression has been detected (Fig. 28). A quantitative analysis of the bands revealed a reduction, respect the bands obtained from not transfected cells or transfected with the empty vector alone, of about 45% in good agreement with those obtained from quantitative PCR experiments. No effects have been observed using MyoD protein both alone and in combination with E47 partner (Fig. 30).

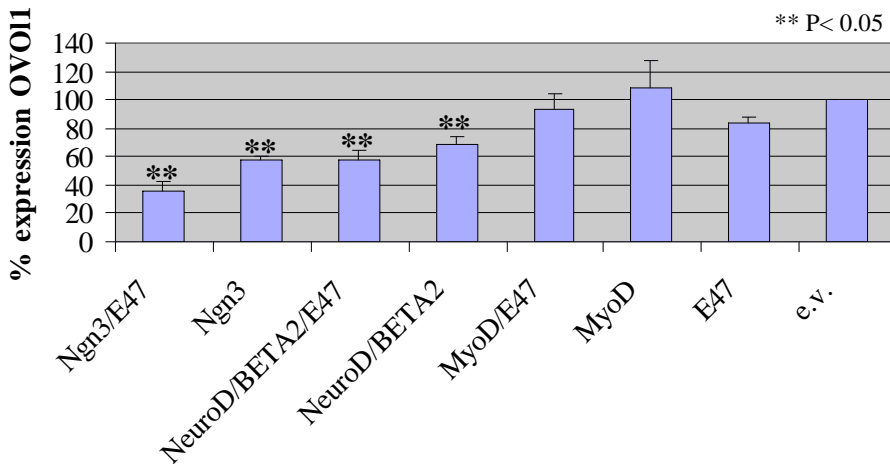


Fig. 30: Real-time analysis on the inhibitory effect of Ngn3, NeuroD/BETA2, MyoD on Ov011. The AR42J and INS-1 cell lines were transfected as described with expression plasmid (pcDNA-Ngn3, pcDNA-NeuroD/BETA2, pCMV-MyoD, pcDNA-E47, e.v. empty pcDNA vector) alone or in combination.

The ability of NeuroD/BETA2 to interact with OV011 promoter has been evaluated *in vivo* performing ChiP analysis as described in “Materials and Methods”. AR42J cell line has been and plated in 2 x 100 mm culture dish at a density of about 75% and with about 20 MOI of AdNeuroD-EGFP and Ad-EGFP. After 48 h post infection the cells were subjected to the ChiP protocol as described in “Materials and Methods” and the DNAs were analyzed by PCR using a specific set of primers (Table 1 Materials and Methods) designed to span the region identified by ChiP-cloning. As showed in Figure 31 the *ov011* gene promoter is coimmunoprecipitated in the NeuroD/BETA2-transduced cells

using an anti-NeuroD/BETA2 antibody but not by addition of IgG. OV011 is not precipitated in the negative control cells transduced with Ad-GFP (data not shown).

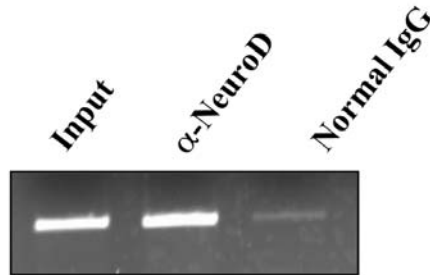


Fig. 31: NeuroD/BETA2 interacts with the promoter region of *ov011* gene. AR42J cells were infected using Ad-NeuroD-EGFP and Ad-EGFP alone or in combination. The ChIP were performed using antibody against NeuroD/BETA2 or generic antibody as control. The PCR analysis was performed using *Ov011*ChIP primers listed in the Material and Methods. 30 cycles of PCR amplification.

4.3 Purification and Characterization the β -Cell Insulin Transcription Factor A2.2 (All information are unpublished proprietary of Joslin Diabeted Center)

The *insulin* gene is specifically expressed in β -cells of the islets of Langerhans and its transcription is regulated by the glucose level. The molecular mechanisms of the β -cell restricted expression of insulin has been extensively studied during the years and leads to the identification of different *cis*-regulating elements on its promoter. Among those, three of them appear exclusively restricted to the pancreas: E1, A3 and RIPE3b/C1. The transcription factors binding to the E1 and A3 sequences have been identified as NeuroD/BETA2 and Pdx1/IPF1/STF1/IDX1/GSF/IUF1, respectively. Previous studies have reported that a β -cell specific nuclear factor, recently identified as the mammalian homologue of avian MafA (Olbrot, Rud et al. 2002), binds to the conserved *cis*-regulatory element RIPE3b/C1, spanning between the nucleotides -139 and -101 of the insulin promoter region. Interestingly it has been showed (Harrington and Sharma 2001; Nishimura, Salameh et al. 2005) that the RIPE3b/C1 (also know as Maf Response Element, MARE element,) binding sequence is composed of distinct overlapping elements. In fact the MARE element is recognized not only by MafA but also by transcription factors binding to an A-element that are present in the insulin-specific MARE sequence and named A2. Three binding complexes have been found until now, named A2.1, A2.2 and A2.3, respectively. Among the three, only A2.2 has been demonstrated to be, like MafA, β -cell specific. Interestingly, A2.2 element acts enhancing the activation potential of MafA, and not its binding activity (Nishimura, Salameh et al. 2005).

Despite the importance for the cell-specific role played by A2.2 in the regulation of *insulin* gene expression its nature is still elusive. According to that it is clear that the fully understanding of insulin regulator elements could have an important impact on possibility to control the *insulin* gene expression as well as the pancreatic development and development.

For those reasons the aims of the present part of the Thesis have been:

- 1) *To characterize and purify the transcription factor A2.2*
- 2) *To identify it as β -cell specific transcription factor*

4.3.1 Characterization of the Transcription Factor A2.2

To address the first aim, EMSA assays have been set-up. Oligonucleotides spanning different region of the rat insulin II promoter region and mutated in the sequence specific for A2.2 have been designed. These sequences have been used for EMSA competition experiments with HIT T-15 nuclear extracts.

The mutated sequences that were unable to compete with the binding activity of the HIT T-15 nuclear extracts toward its oligonucleotide target allowed for the identification of the nucleotides critical for the binding properties. The sequence retrieved was used to perform analyses using the TRANSFAC program that compares a nucleotide query sequence against a nucleotide sequence database. The rat genome database has been analyzed using the default filters setting. The output identified a high homology for bZIP binding elements. In particular the search for transcription factors binding the region of interest identified the C/EBP family as a possible candidate. To confirm the involvement of bZIP transcription factors a set of EMSA competition assays have been performed using HIT T-15 nuclear extract and specific oligonucleotides (Table 1 Materials and MethodsX) containing bZIP binding motifs. For the supershift assays antibody anti-C/EBP α and anti-C/EBP β have been used. The results are reported in Figure 32.

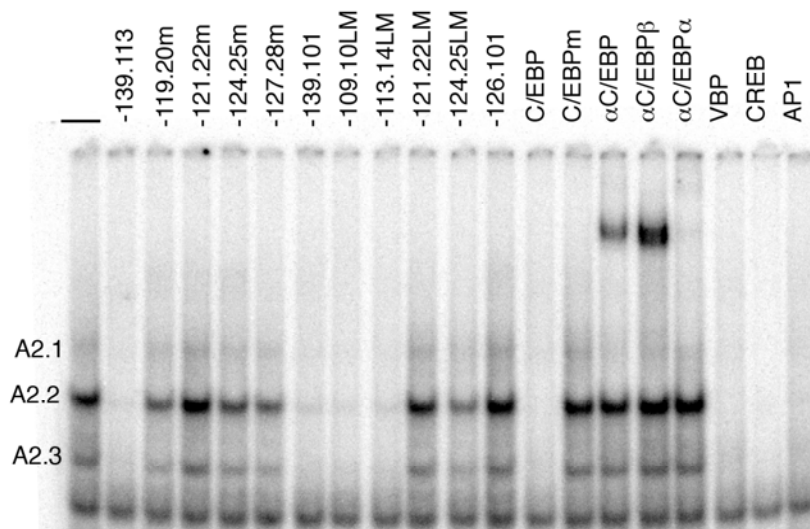


Fig. 32: EMSA analysis. A2.2 wild-type and mutated nucleotide sequence, C/EBP, AP1, VBP, CREB nucleotide sequences were used as cold-competitors. Super shift assay was performed using antibodies against the members belonging to the C/EBP transcription factor family.

As shown (Fig. 32), the bZIP cold-nucleotide sequences, used as competitors, are able to inhibit the formation of the A2 complexes, strongly supporting the hypothesis that A2-binding transcription factor could belong to the bZIP protein family. Unfortunately the super-shift experiments demonstrated that, even though the C/EBP β is part of the complex binding the A2, it does not correspond to the A2.2 transcription factor. In fact even if the presence of the anti- C/EBP β antibody induces a strong supeshift, the signal of the A2.2 complex (Fig. 32, lane 16) has the same intensity of that resulting from the the positive control (Fig. 32, lane 1). Furthermore the same EMSA assays performed in presence of recombinant dominant negatives of different bZIP subfamily members did not show different results (Dr A. Sharma, personal communication). These observations seem suggest that a novel bZIP protein could be involved in the binding the A2 element.

4.3.1 Purification of A2.2 Transcription Factor

The process used for the purification of the A2.2 factor is outlined in Fig. 33. The HIT T-15 protein extracts were first purified on a HiTrap-Heparin column. This column allows for fast, simple and easy separations of nucleic acid binding proteins. Immobilized heparin interacts with nucleic acid-binding proteins as a cation exchanger due to its high content of anionic sulfate groups mimicking the polyanionic structure of the nucleic acid.

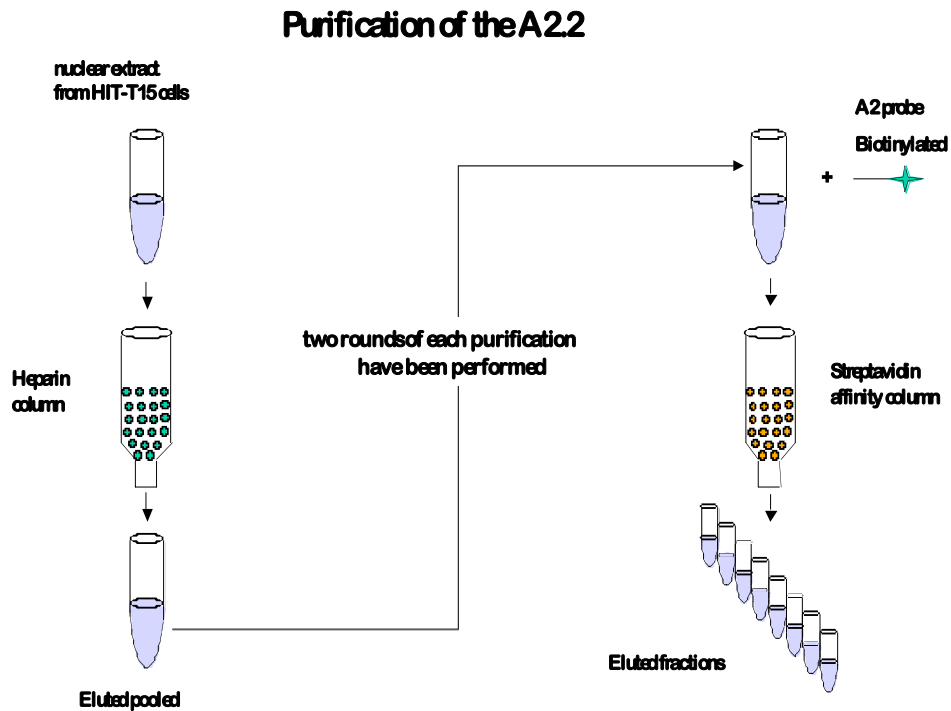


Fig. 33

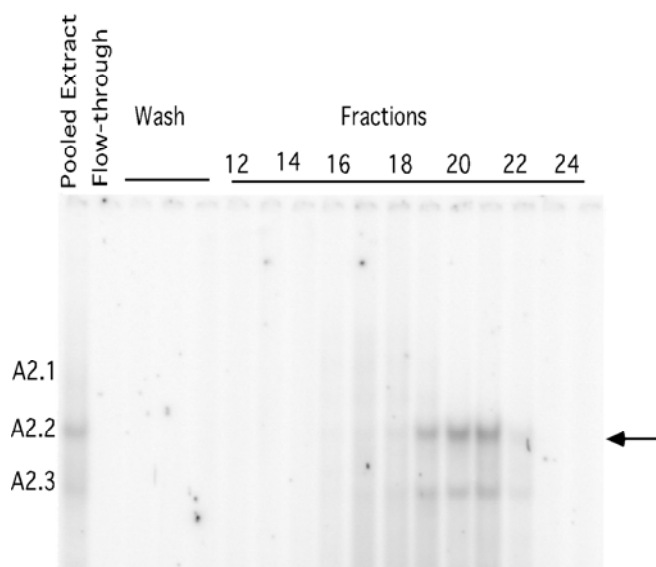


Fig. 34: EMSA assay on purified eluted fractions. Single eluted fractions were assayed for the binding activity using the A2.2 probe. A2.2 complex is highlighted by the arrow.

After extensive washing the bound proteins were eluted with an increased salt gradient and pooled. The eluted pooled fraction was added with a biotinylated A2 oligonucleotide and subjected to an affinity chromatography step on a streptavidin-agarose column. After extensive washing the proteins were eluted using a solution with increased ionic strength and collected in 30 separate fractions. Each fraction was analyzed for their binding activity by EMSA assay, using the A2 probe, as described before. The result is shown in Fig 34.

The A2.2 complex was not present in the flow-through and became detectable only within the fractions from 19 to 21.

The protein concentration was normalized to 2 $\mu\text{g/ml}$ and 20 μl from each fraction were run on 12% SDS-PAGE gel and then blotted on a nitrocellulose membrane. The nitrocellulose membrane was stained, as described in “Materials and Methods”, using colloidal-gold staining. The results are shown in Fig. 35. The fractions 19, 20, 21 showed a specific pattern compared to the control samples. In particular, as indicated by the arrows in Fig. 35, are present low molecular weight bands (corresponding to about 15-20 kDa). The bands from the lanes 16, 19 and 20 were eluted from the membrane and checked for their binding activity by EMSA assay (Fig. 35).

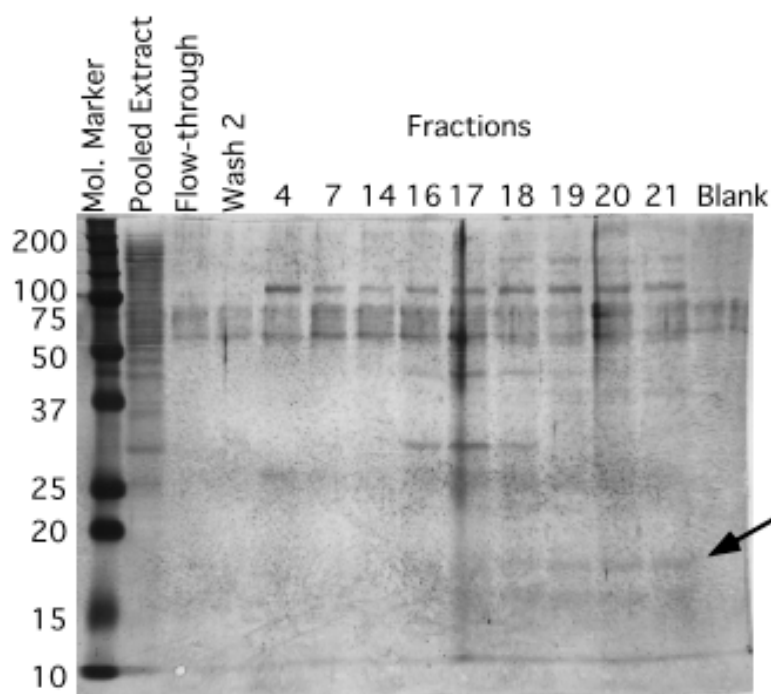


Fig. 35: Western blot on selected eluted fraction. Selected eluted fractions were blotted onto a PVDF membrane and stained using colloidal gold staining procedure. Unique bands on band shift positive fraction are highlighted by the arrow.

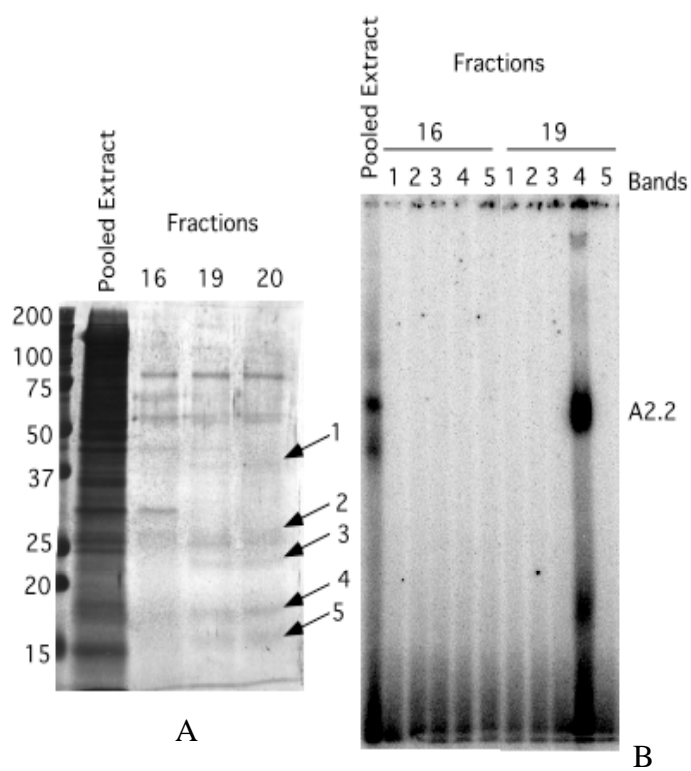


Fig. 36: Western blot and EMSA assay on eluted fractions. A) Low molecular weight bands were eluted from PVDF membrane B) five bands, from fraction 16 and 19, were tested for the binding activity using the A2.2 probe.

The result from the EMSA assay is shown in Fig. 36B. As showed only the eluted-band corresponding to the sample 4 was able to restore the binding activity. This result is in full agreement with the hypothesis that the A2.2 complex could be mediated by a low molecular weight protein (Dr. A Sharma personal communication).

The purified protein from the sample 4 has been subjected to mass spectrometry analysis at the Proteomic Core of the Joslin Diabetes Center. The results are reported in Fig. 37. As showed, among the proteins identified, the small Maf factors showed the higher score.

Reference Scan(s)	Peptide	MH+	z	P (pro) P (pep)	Score XC
IPI00117877.1	Transcription factor MafG			9,92E-12	20,22
2511	R.EPGENGTSLTDEELVTM*SVR.E	2180,00742	2	9,92E-12	4,37
2524	R.EPGENGTSLTDEELVTM*SVR.E	2180,00742	2	4,34E-06	4,05
2744	R.GPLAAGLGPLVPGK.V	1246,75232	2	3,35E-08	4,13
2753	R.GPLAAGLGPLVPGK.V	1246,75232	2	1,85E-06	3,88
IPI00126148.1	Transcription factor MafK			3,54E-09	10,26
IPI00121073.1	Cold-inducible RNA-binding protein			1,15E-07	20,24
IPI00115831.1	Histone deacetylase complex subunit SAP18			2,48E-07	20,18
IPI00225633.2	Activated RNA polymerase II transcriptional coactivator p15 precursor			6,66E-06	30,23
IPI00133583.1	Adult male kidney cDNA, RIKEN full-length enriched library, clone:0610012F23 product:			1,13E-03	10,20
IPI00355031.5	PREDICTED: cortactin binding protein 2 isoform 1 DNAJ			1,98E-03	10,10
IPI00652785.1	Adult male corpus striatum cDNA, RIKEN full-length enriched library, clone:CU30004P21			2,25E-03	14,15
IPI00624384.1	4 kDa protein			1,51E-02	10,13
IPI00620640.2	PREDICTED: similar to WW domain binding protein 2			1,63E-02	2,16
IPI00131373.1	Protein Wnt-4 precursor			1,93E-02	8,14
IPI00677064.1	PREDICTED: hypothetical protein XP_920343			2,16E-02	8,10

Fig. 37: Partial list of the identified proteins from the tandem-MS analysis of the 17 kDa band extracted from the PVDF membrane.

To confirm the presence of small Maf factors within the A2.2 complex, western blot analysis using antibody anti-MafK/G/F have been performed. For the western blot analysis nuclear protein extracts from MIN-6, small Maf factors from HeLa and a small aliquot of the concentrated sample from the 17 kDa band have been used. After TCA precipitation the purified nuclear extracts were run on 15% SDS-PAGE gel and blotted to a PVDF membrane. The membrane was cut horizontally in correspondence of the 25kDa protein of the ladder. The upper portion was incubated with anti-MafA while the lower part was incubated with antibody anti-small Maf factors (MafK/G/F). The results of the western blots are shown in Fig. 38.

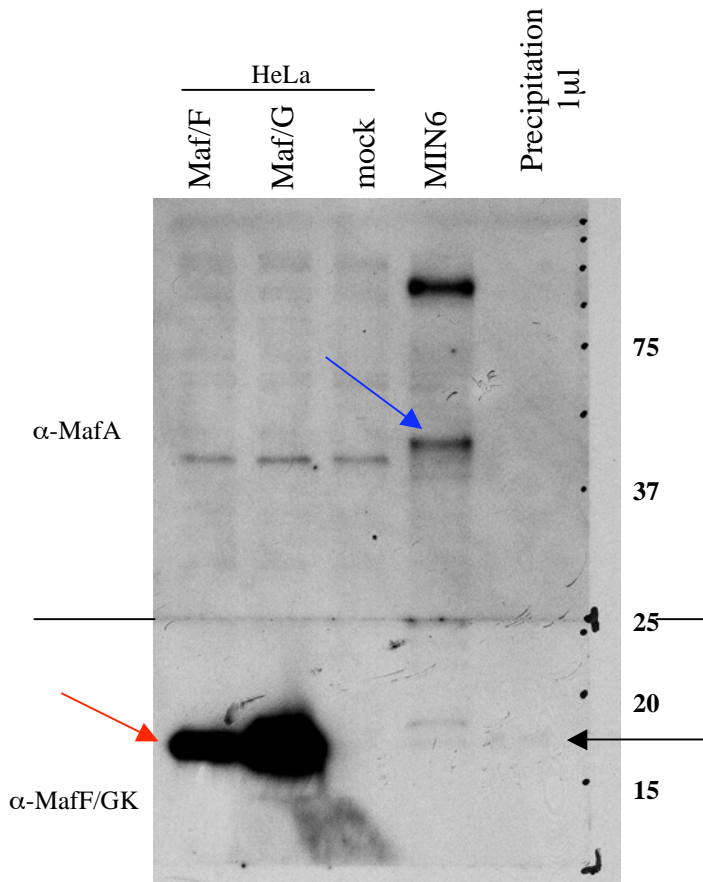


Fig. 38: Western Blot analysis. Nuclear extract from HeLa, MIN6 cell line and from proteins extracted from the 17 kDa band were blotted onto PVDF membrane and stained with antibodies against MafA and small Maf family members. Blu arrow: MafA. Red arrow: small Maf factors. Black arrow: small Maf from eluted band.

The presence of a faint but significant signal for the small Maf proteins is in agreement with the mass spectrometry analyses and identifies small Maf as members of the A2.2 complex.

To test the possibility that small Maf factors could correspond to the A2.2 factor, EMSA and super-shift assays, using small Maf factors from HeLa and HIT-T15 nuclear extract, have been done. The results are reported in Figure 36. As shown in Fig. 39 the oligonucleotides sequences from -139 to -101, -139 to -113, -126 to -101, were used as probes. Small Maf were not able to form the A2.2 binding complex. Furthermore, the antibody used was not able to super shift the A2.2 complex. Along with these results, small Maf factors are present also in cell lines not expressing A2.2 complex, as shown by the western blot assay. Standing the condition that the A2.2 factor “is” β -cell specific it

must be excluded that small Maf is the A2.2 factor *per se* but cannot be excluded, as proved here, that small Maf could interact with it for example as dominant inhibitor.

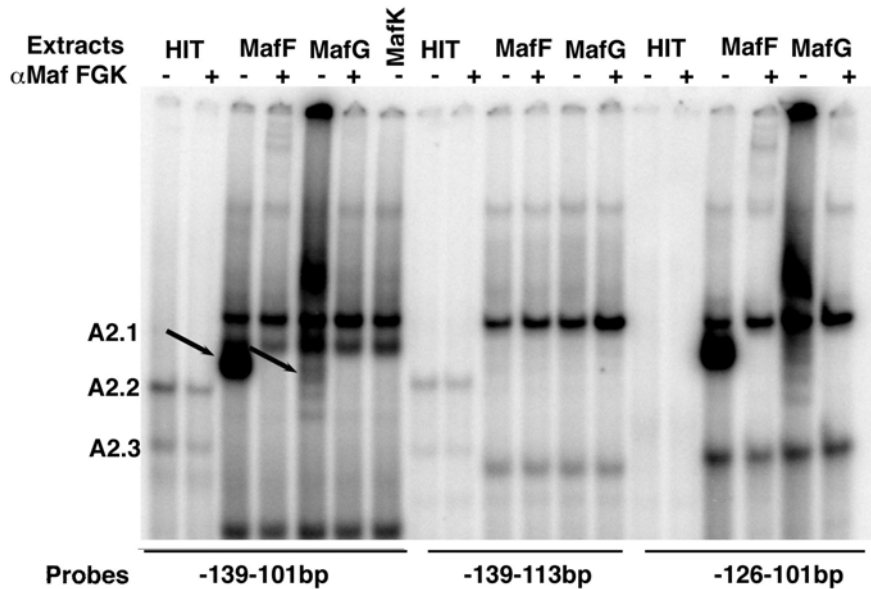


Fig. 39: EMSA and super shift assays. EMSA assay were performed using the -139-101, -139-113, -126-101 probes. Super shift assay was performed using a antibody against small Maf family members.

To select possible A2.2 candidates among the hundred proteins identified by mass spectrometry, a deeper bioinformatics analysis has been performed. The output of the mass spectrometry analysis has been analysed using the database listed in the Materials and Methods, using more stringent limits such as a nuclear localization and a low molecular weight. The possible candidates are reported in Figure 40.

Interestingly at least one candidate, the BATF belongs to the bZIP transcription factor family. The expression of all these potential candidates has been evaluated by RT-PCR. The RNA from α -TC, MIN-6, HeLa cell lines has been extracted using Trizol reagent as described in “Materials and Methods”. One μ g of RNA was DNase treated and reverse transcribed. The level of expression of the above factors was determined by PCR using specific sets of primers (Table 1 Materials and Methods). The results are reported in Fig. 41. None of the above factors were β -cell specific.

IPI00121073.1 Cold-inducible RNA-binding protein
IPI00115831.1 Histone deacetylase complex subunit SAP18
IPI00225633.2 Activated RNA polymerase II transcriptional coactivator p15 precursor
IPI00652785.1 Adult male corpus striatum cDNA, RIKEN full-length enriched library
IPI00620640.2 PREDICTED: similar to WW domain binding protein 2
IPI00323130.3 Transcription elongation factor B polypeptide 1
IPI00662675.1 PREDICTED: similar to High mobility group protein 1
IPI00678394.1 PREDICTED: similar to Wilms tumour 1-associating protein
IPI00662975.1 PREDICTED: PHD finger protein 20-like 1 isoform 11

Fig. 40: Selected A2.2 candidates from the proteins identified by MS analysis.

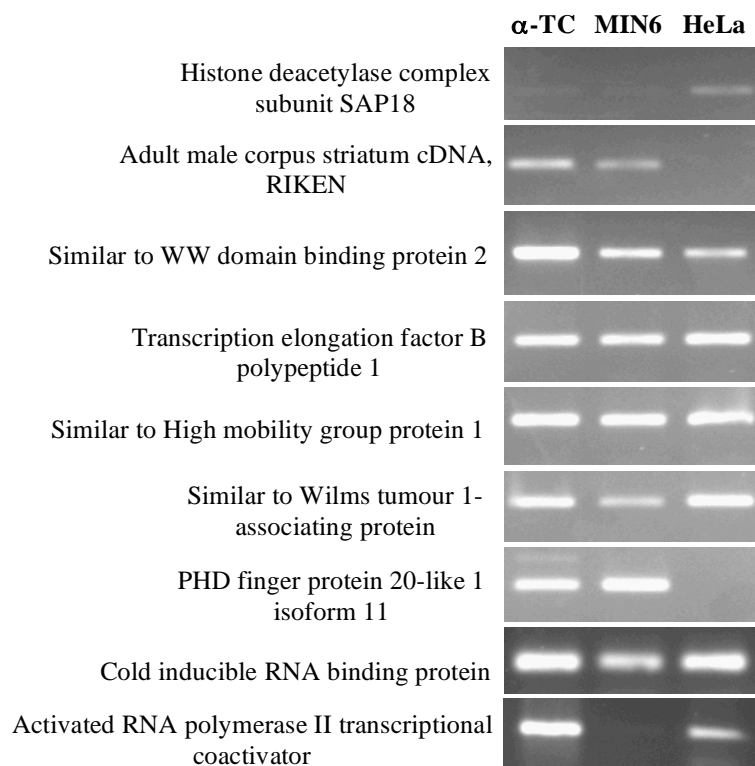


Fig. 41: RT-PCR analysis of the A2.2 candidates. For the analysis RNA extracted, as described in Materials and methods, from HeLa, MIN6 and α -Tc cell lines was used. The results are from 30 cycles of amplification

Loosing further the filters, i.e considering also the proteins with the lowest homology score, and extending the number of databases analysed (see “Materials and Methods”) allowed to indentify another list of possible candidates (Fig. 42). Again the expression level of all these potential candidates has been evaluated in α -TC, MIN-6, and HeLa cell lines by RT-PCR as described above. The results are shown in Fig. 43.

Once again none of the factors identified were β -cell specific.

BATF	basic leucine zipper transcription factor
p21SNFT	Jun dimerization protein
HMGB4	high-mobility group box 4
Csdc2	cold shock domain containing C2, RNA binding
RNABP8	RNA binding protein 8
C1D	nuclear DNA binding protein

Fig. 42: Selected A2.2-candidates after the new bioinformatic analysis.

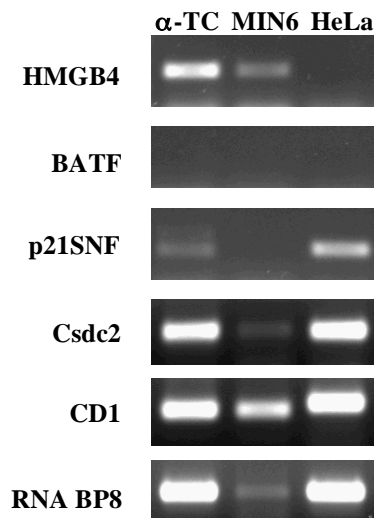


Fig. 42: RT-PCR analysis of the A2.2 candidates. For the analysis RNA extracted, as described in Materials and methods, from HeLa, MIN6 and α -Tc cell lines was used. The results are from 30 cycles of amplification

I would like to thank:

Dr. Michael Li Wan Chu for the OVOL1 staining

Dr. Takuma Kondo and Dr. Ilham El Khattabi for the characterization and isolation of the A2.2 transcription factor.

Chapter 5

DISCUSSION

5 DISCUSSION

There is a growing research interest in the potential of pancreatic stem cell in the treatment of type 1 diabetes. Stem cell-derived β -cells may solve the problem of an inadequate supply of producing β -cells. In this respect it should be stressed that the full understanding of the hierarchy of the molecular events involved in pancreas differentiation is probably necessary for achieving therapeutic success of ES cells treatment. During embryogenesis, transcription factors orchestrate intricate pathways of differentiation by regulating the rate of transcription of an array of genes. Genetic studies have begun to unravel the complex cascade of factors that control the proliferation and differentiation of cells in the developing of pancreas. Some of them are expressed very early, such as Hlxb-9 (Harrison, Thaler et al. 1999) and Pdx-1 (Jonsson, Carlsson et al. 1994). Other transcription factors have a role in the differentiation of specific pancreas cell types. Pax4 (Sosa-Pineda, Chowdhury et al. 1997; Mansouri, Goudreau et al. 1999; Mansouri, St-Onge et al. 1999; Dohrmann, Gruss et al. 2000; Sosa-Pineda 2004) and Pax6 (Ahlgren, Pfaff et al. 1997; St-Onge, Sosa-Pineda et al. 1997; Hill, Asa et al. 1999; Mansouri, Goudreau et al. 1999; Dohrmann, Gruss et al. 2000), Isl1 (Ahlgren, Pfaff et al. 1997), Nkx6.1, Nkx2.2 (Sussel, Kalamaras et al. 1998; Sander, Sussel et al. 2000), NeuroD (Naya, Huang et al. 1997) and Ngn3 (Gradwohl, Dierich et al. 2000) are all necessary for the differentiation of the islet endocrine cells. Nevertheless the specific pathway leading to the development of the insulin-secreting β -cells is particularly important to be able to engineer new β -cells from ES cells and/or progenitor cells. The aim of the first part of this Thesis work has been *collect as much as possible information regarding the pathways leading to the formation of insulin producing cell from undifferentiated cells*. To obtain the goal of the research are necessary

1. A simple cell model for the differentiation.
2. A high throughput experimental approach.

As cell model for the differentiation, the cell line AR42J has been chosen. AR42J cells were originally derived from a chemically induced rat pancreatic tumor and posses

both exocrine and neuroendocrine properties. In the presence of Activin-A (Act-A) plus Hepatocyte Growth Factor (HGF), these cells differentiate into insulin-expressing cells.

In particular the clone named B13 possesses an efficiency >of 90% to differentiate into insulin producing cells when treated with Act-A and HGF (Yasuda, Tanaka et al. 1994; Mashima, Ohnishi et al. 1996; Mashima, Shibata et al. 1996; Ogihara, Watada et al. 2003). For the gene expression and proteomic analyses, microarray and 2D-MS approaches have been used.

Although, the AR42J-B13 cell line is, at the present, the best cell model to mimic as much as possible the physiological differentiation of β -cells it shows some phenotypic “instability” associated with spontaneous differentiation in absence of growth factors stimuli. To recover the original phenotype it has been necessary to expand new colonies from single cells. The colony showing the same characteristics of the original AR42J-B13 clone was used to perform microarray and 2D-MS analysis.

The analysis of the microarray has been challenging due to the platform used for the scanning and the software used for the statistical analysis. In fact, the raw data deriving from the scanning had had to be extensively manipulated to adapt them to the software platform. Moreover, open-source platform requires a deep acquaintance of the environment system use, e.g. **R**-language. Once ready the data were analyzed using two platforms: TIGR-MIDAS/TM4 and R-mArray/S.A.M.-Excel. Although, the results from the two independent platforms were comparable only few genes (three genes) over- and (sixty genes) under-expressed were identified. None of them were specifically related to a pancreatic differentiation event and none of them were transcription factors. Many of the genes identified were, instead, putative transcripts retrieved from the Riken database whereas some others were identified only on the basis of a certain degree of homology with known sequences. Among the under-regulated genes, some were related to neuron phenotype, e.g. *gdpd5* (*glycerophosphodiester phosphodiesterase domain containing 5*) (Yanaka, Nogusa et al. 2007) and *myo1b* (*myosin 1b*) (Ruppert, Kroschewski et al. 1993), just to cite two of them. These results, while disappointing, were not unexpected since the AR42J cells, as stated above, possess neuroendocrine properties and even during the growth factors-mediated differentiation they share with the neurons some molecular traits (Scharfmann 1997). Furthermore it is known that Act-A can induce conversion of these

cell into neuronlike cells (Ohnishi, Ohgushi et al. 1995). Among the gene showing a down-regulation it is worth to mention the amylase. Once again it is important recall the in the undifferentiated state the AR42J shows characteristic of acinar cells. One of the markers of the acinar phenotype is the expression of amylase (Mashima, Ohnishi et al. 1996). The reduction of the expression of the amylase is a direct effect of the presence of Act-A (Yasuda, Tanaka et al. 1994) but it is also an important signal of the starting of a differentiation process. In fact it has been observed in many (trans)differentiative events of the AR42J such as the differentiation into insulin secreting cells (Mashima, Shibata et al. 1996; Zhang, Mashima et al. 2001; Kanno, Ogihara et al. 2006) and the transdifferentiation into hepatocytes (Shen, Slack et al. 2000; Shen, Horb et al. 2003; Burke, Shen et al. 2006).

From the 2D analysis were obtained some, but at the moment, not conclusive informations. Among the few proteins identified to be differentially expressed in the Act-A/HGF-treated samples most of them were related to the response to oxidative stress such as PARK7 (Sekito, Koide-Yoshida et al. 2006) and CH60 (Bonior, Jaworek et al. 2005)) as well as related to the vesicular trafficking, (e.g.VGF8 (Snyder, Peng et al. 2003)) and to the protein folding (e.g. PDIA3 (Elliott, Oliver et al. 1997), GRP75 (Burkart, Liu et al. 2000)).

With the mass spectrometry analysis we had to face many challenges such as the low amount of proteins available. In fact the working hypothesis was to analyze the changes in the protein pattern, essentially in term of transcription factors, which occur during the differentiation. From a technical point of view this means to proceed with an important and necessary step of selective proteins extraction and enrichment. The protocols (or the commercial kits) available show a high efficiency of purification, when used as analytical tools; they loose their efficiency when are scaled-up as requested by a proteomics analysis. The time consuming (and high cost) process of extraction and purification allowed for the preparation of samples for few 2D replicates that, even statistically significant, were too preliminary to give conclusive answers.

Anyway, the identification among the genes and the proteins over-expressed of members involved in the response to “stress” events could be attributed to the treatment with growth factors and in particular to Act-A that is well know to trigger apoptosis

(Chen, Wang et al. 2006). To overcome these problems can be envisaged others promising approaches such as the possibility to use less “aggressive” differentiation agents. Recently a vinca alkaloid, namely conophylline, has been identified as a potent differentiation agent for the AR42J (Kojima and Umezawa 2006). Conophylline-treated cells were found to express insulin as measured at both mRNA and protein levels. By RT-PCR analysis, conophylline-treated cells were shown to express Ngn3 strongly as well as NeuroD/BETA2, Nkx2.2 and insulin (Umezawa, Hiroki et al. 2003).

The difficulty to induce the differentiation process, even in a simple cell model, through the treatment with growth factors that could trigger unexpected epigenetic events prompted us to pursue a more reliable and, for some aspects easy, “molecular” strategy. As already stated in many part of this Thesis during the pancreas differentiation, the Ngn3 plays a central role placing itself at the intersection between an undifferentiated state (upstream the Ngn3) and the beginning of the endocrine pancreas differentiation (downstream the Ngn3) (Figure 3 Introduction)Therefore, all the molecular events that occur immediately downstream to the Ngn3 are truly pancreatic differentiation events. The availability of a simple experimental system in which the Ngn3 is constantly activated or, even better, which expression can be customized could open new interesting perspectives where technical approaches such as microarrays analysis and 2D may play import roles.

From a physiological point of view, it is known that cells that transiently express the basic helix-loop-helix transcription factor Ngn3 are precursors that can become the endocrine pancreas. Ngn3 expression starts at E13.5 and peaks at E15.5 before dissipating, corresponding to the developmental stage of formation of the definitive endoderm. It turn out then that Ngn3 is one of the essential transcription factors expressed early in the differentiation of the islet cells and it is considered to be a marker of endocrine progenitor cell during development of the pancreas.

Despite the important role played by Ngn3 information regarding some aspect of Ngn3 activity is still missing. In particular are still open the following questions;

- 1) *Is the transient expression of Ngn3 a prerequisite to have the proper molecular context to achieve β -cells differentiation in vitro?* In fact, most of the experimental approaches used have had ectopic and constitutive expression of Ngn3 driven by

strong promoter (Heremans, Van De Casteele et al. 2002; Gasa, Mrejen et al. 2004) without any possibility of a feedback regulation of its own expression as happens normally during pancreas differentiation.

- 2) *Is the endoderm patterning a prerequisite to have the proper cellular context to achieve β -cells differentiation?* Development of the pancreatic endocrine cells located in the dispersed islets of Langerhans is only partially characterized, and although some information are available a complete vision of cellular and molecular mechanisms involved are still missing.
- 3) *Which are the direct downstream targets of *Ngn3*?* These information would provide valuable data to set the development of *in vitro* methods for the differentiation either of ES cells (pluripotent), of adult progenitor (multipotent) and, potentially, to drive transdifferentiation of non- β -cells towards insulin-secreting cells.

Therefore the aims of this part of the PhD Thesis have been:

- 1) *To test if transient expression of *Ngn3* is sufficient to act as a genetic switch that specifies the β -cell fate in ES cells in vitro;*
- 2) *To test if the context of endoderm patterning is a prerequisite for β -cells differentiation.*
- 3) *To identify those genes, mainly transcription factor(s), that lie downstream *Ngn3*.*

To attain these goals the following approaches have been used:

1. Simple cell models;
2. A "customized" gene expression for activation of expression of the target gene for specific times using inducible expression vectors
3. A high-throughput experimental approach.

The *cell model* is necessary to mimic as much as possible the physiological behavior of β -cell differentiation. As cells models the F9, AR42J and INS-1 cell lines have been chosen.

The mouse F9 teratocarcinoma cell line is a model that can be manipulated to imitate one of the earliest transitions in mouse development, which is the formation of endoderm. The cells undergo very limited differentiation under normal culture conditions, but can be induced to differentiate into a) parietal endoderm in the presence of retinoic acid (RA) and dibutyryl cyclic AMP; b) visceral endoderm when cultured in suspension as aggregates and treated with RA (Strickland and Mahdavi 1978).

The AR42J cells, as stated before, were originally derived from a chemically induced rat pancreatic tumor and possess both exocrine and neuroendocrine properties. In the presence of Activin-A (Act-A) plus Hepatocyte Growth Factor (HGF), these cells differentiate into insulin-expressing cells.

INS-1 is an insulin-secreting cell line established from an x-ray-induced rat transplantable insulinoma (Asfari, Janjic et al. 1992). This cell line express both proinsulin I and II and display conversion rate of the two precursors hormones similar to those observed in rat islets.

The *customized gene expression* has been obtained by the use of *inducible expression system*.

An *High throughput experimental approach* is necessary to obtain the higher grade of information in term of genes expressed per experiment.

5.1 Are the transient expression of Ngn3 and the endoderm patterning a prerequisite to have the proper molecular context to achieve β -cells differentiation in vitro?

During embryonic development Ngn3 appears *transiently* in scattered undifferentiated cells of *endoderm* origin. Several lines of evidence suggest that the expression of Ngn3 in endoderm progenitor cells directs them to an endocrine cell fate and initiates the program of islet differentiation (Heremans, Van De Casteele et al. 2002; Gasa, Mrejen et al. 2004). Despite this important insight, the approach used was based on a constant expression of Ngn3 even though it is know that Ngn3 is expressed as a pulse.

To ensure both an efficient and regulatable expression of the target gene we used tetracycline-induced transduction system both retroviral and not.

Retroviral gene delivery will offer an efficient *transduction system*; it is generally faster, more reliable, and has broader utility than alternative gene transfer methods.

The RevTet system used is based on the use of 1) a retroviral regulator vector and 2) a retroviral response vector carrying the gene of interest to be regulated. The retroviral regulator vector express the tTA regulatory element, a fusion of the Tet repressor protein (TetR) and the VP16 activation domain of the *Herpes simplex* virus VP16 protein. The tTA expressed by the retroviral regulator vector binds the TRE of the response vector activating transcription in presence or in absence of doxycycline. With the Tet-Off system the gene expression is turned on when doxycycline is removed from the culture medium whereas in the Tet-On system the gene expression is turned on when doxycycline is added to the culture medium since the TetR was modified (obtaining the rtTA) by four amino acid changes that reverse the protein's response to doxycycline.

The use of viral vectors as transduction system to drive the expression of transcription factors in the pancreas is not novel. For example Song et al (Song, Lee et al. 2007) developed a bi- or tri-cistronic constructs in hybrid adenovirus/adeno-associated virus (Ad/AAV) vectors containing Ngn3, NeuroD/BETA2, and RIPE3b1 (MafA), each of which plays a role in islet cell differentiation. These vectors efficiently express multiple transcription factors and stimulate insulin promoter activity in a combinatorial manner. When these multi-cistronic constructs were administered *in vivo*, they induced hepatic expression of islet-specific markers, including Pdx-1, insulin, glucagon, somatostatin, and islet-amyloid peptide. Administration of the Ad/AAV hybrid vectors to streptozotocin-induced diabetic mice reversed hyperglycemia, consistent with the differentiation of functional hepatic insulin-secreting cells (Song, Lee et al. 2007).

With a similar approach Wang et al (Wang, Ehrhardt et al. 2007) delivered Pdx-1 and Ngn3 to the livers of diabetic mice using adeno-associated virus (AAV) serotype 8, a vector that has been shown to result in non-toxic, persistent, high level expression of the transgene. However, they were unable to correct hyperglycemia in mice with streptozotocin-induced diabetes using AAV vectors expressing Pdx-1 and Ngn3 but, when they co-delivered these transcription factor expression cassettes in non-viral vectors

with an irrelevant adenoviral vector, they were able to correct hyperglycemia in diabetic animals. Further studies demonstrated that an antigen-dependent immune response elicited by the adenoviral capsid together with the expression of a pancreatic transcription factor was required for restoration of serum insulin levels by the liver.

In both the cases described the viral transduction driven a constant gene expression. In any case the need of a pulsed expression has been never addressed.

More interestingly Treff et al. (Treff, Vincent et al. 2006) have determined the expression profiles of murine embryonic stem cells (mESCs) uniformly induced to over-express Ngn3. The mESC line created was induced with doxycycline for 72 h and then a genome-wide microarray analysis was performed to identify genes regulated by Ngn3 in a variety of contexts, including undifferentiated ESCs and differentiating embryoid bodies (EBs). Genes regulated by Ngn3 in a context-independent manner were identified and analyzed using systematic gene ontology tools. This analysis revealed Notch signaling as the most significantly regulated signaling pathway. This result is consistent with the hypothesis that Ngn3 expression makes the cell competent for Notch signaling to be activated and, conversely, more sensitive to Notch signaling inhibition. Indeed, EBs induced to express Ngn3 were significantly more sensitive to γ -secretase inhibitor-mediated Notch signaling inhibition when compared with uninduced EBs. Moreover, they found that Ngn3 induction in differentiating ESCs results in significant increases in insulin, glucagon, and somatostatin expression.

These results are very interesting although they addressed only the problem of the Ngn3-gene pattern in term of positive regulation. Furthermore, the data need to be interpreted with caution because the genes identified may not be direct target genes of the overexpressed Ngn3, but might be the result of indirect regulation of gene expression patterns.

As for the embryonic context the importance of the endoderm patterning has been never considered directly in any of the report available although the experiments were always carried out with cell line and/or organ system already committed toward the appropriate phenotype. In an *in vitro* context O'Driscoll et al. (O'Driscoll, Gammell et al. 2004) showed that, by using retinoic acid-derived-endoderm F9 cells as a model, ectopic expression of Ngn3 is able to start the differentiation pathway of endocrine pancreas.

Ngn3 triggers the expression of several pancreatic transcription factors following a well defined temporal activation sequence. The data reported demonstrated that Ngn3 is the first event in the temporal sequence of activation of some transcription factors specifically involved in the differentiation of β -cells. Interestingly, gene activation triggered by Ngn3 was restricted to the cells that underwent endoderm differentiation after RA treatment. In contrast, ectopic Ngn3 expression in cells not treated with RA was not sufficient for inducing expression of genes involved in the signaling pathways that promote endocrine pancreas differentiation.

More recently a very elegant approach demonstrating the importance of the embryonic context for the appropriate competence of Ngn3 in the pancreas has been reported by Johansson et al. (Johansson, Dursun et al. 2007). They generated transgenic mice expressing a tamoxifen-inducible Ngn3. In order to specifically express this inducible protein in the pancreas, they put it under the transcriptional control of the mouse *pdx1* promoter. The transgene was then backcrossed into the *ngn3*^{-/-} background, devoid of endogenous endocrine cells. The expression of Ngn3 was induced by injecting the pregnant females with a single dose of tamoxifen at e8.7, e10.5, e12.5, or e14.5.

Early activation of Ngn3-ERTM (e8.7) almost exclusively induced glucagon⁺ cells, while depleting the pool of pancreas progenitors. As from e11.5, Pdx1⁺ progenitors became competent to differentiate into insulin⁺ and PP⁺ cells.

Somatostatin⁺ cells were generated from e14.5, while the competence to make glucagon⁺ cells was dramatically decreased. Hence, pancreas progenitors, go through competence states that each allow the generation of a subset of cell types (Johansson, Dursun et al. 2007).

It is important to note that most of classical inducible transduction systems, both viral and not, rely on the use of the Tetracycline (Tet) system.

An alternative and most promising approach is that developed by NEB. Analogous to the operation of a rheostat, the RheoSwitch technology from NEB allows induction and adjustable control of gene expression. The precise regulation of gene expression is achieved through the highly specific interaction of a synthetic inducer, the Ligand RSL1, and a chimeric bipartite nuclear receptor. Unlike other systems, which rely on steroids or other drugs, the synthetic ligand RSL1 (synthetic diacylhydrazine) shows

no pleiotropic effects in mammalian cells and exhibits no cross talk with endogenous transcription factors. Moreover, RSL1 is not a steroid and it is invisible to mammalian nuclear receptors and has been shown to be inert within all cell lines tested to date.

5.2 Which are the direct downstream targets of Ngn3?

As already mentioned before the proendocrine basic helix-loop-helix protein Ngn3 plays a critical role in the development of the hormone-producing islets within the pancreas, including the insulin-producing β -cell. Deletion of *ngn3* gene in mice results in complete loss of pancreatic endocrine cells as well as development of diabetes. Despite numerous reports describing its role, our understanding of the regulatory networks surrounding Ngn3 remains limited. It has been reported that the expression of Ngn3 is directly regulated by HNF6 (Jacquemin, Durviaux et al. 2000), Hes (Lee, Smith et al. 2001), and Ngn3 itself (Smith, Gasa et al. 2003). In addition, Ngn3 has been shown to directly regulate the expression of NeuroD/BETA2 (Huang, Liu et al. 2000), Pax4 (Smith, Gasa et al. 2003), Nkx2.2 (Watada, Scheel et al. 2003) and IA1 (Mellitzer, Bonne et al. 2006). Specification and further differentiation of the islet cell types is controlled by target genes of Ngn3, many of which are still unknown.

In order to identify novel target genes of Ngn3, the transcriptome of pancreas-derived cells ectopically expressing Ngn3 was analyzed using a ChiP-cloning approach. To avoid the complexity of the cell types involved present in embryonic pancreas cell models as simple as possible. As cell models the AR42J and INS-1 have been used.

The AR42J cell line has been previously used as a model to study, in general, the changes in the expression of genes (Palgi, Stumpf et al. 2000) and in particular the changes in the expression of transcription factors (Zhang, Mashima et al. 2001) during differentiation into insulin-producing cells.

INS-1 is an insulin-secreting cell line established from an x-ray-induced rat transplantable insulinoma (Asfari, Janjic et al. 1992). This cell line express both

proinsulin I and II and display conversion rate of the two precursors hormones similar to those observed in rat islets.

AR42J and INS-1 cells are the most valuable model available with respect embryonic pancreata that need sophisticated technical skills to be managed.

As technical approach to get a transcriptome analysis of the Ngn3-induced gene(s) the ChiP-cloning technique has been used. This is a technique for isolation and identification of the DNA sequences occupied by specific DNA binding proteins in cells. These binding sites may indicate functions of various transcriptional regulators and help identify their target genes during animal development and disease progression. The identified binding sites may also be used as a basis for annotating functional elements in genomes. The types of functional elements that one can identify using ChiP-cloning include promoters, enhancers, repressor and silencing elements, insulators, boundary elements, and sequences that control DNA replication.

It has been predicted that at least 2000 transcriptional activators are encoded by the human genome (Tupler, Perini et al. 2001). To make use of this information, it is now important to determine the sets of genes regulated by each of these factors. A common approach used to identify the target genes that are regulated by an individual factor is to couple the overexpression or underexpression of that factor to cDNA microarray analysis (Muller, Bracken et al. 2001). Although this approach allows for the isolation of a large set of potential target genes, the data need to be interpreted with caution for several reasons. First, the genes identified may not be direct target genes of the overexpressed factor, but instead may be isolated as the result of indirect regulation due to overall alterations of gene expression patterns. In addition, it is unclear that the genes regulated by a factor at levels vastly greater than normal, biologically relevant concentrations are in fact true target genes. Therefore, many studies have focused on the development of new approaches designed to examine the direct targets of a site-specific transcription factor in the context of physiologically relevant conditions. One of the most promising methods is a modification of the chromatin immunoprecipitation (ChiP) procedure. ChiP has become a popular method in regulatory biology for studying the binding of transcription factors to genomic loci. The ChiP technique has been traditionally used for demonstration of the *in vivo* binding of a given transcription factor to genomic region of interest (Kuo and Allis

1999; Orlando 2000). In recent years, this technique has also become a useful genomic tool for identifying unknown target loci of transcription factors by cloning and analyzing the retrieved DNA fragments that have been bound by the transcription factors (Kirmizis and Farnham 2004). Most of the available ChIP protocols employ a similar overall strategy to retrieve a pool of protein-bound DNAs (Fig X). In brief, chromatin is cross-linked by formaldehyde, fractionated by sonication and purified by the antibody raised against a given protein and later the bound DNAs are reverse cross-linked by heat treatment and phenol/chloroform extraction. Since the initial chromatin for ChIP is sheared with sonication, both ends of the purified DNAs are usually modified further by fill-in reactions to generate blunt-end DNA fragment. This modified pool of genomic DNAs usually becomes the starting DNA material for constructing clone libraries for the identification of unknown target regions for a given DNA-binding protein.

Recent microarray studies have implicated a number of additional candidate targets of Ngn3 in the developing pancreas (Gu, Dubauskaite et al. 2002) and in pancreatic ductal cells (Gasa, Mrejen et al. 2004). These studies were performed on different microarray platforms and were limited to approximately 12,000 or 18,000 genes, respectively. Unfortunately again it is not possible to discriminate specifically those genes activated directly by Ngn3.

Following this modified protocol of ChIP it has been possible to identify at least three genes encoding for proteins belonging to the class of transcription factors that are specifically regulated by Ngn3. They are:

- 1) The basic Helix-Loop-Helix neurogenic factor 4 (NeuroD4).
- 2) The SRY (sex determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal) (Sox9).
- 3) The zinc-finger type transcription factor OVO homolog-like 1 (OVOL1).

The bHLH transcription factor neurogenic differentiation 1 (NeuroD/BETA2) or β -cell E-box transactivator 2 (BETA2) plays a particularly prominent role in pancreatic β -cell function. It is required for both normal islet development and regulation of insulin gene transcription (Naya, Huang et al. 1997). Homozygous NeuroD/BETA2/BETA2-null

mice show a reduction in the number of insulin-producing β -cells and fail to develop mature islets. In addition, heterozygous mutations in the human NeuroD/BETA2/BETA2 gene (NeuroD/BETA2) are associated with a monogenic form of diabetes that resembles maturity-onset diabetes of the young (MODY) in many respects except perhaps for a slightly later age at onset (Malecki, Jhala et al. 1999). A search for other bHLH proteins expressed in the developing pancreas and mature islet cells revealed the presence of several other members of this family, including the NeuroD/BETA2-related protein NeuroD4/Math-3/Ath-3. NeuroD4 is widely expressed in the developing nervous system and, after birth, it is expressed at its highest levels in the retina (Lee 1997; Tsuda, Takebayashi et al. 1998). It is also found in developing pancreas and mature rodent islet cells, and its expression has been confirmed in adult human islets by RT-PCR-based assay (Lee 1997). Interestingly, recently it has been reported that in addition to NeuroD/BETA2, Ngn3 also induced both NeuroD2 and NeuroD4 gene expression (Gasa, Mrejen et al. 2007).

As for Sox9 much more data are available regarding its role in pancreas development. The *sox* gene family represents a group of developmentally regulated transcription factors, which are characterized by a highly conserved DNA-binding domain, the HMG-domain (Wegner 1999; Schepers, Teasdale et al. 2002). Based upon structural similarities, the more than 20 different vertebrate *sox* genes can be assigned to one of the seven subgroups, A-G (Schepers, Teasdale et al. 2002). Recently, several *sox* genes have been recognized as key players in the regulation of embryonic development and cell fate determination. The analysis of *sox* gene mutations in humans, mice, and zebrafish have demonstrated a role for *sox* genes in endoderm specification, as well as the development of gonads, lens, heart, lymphocytes, bone, and glial cells (Schilham, Oosterwegel et al. 1996; Britsch, Goerich et al. 2001; Dickmeis, Mourrain et al. 2001; Kamachi, Uchikawa et al. 2001). Specifically related to the pancreas it has been shown that at least eight different members of the *sox* gene family in the embryonic pancreas, and three members in adult islet cells (Lioubinski, Muller et al. 2003) were expressed. Among those, Sox9 was confined to the pancreatic epithelium and later to islets. The expression of Sox9 in the embryonic pancreatic ducts is of particular interest, because ducts have been suggested to contain the endocrine progenitor cells (Schwitzgebel,

Scheel et al. 2000; Lee and Saint-Jeannet 2003). In fact recently Sox9 has been identified as the first specific marker of multipotential progenitors during pancreas organogenesis. It acts stimulating the proliferation, survival, and persistence in an undifferentiated state of the multipotential progenitor cells (Lynn, Smith et al. 2007; Seymour, Freude et al. 2007).

Expression of Sox9 has also been reported in the human pancreas (Piper, Ball et al. 2002).

Interestingly the *sox9* gene shows extended regulatory cis-regulatory regions. Most of these elements reside within a 600-kb interval extending 350 kb 5' and 250 kb 3' to *sox9* (Wunderle, Critcher et al. 1998; Bagheri-Fam, Barrionuevo et al. 2006; Bien-Willner, Stankiewicz et al. 2007). According to these observations we included *sox9* as a potential gene regulated by Ngn3 even though the long range localization from the 5' side of the *sox9* gene (89200 bp).

Among the genes prominently affected by Ngn3 we focused the attention on that encoding the protein OVO homolog-like 1 (OVO11). So far neither the expression pattern of OVO11 nor its function in the pancreas has been addressed directly.

The molecular structure of the *ovo* gene product contains four C₂H₂ zinc finger motifs, characteristic of zinc-finger type transcription factor (Mével-Ninio, Terracol et al. 1991; Garfinkel, Wang et al. 1994; Mével-Ninio, Terracol et al. 1995; Mével-Ninio, Fouilloux et al. 1996). *ovo* gene has been described in *Drosophila* to be required for survival and differentiation of female germ line cells where it plays an important role in germ line sex determination (Oliver 1987; Oliver, Pauli et al. 1990; Oliver, Singer et al. 1994; Horabin, Bopp et al. 1995). Homologous null mutants produce rudimentary ovaries in which germ cells have completely degenerated, leaving only somatic tissues (Oliver 1987). Less severe mutations produce viable germ cells that exhibit defective oogenesis (Oliver 1987) or ovarian tumors with male germ line features (Oliver 1987; Oliver, Pauli et al. 1990). Later on (Masu, Ikeda et al. 1998) a cDNA encoding for a 274-amino acid protein containing an acidic amino acid and serine-rich domain and a zinc finger domain which shows high sequence homology to that of *Drosophila* Ovo protein has been cloned. The mRNA is predominantly expressed in testis, and the expression increases from 3 weeks postnatal, implying that these proteins may play important roles in the

development of the testes where it is present in spermatocytes and spermatids (Masu, Ikeda et al. 1998). To assign definitively OVO to a transcriptional factor, the DNA-binding and transactivation properties have characterized. Random oligonucleotide selection, electrophoretic mobility shift assay and footprinting indicated that OVO bound to the sequence, 5'-G(G/C/T)GGGGG-3'. These motifs were found in the 5'-flanking regions of *ovo* and other testis-specific genes (Unezaki, Nishizawa et al. 2004).

Using gene targeting, it has been show that *ovo* is required for proper development of both hair and sperm. *Ovo*^{-/-} mice are small, produce aberrant hairs, and display hypogenitalism, with a reduced ability to reproduce. These mice also develop abnormalities in kidney, where *ovo1* is also expressed.

Ovoll seems to regulate meiotic pachytene progression during spermatogenesis. In fact *ovoll*-deficient germ cells are defective in progressing through the pachytene stage. The pachytene arrest was accompanied by an inefficient exit from proliferation, increased apoptosis and an abnormal nuclear localization of the G2-M cell cycle regulator cyclin B1, but was not associated with apparent chromosomal or recombination defects. Transcriptional profiling revealed reduced expression of pachytene markers in the mutant, providing molecular evidence that pachytene differentiation was defective. In addition, the expression of Id2 (inhibitor of differentiation 2), a known regulator of spermatogenesis, was upregulated in *ovoll*-deficient pachytene spermatocytes and repressed by *Ovoll* in reporter assays (Li, Nair et al. 2005).

In mice, *ovoll* gene is also expressed in skin, where it localizes to the differentiating cells of epidermis and hair follicles. In its absence, developing epidermis fails to properly restrict the proliferative potential of progenitor cells, and cultured keratinocytes fail to efficiently undergo growth arrest in response to extrinsic growth-inhibitory signals. There are many evidences showing that c-myc expression is up-regulated in *Ovoll*-deficient suprabasal cells and that *Ovoll* represses c-myc transcription by directly binding to its promoter. c-Myc proto-oncoprotein is one the most important transcription factors that control the balance between proliferation and differentiation of keratinocytes (Nair, Teng et al. 2006).

As for the positive regulation of the *ovoll* gene, recently the *wnt* signal transduction pathway has been involved (Li, Mackay et al. 2002). Wnt proteins are a

family of highly conserved secreted proteins that regulate multiple developmental processes, including proliferation of organ-specific stem/progenitor cell populations, growth control and cell fate determination in diverse organs, and tissue patterning (Logan and Nusse 2004). In the canonical Wnt signaling pathway, binding of Wnt ligand to cognate membrane-spanning receptor proteins called Frizzleds (Fz), triggers a signaling cascade that results in the stabilization and nuclear localization of β -catenin, whose interactions with T cell-specific factor/lymphoid enhancer-binding factor (TCF/LEF) transcription factors control transcription of target genes like Pitx2 (Kioussi, Briata et al. 2002; Baek, Kioussi et al. 2003). In the absence of Wnt signaling, cytoplasmic β -catenin in Wnt-responsive cells is targeted for proteasome-mediated degradation by a heteromeric protein complex that includes Axin, glycogen synthase kinase 3 β (GSK-3 β), and other proteins (Rulifson, Karnik et al. 2007).

Components of the *wnt* signal transduction pathway have been functionally implicated in multiple events during hair follicle morphogenesis and differentiation. For example, ectopic expression of *wnt3a* in mice leads to hair defects (Li, Mackay et al. 2002). LEF1 knockout mice display a reduced number of hair follicles, and residual follicles fail to produce normal hair shafts (van Genderen, Okamura et al. 1994). Conversely, ectopic expression of LEF1 leads to hair formation in ectopic locations (Zhou, Byrne et al. 1995). Expression of a stable form of β -catenin in skin leads to overt phenotypes such as *de novo* hair follicle morphogenesis and abnormal angling of protruding hairs (Gat, DasGupta et al. 1998), whereas conditional ablation of β -catenin in the epidermis or expression of an N-terminally truncated LEF1 that cannot associate with β -catenin results in defective hair morphogenesis (Huelsenken, Vogel et al. 2001; Merrill, Gat et al. 2001). These and other studies suggest that *wnt* signaling, LEF1, and β -catenin are required for formation of hair follicles during embryogenesis as well as for postnatal hair production. An analysis of the gene regulatory region of *ovoll* promoter showed that it contains a LEF1 binding site that is activated by the ectopic expression of the LEF1/ β -catenin complex (Li, Mackay et al. 2002)

As for the negative regulation some genetic evidence suggests that *Drosophila ovo* auto-regulates (Oliver, Singer et al. 1994; Lu, Andrews et al. 1998) its own expression by binding to and repressing the *ovoll* promoter. *Ovoll* represses transcription

using both passive (competing with the c-Myb transcriptional activator, a known positive regulator of proliferation) and active (recruiting histone deacetylases 1) repression mechanisms (Nair, Bilanchone et al. 2007). Furthermore, as showed above, *Ovo11* represses c-Myb and *Id2* transcription by binding to their promoters (Li, Nair et al. 2005; Nair, Teng et al. 2006). Interestingly c-Myb is a known target of c-Myb (Evans, Moore et al. 1990; Nakagoshi, Kanei-Ishii et al. 1992; Cogswell, Cogswell et al. 1993), and c-Myb indeed binds to an *Ovo11* site in the *Id2* promoter. Therefore, the passive and active repression mechanisms of *Ovo11* may apply to *Ovo11*-mediated regulation of its other target genes.

A high number of evidences and data seem to confine the role of *ovoll* gene in controlling the development and differentiation of hair formation and spermatogenesis, therefore a possible role of *ovoll* gene in controlling pancreas development and differentiation that has even a different embryonic origin (endoderm vs. ectoderm) could seem improbable. However, interestingly, some of the target genes that act downstream *Ovo11* are a common trait of the pancreas development. Among those an interesting role is played by *Id2* and c-Myb. *Id2* belongs to the *Id* family of helix-loop-helix (HLH) proteins, which upon heterodimerization with basic HLH proteins prevent basic HLH proteins from DNA binding (Norton 2000; Lasorella, Uo et al. 2001). Proteins of the *Id* family act as negative regulatory transcriptional factors, and their expression correlates with cell proliferation and arrested differentiation in many cell lineages.

As for pancreas-related function it has been showed that *Id2* protein inhibited NeuroD/BETA2 binding to its target sequence E-box inhibiting its stimulatory activity. As mentioned earlier (see “Introduction”) NeuroD/BETA2 regulates the survival and terminal differentiation of pancreatic beta cells. Gene-targeting experiments revealed that deletion of the NeuroD/BETA2 gene resulted in defective pancreatic morphogenesis and abnormal enteroendocrine differentiation, which leads to the development of early diabetes (Naya, Huang et al. 1997). Further upstream it has been showed that *Id2* in the pancreas could be activated by Bone Morphogenic Protein 4 (BMP-4). BMP signaling molecules, including SMADs 1, 2, and 4, are expressed in neonatal and adult pancreatic islets (Brorson, Hougaard et al. 2001), and BMP4 promote the proliferation and development of epithelial islet-like structures *in vitro* (Jiang, Stanley et al. 2002). A

SMAD-binding element has been found in the promoter of Id2 protein (Hollnagel, Oehlmann et al. 1999). In the presence of BMP4 ligand, BMP receptors transmit the signal that leads to phosphorylation (*P*) of SMAD-1 and -5 as well as co-SMAD-4. Activated SMADs bind to SMAD-binding elements on the Id2 promoter resulting in transcription of Id2 protein. This binds to NeuroD/BETA2, preventing it from binding to the E-box elements of endocrine-specific genes (Hua, Zhang et al. 2006).

As for c-Myc it has been showed that it controls the inverse relationship between proliferation and differentiation in human pancreatic endocrine cells. c-Myc is expressed in human fetal and adult pancreatic tissue (Schmid, Schulz et al. 1989), but not in differentiated endocrine cells. c-Myc is a target gene of Wnt/ β -catenin pathway (Evan and Littlewood 1993), which is activated during pancreas development (Amati and Land 1994). β -Catenin is identified as the component of the E-cadherin-mediated cell-cell adhesion system (Blackwell, Kretzner et al. 1990) and a key effector of the Wnt signaling pathway, which plays a critical role in growth, division, and cell fate at early and late developmental stage (Prendergast, Lawe et al. 1991; Blackwood, Luscher et al. 1992; Blackwell, Huang et al. 1993). In respond to Wnt signals, β -catenin complexes with T cell factor (Tcf)/lymphoid-enhancing factors and p300 to induce transcription of target genes known to be important for normal cell proliferation, such as c-Myc (He, Sparks et al. 1998) and cyclin D1 (Shtutman, Zhurinsky et al. 1999).

It is well known that during development the differentiation into hormone-expressing cells requires cell-cell contact acting in synergy with the homeodomain transcription factor pancreatic duodenal homeobox-1 (Pdx-1). This event is associated with a decrease in cell proliferation that has been showed to be mediated a down-regulation of the *c-myc* protooncogene (Demeterco, Itkin-Ansari et al. 2002). Furthermore, conditional activation of c-Myc using a fusion protein between c-Myc and a hormone-binding domain derived from the estrogen receptor leads to the ablation of hormone expression and an increase in cellular proliferation.

From a molecular point of view recently it has been showed that this downregulation of c-Myc expression is mediated by Pdx-1. Pdx-1 can suppress β -catenin/Tcf-induced c-Myc transcription by inhibiting the level of β -catenin/Tcf and p300

complex formation and reducing their binding capacities to endogenous c-Myc promoter region (Chen, Yan et al. 2007).

Therefore, it is believed that c-Myc, through Pdx-1, could play a role in the switch mechanism that controls the inverse relationship between proliferation and differentiation in human pancreatic endocrine cells. Proliferation of postnatal islet cells is relatively low compared with tissues such as intestines or bone marrow, except in states favoring robust β cell growth, such as pregnancy, insulin resistance, and obesity (Heit, Karnik et al. 2006).

The mechanism by which this equilibrium is finely tuned shows common trait with that one encountered in the differentiation/growth of epidermal progenitor cells.

In both epidermal and pancreatic progenitor cells the delicate balance between proliferation and differentiation are mediated by the *wnt* signal transduction pathway. Multiple Wnt ligands, receptors, and signal transduction factors are expressed in the embryonic and adult pancreas (Murtaugh, Law et al. 2005; Papadopoulou and Edlund 2005). Although prior studies provided evidence that pancreatic growth and differentiation are regulated by Wnt signaling (Heller, Dichmann et al. 2002; Heller, Klein et al. 2003; Pedersen and Heller 2005; Rulifson, Karnik et al. 2007) these reports did not present a mechanism for the Wnt-mediated action nor did they test whether Wnt signaling was sufficient to stimulate β -cell proliferation in pancreatic islets (Rulifson, Karnik et al. 2007). In the seminal work of Rulifson et al. (Rulifson, Karnik et al. 2007) it has been shown that Wnt signaling stimulates islet β - cell proliferation. In fact, the addition of purified Wnt3a protein to cultured β -cells or islets promoted expression of Pitx2, a direct target of Wnt signaling, and Cyclin D2, an essential regulator of β -cell cycle progression, and led to increased β -cell proliferation *in vitro*. Conditional pancreatic β -cell expression of activated β -catenin, a crucial Wnt signal transduction protein, produced similar phenotypes *in vivo*, leading to β -cell expansion, increased insulin production and serum levels, and enhanced glucose handling. Conditional β -cell expression of Axin, a potent negative regulator of Wnt signaling, led to reduced Pitx2 and Cyclin D2 expression by β -cells, resulting in reduced neonatal β -cell expansion and mass and impaired glucose tolerance. Thus it turned out that Wnt signaling is both necessary and sufficient for islet β -cell proliferation providing evidence of a mechanism governing

endocrine pancreas growth and function. The Wnt signaling acts, in *Xenopus* embryos, very early during the differentiation process, at level of organ patterning. It has been showed that Wnt/ β -catenin activity must be repressed in the anterior endoderm to maintain foregut identity and to allow pancreas development. By contrast, high β -catenin activity in the posterior endoderm inhibits foregut fate while promoting intestinal development. Experimentally repressing β -catenin activity in the posterior endoderm was sufficient to induce ectopic organ buds that express early liver and pancreas markers. β -Catenin acts in part by inhibiting expression of the homeobox gene *hhex*, which is one of the earliest foregut markers and is essential for liver and pancreas development. Promoter analysis indicates that β -catenin represses *hhex* transcription indirectly via the homeodomain repressor Vent2. Later in development, β -catenin activity has the opposite effect and enhances liver development. These results illustrate that turning Wnt signaling off and on in the correct temporal sequence is essential for organ formation, a finding that might directly impact efforts to differentiate liver and pancreas tissue from stem cells (Rulifson, Karnik et al. 2007). In the attempt to define precisely the topography of involvement of the wnt signaling in pancreas Murtaugh et al. (Murtaugh, Law et al. 2005) observed that that the early pancreatic epithelium exhibits a specific enrichment in unphosphorylated β -catenin protein. To clarify the role of wnt signaling in pancreas development they specifically deleted the *β -catenin* gene in these cells. Pancreata developing without β -catenin are hypoplastic, although their early progenitors appear normal and exhibit no premature differentiation or death. Surprisingly, and in marked contrast to its role in the intestine, loss of β -catenin does not significantly perturb islet endocrine cell mass or function. The major defect of the β -catenin-deficient pancreas is an almost complete lack of acinar cells, which normally comprise the majority of the organ (Murtaugh, Law et al. 2005).

The discovery of a hypothetical role of OVO11 transcription factor in the pancreas could open new hypothesis of the molecular mechanism involved in the differentiation of pancreas. Previously no information of the expression pattern of OVO11 during pancreas development had been known. Our data show that OVO11 is expressed in islets and in pancreas-related cells (INS-1 and AR42J) and that Ngn3 seems to act as a negative regulator of its expression. Interestingly Ngn3 is active during a proliferative stage that

ends with the appearance of almost definitive islets clusters. The proliferation burst induced by Ngn3 seems incompatible with the differentiation one and thus should require a “silencing” of the factors inducing it.

Our hypothesis is that OVO11 is active in this time of transition to differentiation in the pancreas. These discoveries, albeit preliminary, suggest a new interesting scenario in pancreas development involving another layer of regulation by new unsuspected “players”. OVO11 and Ngn3 in DNA binding and transcriptional regulation are particularly exciting in light of their apparently opposing biological functions. Ngn3 positively regulates the proliferation of immature cells that are committed to differentiation, while OVO11 is expressed later in a differentiation pathway and is necessary for efficient proliferation arrest (Dai, Schonbaum et al. 1998; Li, Dai et al. 2002; Li, Nair et al. 2005; Nair, Teng et al. 2006). Based on the data presented here, it is tempting to speculate that Ngn3 and OVO11 transcriptional regulation represent two consecutive steps in a relevant differentiation pathway, where the balance of the opposing effects of these two proteins coordinates proliferation with differentiation. Since Ngn3 is expressed in proliferating pancreatic progenitor cells, a possible scenario in pancreas development is that OVO11 activates genes such as c-Myc and Id2 to up-regulate the transient proliferation of progenitor cells that have committed to differentiate, but also turns on the expression of its own antagonist, OVO11, to restrict proliferation to allow terminal differentiation to actually occur.

The activity of Ngn3 emerging from our experimental data appear quite unusual. Normally Ngn3 has been considered an activator of gene transcription. The present study reveals, instead, that OVO11 is expressed in similar cells as Ngn3 during pancreatic development and shows that the *ovoll* gene is directly negatively regulated by Ngn3.

As mentioned earlier (see “Introduction”) *ngn3* encodes a class B basic helix-loop-helix factor, which has been shown by loss-of-function studies to be required for the development of the four endocrine cell lineages of the pancreas (Gradwohl, Dierich et al. 2000). The pro-endocrine role of *ngn3* has also been demonstrated in gain-of-function studies. Ectopic *ngn3* expression (Apelqvist, Li et al. 1999; Jensen, Pedersen et al. 2000; Schwitzgebel, Scheel et al. 2000), as well as lineage tracing experiments (Gu, Dubauskaite et al. 2002), indicates that *ngn3* is a cell-autonomous determinant and true

marker of endocrine progenitor cells. The adoption of each endocrine fate within the islet (α -, β -, δ -, and PP cells) occurs preferentially at specific time points during embryogenesis, suggesting that *ngn3*-positive cells adapt their responses to an evolving milieu of signals. Premature expression of the *ngn3* gene in early pancreatic progenitor cells (e8.5–e9) results in their differentiation into glucagon-producing cells (Apelqvist, Li et al. 1999). In all of the reports so far available the activity of Ngn3 relies as a positive regulator of gene expression at least for its direct known target (NeuroD/BETA2, Pax4, Nkx2.2 and IA1). The activation capacity of Ngn3 has been mapped at the carboxyl terminus of the protein using one hybrid analysis that, instead, does not show any significant intrinsic transcriptional repression capacity. Furthermore, studies of the human *ngn3* promoter have demonstrated that fragments of the promoter itself contain a critical E box which presence is required for activation of the *ngn3* gene (Smith, Watada et al. 2004).

Surprisingly in presence of exogenous Ngn3 this promoter is strongly repressed suggesting that a ubiquitous transcriptional activator may bind to this site, when Ngn3 expression is needed, and that Ngn3 could act as a competitive inhibitor of this activator. This hypothesis was supported by the lack of evidence for significant intrinsic transcriptional repression capacity in the Ngn3 protein, and by the ability of isolated DNA-binding basic helix-loop-helix domains to repress the *ngn3* promoter. Ngn3 produced additional repression, however, when the protein included an intact transcriptional activation domain, suggesting that it may also induce the expression of a downstream transcriptional repressor (Smith, Watada et al. 2004).

The predominant mode of action of a sequence-specific DNA-binding transcriptional repressor in eukaryotes is to recruit co-repressor complexes to its target promoters (active repression). Many sequence-specific repressors recruit histone deacetylases (HDACs), either directly or via adaptor proteins such as Sin3 (Burke and Baniahmad 2000; Thiel, Lietz et al. 2004). HDACs, opposing the function of histone acetylases, catalyze the deacetylation of lysine residues of core histone tails. This results in a more compact chromatin structure and consequently decreased accessibility for transcription factors. Two of the class I HDACs, HDAC1 and HDAC2, has been most widely implicated in transcriptional repression by myriad DNA-binding repressors

(Verdin, Dequiedt et al. 2003). Transcriptional repression can also occur by a passive mechanism, where repressors interfere with the function of transcriptional activators, for example by competing for binding to common DNA sequences [reviewed in (Johnson 1995; Thiel, Lietz et al. 2004)].

The mechanism by which Ngn3 negatively control the activity of *ovoll* gene is in progress.

In conclusion, the ability of Ngn3 to acts as transcriptional repressor could open a novel vision of the pancreas differentiation where start to appear important a correct and fine balance between gene activation and repression.

5.3 Identification of the Transcription factor A2.2

The diabetes mellitus is one of the most common endocrine disorders today, with 180 million people affected in the world in 2003 to approximately 220 million people in 2010 (Zimmet 2000; Zimmet, Alberti et al. 2001). The diabetes is a group of disorders of multiple etiologies characterized by chronic hyperglycemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action or both (WHO 1999). The two most common types of diabetes mellitus are the type 1 and the type 2.

Type 1 diabetes is characterized by an absolute insulin deficiency due to an autoimmune destruction of the insulin producing pancreatic β -cells. Patients with type 1 Diabetes require daily injection of insulin or islet transplants to maintain blood glucose levels within physiological levels. Due to the shortage of islet donors, the most promising approach to cure type 1 Diabetes is the generation of insulin-producing cells from ES cells or from adult/stem cells progenitors.

Type 2 Diabetes is, instead, characterized by a reduction in the insulin secretory cell population and consequently a reduction in insulin secretion that are, along with insulin resistance in liver, muscles and adipose tissues, the main cause of pathological traits of type 2 Diabetes.

For both the pathology the understanding of the fine molecular tuning of insulin gene expression is a prerequisite to envisage new therapeutic approaches.

The molecular mechanism of the β -cell-restricted expression of insulin has been extensively studied during the past years and many of the transcription factors involved in its regulation have been identified (Wang, Shen et al. 1997; Ohneda, Ee et al. 2000). Nevertheless many aspects of the insulin-gene regulation and hormone production are still unclear. For example, the restriction of the proinsulin expression to the β -cell is still unexplained.

The insulin gene is a small gene, located on the chromosome 11q15.5 in humans and consisting of three exons and two introns (Owerbach, Bell et al. 1980). The sequences immediately upstream of the transcription start site are important control elements for its expression in the β -cell (Walker, Edlund et al. 1983; Edlund, Walker et al. 1985; Hanahan 1985). Most of the data available identified a short region of about 400 bp in the proximal insulin promoter as responsible for the “segregation” of insulin expression into the β -cells (Ohneda, Ee et al. 2000). Many *cis*-regulatory elements have been identified; among them the most important are the conserved E1 (-100 to -91 bp), A3 (-201 to -196 bp), and RIPE3b/C1-A2 (-126 to -101 bp) elements (Karlsson, Edlund et al. 1987; Whelan, Poon et al. 1989; Shieh and Tsai 1991). The islet-restricted transcription factors NeuroD/BETA2 and Pdx-1/IPF1/STF1/IDX1/GSF/IUF1, which bind to the E1 and A3 elements, respectively, have been identified as well. Gene disruption experiments in mice have revealed that both NeuroD/BETA2 and PDX-1 play critical roles in insulin gene expression as well as in islet development and function (Jonsson, Carlsson et al. 1994; Naya, Stellrecht et al. 1995; Ahlgren, Jonsson et al. 1998). Furthermore mutations in *beta2* and *pdx1*, genes are found in some population of patients with type 2 diabetes (Stoffers, Ferrer et al. 1997; Stoffers, Zinkin et al. 1997; Malecki, Jhala et al. 1999).

The third regulatory element, RIPE3b/C1, has been shown to play an important role in *insulin* gene transcription and in its glucose-regulated expression (Shieh and Tsai 1991; Sharma and Stein 1994; Wang, Brun et al. 2007). Previous studies have identified a β -cell-restricted RIPE3b-binding factor, called RIPE3b1 activator that appears in response to glucose in pancreatic nuclear extract (Sharma, Fusco-DeMane et al. 1995).

Using biochemical purifications and band-shift assay approaches, the transcription factor RIPE3b1 has been recently identified as a mammalian homologue of avian MafA/L-Maf (Olbrot, Rud et al. 2002). Maf family proteins belong to the basic leucine zipper (bZIP) family of transcription factors. To date, several *maf*-related genes have been identified in various species including human, mouse, rat, chicken, frog and zebrafish. The DNA consensus sequence of the Maf family of transcription factors has been determined by independent works (Kataoka, Noda et al. 1994; Kerppola and Curran 1994; Kerppola and Curran 1994) to be a palindromic sequence of 13-14 bp (TGC(N)₆₋₇GCA), now called Maf Recognition Element (MARE). MARE sequences have been identified in different promoter regions (Kataoka, Noda et al. 1994) that turned out to be targets of the Maf family members (*e.g.* α - and β -globin, opsin, crystalline genes and insulin). The RIPE3b/C1-A2 region, within the insulin promoter region, has been characterized by elegant and complete studies (Harrington and Sharma 2001; Nishimura, Salameh et al. 2005). The insulin RIPE3b/C1-A2 enhancer element contains the C1 element (-116 to 107 bp), bound by MafA, and the A2 element (-126 to -113 bp) (Shieh and Tsai 1991; German, Ashcroft et al. 1995; Tomonari, Yoshimoto et al. 1996; Zhao, Guo et al. 2005). Furthermore, the A2 *insulin* enhancer element was classified, based on the sequence homology, as a A-elements (German, Ashcroft et al. 1995). Transcription factors belonging to the homeodomain transcription factor family, such PDX-1, recognize these sequences (Karlsson, Thor et al. 1990; Rudnick, Ling et al. 1994). In addition, the region RIPE3b/C1-A2 contains also an element recognized by transcription factor belonging to the Nkx2 family (Damante, Fabbro et al. 1994; Weiler, Gruschus et al. 1998). According to that it was first hypothesized that both PDX-1 and Nkx2.2 could correspond to the A2 element. However, the work of Harrington et al. (Harrington and Sharma 2001) has demonstrated that neither of them were the A2-binding complex. Instead, when analyzed using DNA shift assays, the RIPE3b/C1-A2 element region showed the ability to form five different DNA/protein complexes named RIPE3b1, RIPE3b2, A2.1, A2.2 and A2.3 (Olbrot, Rud et al. 2002). All the complexes were formed using nuclear extracts from non-insulin producing cells while only RIPE3b1 (MafA) and A2.2 were formed exclusively using nuclear extracts from insulin-producing cells (Harrington and Sharma 2001). Interestingly it has been demonstrated that both the C1 and A2 element share

overlapping nucleotides, in particular two base pairs (-122 and -121) essential for the binding activity of both the complexes (Nishimura, Salameh et al. 2005). Moreover, it emerged that C1 and A2 binding complexes can act cooperatively to regulate the insulin gene expression (Harrington and Sharma 2001).

Transcriptional regulation of insulin expression results not only from the specific combinations of the C1 and A2.2 activators through DNA-protein and protein-protein interactions, but also from their relative nuclear concentrations, generating a cooperatively transcriptional synergism unique for the insulin gene.

Despite the importance for the β -cell-specific role played by A2.2 in the regulation of *insulin* gene expression its nature is still elusive. According to that it is clear that the fully understanding of insulin regulator transcription factors will have fundamental impact for the approaches to cure type 1 and type 2 Diabetes.

For these reasons the aims of the present part of the Thesis have been:

- 1) *To characterize and purify the transcription factor A2.2*
- 2) *To identify it as β -cell specific transcription factor*

5.3.1 Characterization of A2.2

The 5' flanking region of the rat *insulin II* gene corresponding to the region containing the A2 element has been analyzed for the relative homology with the corresponding region of rat insulin I and human insulin. The consensus sequence retrieved has been then used for the identification of putative transcription factors by using the TRANSFAC software (Nishimura, Salameh et al. 2005). The results identified the A2 sequence as highly homologous to those recognized by bZIP transcription factors (Nishimura, Salameh et al. 2005). Furthermore among the protein identified, the higher scores were attributed to the proteins belonging to the C/EBP transcription factors family.

The high homology of A2 sequence with those recognized by bZIP proteins is in good agreement with the observation that many bZIP domain-containing proteins play

important roles during pancreas development (Shih and Stoffel 2001; Wang, Webb et al. 2003; Towle 2005; Tsuchiya, Taniguchi et al. 2006).

This has been the starting point to develop an experimental strategy aimed to address the first point of the proposed aims (*To characterize and purify the transcription factor A2.2*). The band shift competition assays described in the “Materials and Methods” section and discussed in details in the “Results” section confirmed the originality of the working hypothesis. In fact, the oligonucleotide probes used (containing sequences recognized by several bZIP transcription factors) were able to compete with the A2 sequence for the A2.2 binding.

Furthermore the high score attributed by the TRANSFAC analysis to the proteins belonging to the C/EBP transcription factors family resulted almost “unusual” since it is known that C/EBP act as negative regulator of the transcription (Lu, Seufert et al. 1997; Seufert, Weir et al. 1998) while the A2.2 transcription factor acts, undoubtedly, as a positive regulator for the *insulin* gene expression. Nevertheless, band shift competition assays using the DNA binding domain for C/EBP transcription factors and super-shift assays, using anti-C/EBP α and anti-C/EBP β antibodies, were performed. Once again, the results confirmed the ability of oligonucleotides containing C/EBP-DNA binding elements to compete for the formation of the A2.2 complex. Furthermore, a strong super shift signal, but not the disappearance of the A2.2 complex, was observed in presence of anti-C/EBP α and anti-C/EBP β antibodies. Further experiments were performed using plasmids expressing dominant negative of the bZIP proteins that selectively inhibit transcriptional activation mediated by different bZIP sub-family members. None of these recombinant proteins resulted in an inhibition of the A2.2 formation.

The first two observations strongly support the hypothesis that in the formation of the A2.2 complex the C/EBP transcription factors play an important role. Unfortunately they appear to act more as a “co-actor” rather than as a specific and unique “actor” of the formation of the A2.2 complex.

The third observation opens the possibility of a more challenging “scenario”; the possibility that the A2.2 complex belongs to the C/EBP transcription factors family but corresponding to a “brand new” protein with new and unexpected behavior(s)

Although “disappointing”, from certain point of view, these observations added important pieces of the A2.2 puzzle prompting us to challenge the purification “from scratch” of the A2.2 protein.

5.3.2 Purification of A2.2

For the purification of the A2.2 transcription factor the cell line HIT T-15 has been chosen. The HIT T-15 cells derive from dispersed Syrian hamster pancreatic islets following transformation with SV40 (Santerre, Cook et al. 1981). This cell line was chosen because it shows the highest binding activity to the A2 binding element (over 10 folds) respect other insulin producing cell lines (Harrington and Sharma 2001). To obtain the highest degree of purity with the lowest number of purification process a two-step chromatography strategy has been developed. The first chromatographic separation relies on the use of the HiTrap-Heparin column. The chromatographic support of the HiTrap-Heparin column allows for fast, simple and easy separations of nucleic acid binding proteins. In fact, immobilized heparin interacts with nucleic acid-binding proteins as a cation exchanger due to its high content of anionic sulfate groups mimicking the polyanionic structure of the nucleic acid. The second chromatographic separation relies on the use of an affinity chromatography step on a streptavidin-agarose column. This chromatographic support is able to bind irreversibly, in mild biochemical conditions, its ligand, namely the biotin. For this highly selective purification step the purified and pooled protein fractions from the previous steps were incubated with a biotinylated A2 oligonucleotide. This acts as a bait to “capture” the DNA binding protein(s) that recognize with high specificity the A2 consensus sequence. The complexes DNA/protein(s) are then captured by the specific interaction biotin/streptavidin. The highly stringent condition of washing should allow removing all spurious proteins. The elution of the bound proteins was obtained by competition of the protein(s) with untagged A2 oligonucleotides. Each of the fraction obtained (about 30) were assayed for their ability to form A2.2 complexes in EMSA assays. As reported in the “results” section at least three fractions showed a significant A2-binding specificity. A further

characterization of these fractions has been obtained running a SDS-PAGE. Although the results showed a certain unexpected heterogeneity (at least 5 distinct bands were obtained) it was not so dramatic. In fact it has been possible to easily extract with high purity some of the bands of greater interest. In particular our attention was concentrated on the bands showing a low molecular weight. According to unpublished data from Dr. Sharma the mass range of the expected protein(s) should be between 15 and 20 kDa (Dr. Sharma, personal communication). Only one band present in this mass range showed a significant A2-binding specificity. This band has been sent to the Joslin Proteomic Core for mass spectrometry analysis.

The output of the mass spectrometry analysis identified with the higher score of homology proteins belonging to the small Maf family factors. The small Maf proteins consist of only three members: MafF, MafG and MafK.

The small Maf proteins have very similar function (Onodera, Shavit et al. 1999); they recognize the same MARE sequences of the Maf factors and tend to form homodimers or heterodimers with other bZIP factors, such as CNC (“cap’n’collar”) or Bach family proteins (Andrews, Erdjument-Bromage et al. 1993; Igarashi, Kataoka et al. 1994; Itoh, Igarashi et al. 1995; Oyake, Itoh et al. 1996; Marini, Chan et al. 1997). The Small Maf factors lack the trans-activation domain but CNC and Bach proteins have a transactivation (CNC) or a repression (Bach) domain unique to each molecule. Therefore through heterodimerization, the small Maf proteins acquire cis-acting properties (activation or repression).

The presence of small Maf factors within eluted A2.2 positive fraction has been experimentally confirmed. As observed for the in the C/EBP transcription factors also for the small Maf could be hypothesized a contribution in the formation of the A2.2 complex as “guest star” rather than as the only responsible for the formation of the A2.2 complex.

Next the bioinformatics analysis was extended also to those proteins showing lower homology but retaining some constrains:

- 1) The nuclear localization.
- 2) The low molecular weight (15-20kDa).
- 3) The cell specificity (β -cells, if the information was available from the database).

The expression pattern for the selected candidates was then analyzed by RT-PCR on insulin and non-insulin producing cells. All the candidates failed to satisfy the third constrain, the β -cell specificity.

In the attempt to broaden the analysis the results of the mass spectrometry some protein database(see “Materials and Methods”) have been analyzed loosening the homology allowing a limited degree of degeneracy among the amino acids sequences and adding, to the constrains described above, the belonging to the bZIP proteins family. A list of nine proteins was obtained. Among them six possible candidates were selected to assess their expression pattern by RT-PCR on insulin and non-insulin producing cells.

Once again the candidates failed to satisfy the β -cell specificity.

In conclusion the identification of the A2.2 factors is still an open question although some important information has been collected. The A2.2 complex is probably the result of the cooperative activity of different proteins that include member of the bZIP and small Maf families. A serious drawback that impaired the analysis is the lack of a sufficiently complete nucleotide and protein database for hamster. All the bioinformatics analyses have been performed assuming a good degree of homology between hamster and rat. Despite the close relation between the two species (both belonging to the Rodents order) the evolutive pressure could had act selecting protein in which the degree of homology (at least in term of amino acids sequences) could be low enough to complicate the cross-species homology analyses.

As an alternative strategy could be worth to consider the possibility to start over the purification steps from nuclear extracts obtained from rat or mouse which database (both DNA and proteins) are as far the most complete available.

Chapter 6

CONCLUSIONS

6 CONCLUSIONS

Diabetes currently affects close to 50 million people in Europe and this number is expected to increase to almost 60 million by 2025. There is, therefore, an urgent need to take action to address this diabetic time bomb. Apart the social and medical issues arising from this disease there are also challenging aspects for basic and clinical research. From a medical point of view maintaining glycemic control with exogenous insulin is the classical therapeutic approach for type 1 Diabetes. However, this approach is associated with an increased risk of severe hypoglycemia. One alternative treatment for type 1 Diabetes is the whole organ pancreas transplantation. While the results of these studies are very encouraging, multiple donor pancreata were required to achieve insulin independence, and immunosuppression is still needed.

Contribution from the research in the challenging field of the embryonic stem (ES) cells could be very important. Insulin secreting-cells derived from ES cells may offer a solution to the problem of an inadequate supply of insulin-producing cells. Theoretically, ES cells can be differentiated by the use of genetic and epigenetic manipulations.

There is a growing research interest for possible use of pluripotent ES cells to be used for the treatment of type 1 Diabetes. Recent studies showed that cultured ES cells can differentiate, at low frequencies, into insulin-producing cells [43, 44]. Unfortunately these approaches are quite empirical and are based merely on a so-called “Medium Engineering”. This approach is based on use of media with different formulation obtained by addition or withdrawal of supplements, growth factors, etc.

Almost two aspects impair the use of the “Medium Engineering” approach:

1. Insulin-secreting cells obtained not always show a physiological response to the variation of glucose concentration.
2. The population heterogeneity is still too high.

Inducing β -cell differentiation by expression of transcription factors that are necessary for β -cells development is a possible approach for gene therapy in diabetes.

The specific pathway leading to the development of the insulin-secreting β -cells is particularly important to be able to engineer new β -cells.

It is clear that endocrine cell fate appears to be governed by the expression of the transcription factor Ngn3 that acts as a switch for differentiation activating, downstream, a molecular cascade of factors that commit the progenitor cell to the endocrine fate.

On this aspect is based the novelty of the approach developed in this Thesis. Understanding the role of this molecular switch could have potential important outcomes for the diabetes. In fact understanding the gene pathway activated by Ngn3 during β -cell differentiation will *increase the basic knowledge on β -cell function* that could allow for development of reproducible *in vitro* methods for differentiation to:

1. Development of reproducible *in vitro* methods for differentiation to:
 - Generation of glucose-responsive and insulin-secreting cells from ES cells;
 - Generation of glucose-responsive and insulin-secreting cells from adult progenitor cells;
 - Generation of glucose-responsive and insulin-secreting cells from non-b cells;
2. Development of clinical and preclinical trials;
3. Development of insulin-secreting cell lines suitable for therapeutic use in humans.

Yet, *Insulin* gene therapy is another approach for the possible cure of type 1 Diabetes that might overcome the weakness of islet transplantation and β -cell regeneration with respect their vulnerability towards autoimmune attack. Many attempts have been made to replace the function of β -cells by introducing various components of the insulin secretory machinery into non β -cells, which are no targets of the β -cell-specific autoimmune response. However, the simplest replacement of insulin gene expression by genetic engineering is not likely to be useful as a therapy for type 1 Diabetes unless it contains an appropriate system to regulate insulin levels within extremely narrow physiological limits in response to glucose. Successful insulin gene therapy requires several important components including an effective insulin gene delivery system, a regulatory system tightly controlled by glucose, an appropriate processing system and an appropriate target cell with biochemical characteristics similar

to β -cells. In this respect it becomes obvious that a fully understanding of insulin gene regulation in β -cell must be achieved. Many studies have contributed to expand our knowledge regarding how insulin expression is restricted to pancreatic β -cell and how its production is regulated by physiologic signals such as glucose. A picture is emerging in which an elaborate set of transcription factors bind to specific sequences along the promoter, recruiting additional “co-actors” to build a functional transcriptional-activation-complex that is unique of the β -cells. Nonetheless the complete scenario of insulin expression within the β -cell is still full of shade. On this respect the setup of a **reproducible method of purification and the characterization/identification of the A2.2 transcription factor**, still “in progress”, will help so far to add a new piece to the puzzle on the wisdom of *insulin* gene regulation.

In conclusion, the project developed in this Thesis was strictly oriented to contribute to find a solution of an emerging social and medical problem such as Type 1 Diabetes starting from *original* idea that are to mimic as much as possible physiological events and identify the and using *innovative* cell and gene therapy approaches building a new concept of Medicine that is the Molecular Medicine.

REFERENCES

REFERENCES

- Ahlgren, U., J. Jonsson, et al. (1996). "The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice." Development **122**(5): 1409-16.
- Ahlgren, U., J. Jonsson, et al. (1998). "beta-cell-specific inactivation of the mouse *Ipf1/Pdx1* gene results in loss of the beta-cell phenotype and maturity onset diabetes." Genes Dev **12**(12): 1763-8.
- Ahlgren, U., S. L. Pfaff, et al. (1997). "Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells." Nature **385**(6613): 257-60.
- Aiello, L. P. (2005). "Angiogenic pathways in diabetic retinopathy." N Engl J Med **353**(8): 839-41.
- Akerblom, H. K., O. Vaarala, et al. (2002). "Environmental factors in the etiology of type 1 diabetes." Am J Med Genet **115**(1): 18-29.
- Amati, B. and H. Land (1994). "Myc-Max-Mad: a transcription factor network controlling cell cycle progression, differentiation and death." Curr Opin Genet Dev **4**(1): 102-8.
- Andrews, N. C., H. Erdjument-Bromage, et al. (1993). "Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein." Nature **362**(6422): 722-8.
- Ang, S. L. and J. Rossant (1994). "HNF-3 beta is essential for node and notochord formation in mouse development." Cell **78**(4): 561-74.
- Ang, S. L., A. Wierda, et al. (1993). "The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins." Development **119**(4): 1301-15.
- Apelqvist, A., H. Li, et al. (1999). "Notch signalling controls pancreatic cell differentiation." Nature **400**(6747): 877-81.
- Appel, B., L. A. Givan, et al. (2001). "Delta-Notch signaling and lateral inhibition in zebrafish spinal cord development." BMC Dev Biol **1**: 13.
- Aronheim, A., H. Ohlsson, et al. (1991). "Distribution and characterization of helix-loop-helix enhancer-binding proteins from pancreatic beta cells and lymphocytes." Nucleic Acids Res **19**(14): 3893-9.
- Asfari, M., D. Janjic, et al. (1992). "Establishment of 2-mercaptoethanol-dependent differentiated insulin-secreting cell lines." Endocrinology **130**(1): 167-78.
- Baek, S. H., C. Kiousi, et al. (2003). "Regulated subset of G1 growth-control genes in response to derepression by the Wnt pathway." Proc Natl Acad Sci U S A **100**(6): 3245-50.
- Bagheri-Fam, S., F. Barrionuevo, et al. (2006). "Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal *Sox9* expression pattern." Dev Biol **291**(2): 382-97.
- Bajaj, M. and R. A. Defronzo (2003). "Metabolic and molecular basis of insulin resistance." J Nucl Cardiol **10**(3): 311-23.

- Beck-Nielsen, H. and L. C. Groop (1994). "Metabolic and genetic characterization of prediabetic states. Sequence of events leading to non-insulin-dependent diabetes mellitus." *J Clin Invest* **94**(5): 1714-21.
- Bevington, P. R. and D. K. Robinson (1991). "Reduction and Error Analysis for the Physical Sciences." *McGraw-Hill, New York*.
- Bhushan, A., N. Itoh, et al. (2001). "Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis." *Development* **128**(24): 5109-17.
- Bien-Willner, G. A., P. Stankiewicz, et al. (2007). "SOX9^{cre1}, a cis-acting regulatory element located 1.1 Mb upstream of SOX9, mediates its enhancement through the SHH pathway." *Hum Mol Genet* **16**(10): 1143-56.
- Bitgood, M. J. and A. P. McMahon (1995). "Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo." *Dev Biol* **172**(1): 126-38.
- Blackwell, T. K., J. Huang, et al. (1993). "Binding of myc proteins to canonical and noncanonical DNA sequences." *Mol Cell Biol* **13**(9): 5216-24.
- Blackwell, T. K., L. Kretzner, et al. (1990). "Sequence-specific DNA binding by the c-Myc protein." *Science* **250**(4984): 1149-51.
- Blackwood, E. M., B. Luscher, et al. (1992). "Myc and Max associate in vivo." *Genes Dev* **6**(1): 71-80.
- Bohlender, J. M., S. Franke, et al. (2005). "Advanced glycation end products and the kidney." *Am J Physiol Renal Physiol* **289**(4): F645-59.
- Bonior, J., J. Jaworek, et al. (2005). "Increase of heat shock protein gene expression by melatonin in AR42J cells." *J Physiol Pharmacol* **56**(3): 471-81.
- Bonner-Weir, S. (2001). "beta-cell turnover: its assessment and implications." *Diabetes* **50 Suppl 1**: S20-4.
- Bratanova-Tochkova, T. K., H. Cheng, et al. (2002). "Triggering and augmentation mechanisms, granule pools, and biphasic insulin secretion." *Diabetes* **51 Suppl 1**: S83-90.
- Breslin, M. B., M. Zhu, et al. (2003). "NeuroD1/E47 regulates the E-box element of a novel zinc finger transcription factor, IA-1, in developing nervous system." *J Biol Chem* **278**(40): 38991-7.
- Brink, C. (2003). "Promoter elements in endocrine pancreas development and hormone regulation." *Cell Mol Life Sci* **60**(6): 1033-48.
- Briscoe, J., A. Pierani, et al. (2000). "A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube." *Cell* **101**(4): 435-45.
- Britsch, S., D. E. Goerich, et al. (2001). "The transcription factor Sox10 is a key regulator of peripheral glial development." *Genes Dev* **15**(1): 66-78.
- Brooke, N. M., J. Garcia-Fernandez, et al. (1998). "The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster." *Nature* **392**(6679): 920-2.
- Brorson, M., D. M. Hougaard, et al. (2001). "Expression of SMAD signal transduction molecules in the pancreas." *Histochem Cell Biol* **116**(3): 263-7.
- Burkart, V., H. Liu, et al. (2000). "Natural resistance of human beta cells toward nitric oxide is mediated by heat shock protein 70." *J Biol Chem* **275**(26): 19521-8.
- Burke, L. J. and A. Baniahmad (2000). "Co-repressors 2000." *Faseb J* **14**(13): 1876-88.

- Burke, Z. D., C. N. Shen, et al. (2006). "Characterization of liver function in transdifferentiated hepatocytes." J Cell Physiol **206**(1): 147-59.
- Burke, Z. D., C. N. Shen, et al. (2004). "Bile ducts as a source of pancreatic beta cells." Bioessays **26**(9): 932-7.
- Busch, C. P. and R. A. Hegele (2001). "Genetic determinants of type 2 diabetes mellitus." Clin Genet **60**(4): 243-54.
- Catasti, P., X. Chen, et al. (1996). "Structure-function correlations of the insulin-linked polymorphic region." J Mol Biol **264**(3): 534-45.
- Chae, J. H., G. H. Stein, et al. (2004). "NeuroD: the predicted and the surprising." Mol Cells **18**(3): 271-88.
- Chan, S. J., P. Keim, et al. (1976). "Cell-free synthesis of rat preproinsulins: characterization and partial amino acid sequence determination." Proc Natl Acad Sci U S A **73**(6): 1964-8.
- Chatterjee, S. and B. Price (1991). "Regression Analysis by Example." John Wiley & Sons, New York.
- Chen, L., H. X. Yan, et al. (2007). "Negative regulation of c-Myc transcription by pancreas duodenum homeobox-1." Endocrinology **148**(5): 2168-80.
- Chen, Y., E. R. Dougherty, et al. (1997). "Ratio-based decisions and the quantitative analysis of cDNA microarray images." J. Biomed. Optics **2**: 11.
- Chen, Y. G., Q. Wang, et al. (2006). "Activin signaling and its role in regulation of cell proliferation, apoptosis, and carcinogenesis." Exp Biol Med (Maywood) **231**(5): 534-44.
- Chiang, M. K. and D. A. Melton (2003). "Single-cell transcript analysis of pancreas development." Dev Cell **4**(3): 383-93.
- Cogswell, J. P., P. C. Cogswell, et al. (1993). "Mechanism of c-myc regulation by c-Myb in different cell lineages." Mol Cell Biol **13**(5): 2858-69.
- Crisera, C. A., T. S. Maldonado, et al. (2000). "Transforming growth factor-beta 1 in the developing mouse pancreas: a potential regulator of exocrine differentiation." Differentiation **65**(5): 255-9.
- da Silva Xavier, G., A. Varadi, et al. (2000). "Regulation of gene expression by glucose in pancreatic beta -cells (MIN6) via insulin secretion and activation of phosphatidylinositol 3'-kinase." J Biol Chem **275**(46): 36269-77.
- Dahlquist, G. (1998). "The aetiology of type 1 diabetes: an epidemiological perspective." Acta Paediatr Suppl **425**: 5-10.
- Dai, X., C. Schonbaum, et al. (1998). "The ovo gene required for cuticle formation and oogenesis in flies is involved in hair formation and spermatogenesis in mice." Genes Dev **12**(21): 3452-63.
- Damante, G., D. Fabbro, et al. (1994). "Sequence-specific DNA recognition by the thyroid transcription factor-1 homeodomain." Nucleic Acids Res **22**(15): 3075-83.
- Davies, P. O., C. Poirier, et al. (1994). "Genetic reassignment of the insulin-1 (Ins1) gene to distal mouse chromosome 19." Genomics **21**(3): 665-7.
- De Vos, A., H. Heimberg, et al. (1995). "Human and rat beta cells differ in glucose transporter but not in glucokinase gene expression." J Clin Invest **96**(5): 2489-95.
- Demeterco, C., P. Itkin-Ansari, et al. (2002). "c-Myc controls proliferation versus differentiation in human pancreatic endocrine cells." J Clin Endocrinol Metab **87**(7): 3475-85.

- Dickmeis, T., P. Mourrain, et al. (2001). "A crucial component of the endoderm formation pathway, CASANOVA, is encoded by a novel sox-related gene." Genes Dev **15**(12): 1487-92.
- Dieterlen-Lievre, F. and D. Beaupain (1974). "[Immunocytological study of the differentiation of the endocrine pancreas in the chick embryo. I. Islands of Langerhans]." Gen Comp Endocrinol **22**(1): 62-9.
- Do, H., D. Falcone, et al. (1996). "The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process." Cell **85**(3): 369-78.
- Dodson, G. and D. Steiner (1998). "The role of assembly in insulin's biosynthesis." Curr Opin Struct Biol **8**(2): 189-94.
- Dohrmann, C., P. Gruss, et al. (2000). "Pax genes and the differentiation of hormone-producing endocrine cells in the pancreas." Mech Dev **92**(1): 47-54.
- Dominguez-Bendala, J., D. Klein, et al. (2005). "TAT-mediated neurogenin 3 protein transduction stimulates pancreatic endocrine differentiation in vitro." Diabetes **54**(3): 720-6.
- Dudek, R. W. and I. E. Lawrence, Jr. (1988). "Morphologic evidence of interactions between adult ductal epithelium of pancreas and fetal foregut mesenchyme." Diabetes **37**(7): 891-900.
- Dudek, R. W., I. E. Lawrence, Jr., et al. (1991). "Induction of islet cytodifferentiation by fetal mesenchyme in adult pancreatic ductal epithelium." Diabetes **40**(8): 1041-8.
- Dudoit, S., R. C. Genteman, et al. (2003). "Open Source Software for the Analysis of Microarray Data." BioTechniques **19**(2): 9.
- Duvillie, B., D. Bucchini, et al. (1998). "Imprinting at the mouse Ins2 locus: evidence for cis- and trans-allelic interactions." Genomics **47**(1): 52-7.
- Echelard, Y., D. J. Epstein, et al. (1993). "Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity." Cell **75**(7): 1417-30.
- Edlund, H. (1998). "Transcribing pancreas." Diabetes **47**(12): 1817-23.
- Edlund, H. (2001). "Developmental biology of the pancreas." Diabetes **50 Suppl 1**: S5-9.
- Edlund, T., M. D. Walker, et al. (1985). "Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5' flanking elements." Science **230**(4728): 912-6.
- Efrat, S., M. Tal, et al. (1994). "The pancreatic beta-cell glucose sensor." Trends Biochem Sci **19**(12): 535-8.
- Elliott, J. G., J. D. Oliver, et al. (1997). "The thiol-dependent reductase ERp57 interacts specifically with N-glycosylated integral membrane proteins." J Biol Chem **272**(21): 13849-55.
- Emens, L. A., D. W. Landers, et al. (1992). "Hepatocyte nuclear factor 1 alpha is expressed in a hamster insulinoma line and transactivates the rat insulin I gene." Proc Natl Acad Sci U S A **89**(16): 7300-4.
- Esni, F., B. Ghosh, et al. (2004). "Notch inhibits Ptf1 function and acinar cell differentiation in developing mouse and zebrafish pancreas." Development **131**(17): 4213-24.
- Evan, G. I. and T. D. Littlewood (1993). "The role of c-myc in cell growth." Curr Opin Genet Dev **3**(1): 44-9.

- Evans, J. L., T. L. Moore, et al. (1990). "Functional analysis of c-Myb protein in T-lymphocytic cell lines shows that it trans-activates the c-myc promoter." Mol Cell Biol **10**(11): 5747-52.
- Fajans, S. S., G. I. Bell, et al. (2001). "Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young." N Engl J Med **345**(13): 971-80.
- Foster, L. J. and A. Klip (2000). "Mechanism and regulation of GLUT-4 vesicle fusion in muscle and fat cells." Am J Physiol Cell Physiol **279**(4): C877-90.
- Francis, J., S. K. Chakrabarti, et al. (2005). "Pdx-1 links histone H3-Lys-4 methylation to RNA polymerase II elongation during activation of insulin transcription." J Biol Chem **280**(43): 36244-53.
- Freiesleben De Blasio, B., P. Bak, et al. (1999). "Onset of type 1 diabetes: a dynamical instability." Diabetes **48**(9): 1677-85.
- Froguel, P., H. Zouali, et al. (1993). "Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus." N Engl J Med **328**(10): 697-702.
- Furukawa, M., Y. Eto, et al. (1995). "Expression of immunoreactive activin A in fetal rat pancreas." Endocr J **42**(1): 63-8.
- Gao, Y., J. Miyazaki, et al. (2003). "The transcription factor PDX-1 is post-translationally modified by O-linked N-acetylglucosamine and this modification is correlated with its DNA binding activity and insulin secretion in min6 beta-cells." Arch Biochem Biophys **415**(2): 155-63.
- Garfinkel, M. D., J. Wang, et al. (1994). "Multiple products from the shavenbaby-ovo gene region of *Drosophila melanogaster*: relationship to genetic complexity." Mol Cell Biol **14**(10): 6809-18.
- Gasa, R., C. Mrejen, et al. (2004). "Proendocrine genes coordinate the pancreatic islet differentiation program in vitro." Proc Natl Acad Sci U S A **101**(36): 13245-50.
- Gasa, R., C. Mrejen, et al. (2007). "Induction of pancreatic islet cell differentiation by the neurogenin-neuroD cascade." Differentiation.
- Gat, U., R. DasGupta, et al. (1998). "De Novo hair follicle morphogenesis and hair tumors in mice expressing a truncated beta-catenin in skin." Cell **95**(5): 605-14.
- Genteman, R. C., V. J. Carey, et al. (2004). "Bioconductor: open software development for computational biology and bioinformatics." Genome Biol **5**:R80.
- German, M., S. Ashcroft, et al. (1995). "The insulin gene promoter. A simplified nomenclature." Diabetes **44**(8): 1002-4.
- Giddings, S. J. and L. R. Carnaghi (1988). "The two nonallelic rat insulin mRNAs and pre-mRNAs are regulated coordinately in vivo." J Biol Chem **263**(8): 3845-9.
- Gierl, M. S., N. Karoulias, et al. (2006). "The zinc-finger factor Insm1 (IA-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells." Genes Dev **20**(17): 2465-78.
- Gilmore, R., P. Walter, et al. (1982). "Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor." J Cell Biol **95**(2 Pt 1): 470-7.
- Gittes, G. K., P. E. Galante, et al. (1996). "Lineage-specific morphogenesis in the developing pancreas: role of mesenchymal factors." Development **122**(2): 439-47.

- Gittes, G. K. and W. J. Rutter (1992). "Onset of cell-specific gene expression in the developing mouse pancreas." Proc Natl Acad Sci U S A **89**(3): 1128-32.
- Glick, E., D. Leshkowitz, et al. (2000). "Transcription factor BETA2 acts cooperatively with E2A and PDX1 to activate the insulin gene promoter." J Biol Chem **275**(3): 2199-204.
- Golosow, N. and C. Grobstein (1962). "Epitheliomesenchymal interaction in pancreatic morphogenesis." Dev Biol **4**: 242-55.
- Goodge, K. A. and J. C. Hutton (2000). "Translational regulation of proinsulin biosynthesis and proinsulin conversion in the pancreatic beta-cell." Semin Cell Dev Biol **11**(4): 235-42.
- Goto, Y., M. G. De Silva, et al. (1992). "A novel human insulinoma-associated cDNA, IA-1, encodes a protein with "zinc-finger" DNA-binding motifs." J Biol Chem **267**(21): 15252-7.
- Gradwohl, G., A. Dierich, et al. (2000). "neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas." Proc Natl Acad Sci U S A **97**(4): 1607-11.
- Grapin-Botton, A., A. R. Majithia, et al. (2001). "Key events of pancreas formation are triggered in gut endoderm by ectopic expression of pancreatic regulatory genes." Genes Dev **15**(4): 444-54.
- Gu, G., J. Dubauskaite, et al. (2002). "Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors." Development **129**(10): 2447-57.
- Guz, Y., M. R. Montminy, et al. (1995). "Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny." Development **121**(1): 11-8.
- Habener, J. F. and D. A. Stoffers (1998). "A newly discovered role of transcription factors involved in pancreas development and the pathogenesis of diabetes mellitus." Proc Assoc Am Physicians **110**(1): 12-21.
- Hald, J., J. P. Hjorth, et al. (2003). "Activated Notch1 prevents differentiation of pancreatic acinar cells and attenuate endocrine development." Dev Biol **260**(2): 426-37.
- Hale, M. A., H. Kagami, et al. (2005). "The homeodomain protein PDX1 is required at mid-pancreatic development for the formation of the exocrine pancreas." Dev Biol **286**(1): 225-37.
- Hanahan, D. (1985). "Heritable formation of pancreatic beta-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes." Nature **315**(6015): 115-22.
- Harper, M. E., A. Ullrich, et al. (1981). "Localization of the human insulin gene to the distal end of the short arm of chromosome 11." Proc Natl Acad Sci U S A **78**(7): 4458-60.
- Harrington, R. H. and A. Sharma (2001). "Transcription factors recognizing overlapping C1-A2 binding sites positively regulate insulin gene expression." J Biol Chem **276**(1): 104-13.
- Harrison, K. A., J. Thaler, et al. (1999). "Pancreas dorsal lobe agenesis and abnormal islets of Langerhans in Hlx9-deficient mice." Nat Genet **23**(1): 71-5.

- Hart, A., S. Papadopoulou, et al. (2003). "Fgf10 maintains notch activation, stimulates proliferation, and blocks differentiation of pancreatic epithelial cells." *Dev Dyn* **228**(2): 185-93.
- Hayashi, S. and A. P. McMahon (2002). "Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse." *Dev Biol* **244**(2): 305-18.
- He, T. C., A. B. Sparks, et al. (1998). "Identification of c-MYC as a target of the APC pathway." *Science* **281**(5382): 1509-12.
- Healey, D., P. Ozegbe, et al. (1995). "In vivo activity and in vitro specificity of CD4+ Th1 and Th2 cells derived from the spleens of diabetic NOD mice." *J Clin Invest* **95**(6): 2979-85.
- Hebrok, M., S. K. Kim, et al. (1998). "Notochord repression of endodermal Sonic hedgehog permits pancreas development." *Genes Dev* **12**(11): 1705-13.
- Heit, J. J., S. K. Karnik, et al. (2006). "Intrinsic regulators of pancreatic beta-cell proliferation." *Annu Rev Cell Dev Biol* **22**: 311-38.
- Heller, R. S., D. S. Dichmann, et al. (2002). "Expression patterns of Wnts, Frizzleds, sFRPs, and misexpression in transgenic mice suggesting a role for Wnts in pancreas and foregut pattern formation." *Dev Dyn* **225**(3): 260-70.
- Heller, R. S., T. Klein, et al. (2003). "Expression of Wnt, Frizzled, sFRP, and DKK genes in adult human pancreas." *Gene Expr* **11**(3-4): 141-7.
- Heller, R. S., D. A. Stoffers, et al. (2004). "The role of Brn4/Pou3f4 and Pax6 in forming the pancreatic glucagon cell identity." *Dev Biol* **268**(1): 123-34.
- Henquin, J. C., N. Ishiyama, et al. (2002). "Signals and pools underlying biphasic insulin secretion." *Diabetes* **51 Suppl 1**: S60-7.
- Heremans, Y., M. Van De Casteele, et al. (2002). "Recapitulation of embryonic neuroendocrine differentiation in adult human pancreatic duct cells expressing neurogenin 3." *J Cell Biol* **159**(2): 303-12.
- Herrera, P. L., L. Orci, et al. (1998). "Two transgenic approaches to define the cell lineages in endocrine pancreas development." *Mol Cell Endocrinol* **140**(1-2): 45-50.
- High, S., B. Martoglio, et al. (1993). "Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence." *J Biol Chem* **268**(35): 26745-51.
- Hill, M. E., S. L. Asa, et al. (1999). "Essential requirement for Pax6 in control of enteroendocrine proglucagon gene transcription." *Mol Endocrinol* **13**(9): 1474-86.
- Hogan, B. L. and P. A. Kolodziej (2002). "Organogenesis: molecular mechanisms of tubulogenesis." *Nat Rev Genet* **3**(7): 513-23.
- Holland, A. M., L. J. Gonez, et al. (2005). "Conditional expression demonstrates the role of the homeodomain transcription factor Pdx1 in maintenance and regeneration of beta-cells in the adult pancreas." *Diabetes* **54**(9): 2586-95.
- Hollnagel, A., V. Oehlmann, et al. (1999). "Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells." *J Biol Chem* **274**(28): 19838-45.
- Horabin, J. I., D. Bopp, et al. (1995). "Selection and maintenance of sexual identity in the *Drosophila* germline." *Genetics* **141**(4): 1521-35.

- Horikawa, Y., N. Iwasaki, et al. (1997). "Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY." *Nat Genet* **17**(4): 384-5.
- Hounsom, L., R. Corder, et al. (2001). "Oxidative stress participates in the breakdown of neuronal phenotype in experimental diabetic neuropathy." *Diabetologia* **44**(4): 424-8.
- Hua, H., Y. Q. Zhang, et al. (2006). "BMP4 regulates pancreatic progenitor cell expansion through Id2." *J Biol Chem* **281**(19): 13574-80.
- Huang, H. P., M. Liu, et al. (2000). "Regulation of the pancreatic islet-specific gene BETA2 (neuroD) by neurogenin 3." *Mol Cell Biol* **20**(9): 3292-307.
- Huber, W., A. von Heydebreck, et al. (2002). "Variance stabilization applied to microarray data calibration and to the quantification of differential expression." *Bioinformatics* **18 Suppl 1**: S96-104.
- Huelsken, J., R. Vogel, et al. (2001). "beta-Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin." *Cell* **105**(4): 533-45.
- Hussain, M. A., J. Lee, et al. (1997). "POU domain transcription factor brain 4 confers pancreatic alpha-cell-specific expression of the proglucagon gene through interaction with a novel proximal promoter G1 element." *Mol Cell Biol* **17**(12): 7186-94.
- Hwung, Y. P., Y. Z. Gu, et al. (1990). "Cooperativity of sequence elements mediates tissue specificity of the rat insulin II gene." *Mol Cell Biol* **10**(4): 1784-8.
- Igarashi, K., K. Kataoka, et al. (1994). "Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins." *Nature* **367**(6463): 568-72.
- Imhof, A., X. J. Yang, et al. (1997). "Acetylation of general transcription factors by histone acetyltransferases." *Curr Biol* **7**(9): 689-92.
- Itkin-Ansari, P., E. Marcora, et al. (2005). "NeuroD1 in the endocrine pancreas: localization and dual function as an activator and repressor." *Dev Dyn* **233**(3): 946-53.
- Itoh, K., K. Igarashi, et al. (1995). "Cloning and characterization of a novel erythroid cell-derived CNC family transcription factor heterodimerizing with the small Maf family proteins." *Mol Cell Biol* **15**(8): 4184-93.
- Jacquemin, P., S. M. Durviaux, et al. (2000). "Transcription factor hepatocyte nuclear factor 6 regulates pancreatic endocrine cell differentiation and controls expression of the proendocrine gene ngn3." *Mol Cell Biol* **20**(12): 4445-54.
- Jenny, M., C. Uhl, et al. (2002). "Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium." *Embo J* **21**(23): 6338-47.
- Jensen, J. (2004). "Gene regulatory factors in pancreatic development." *Dev Dyn* **229**(1): 176-200.
- Jensen, J., R. S. Heller, et al. (2000). "Independent development of pancreatic alpha- and beta-cells from neurogenin3-expressing precursors: a role for the notch pathway in repression of premature differentiation." *Diabetes* **49**(2): 163-76.
- Jensen, J., E. E. Pedersen, et al. (2000). "Control of endodermal endocrine development by Hes-1." *Nat Genet* **24**(1): 36-44.
- Jiang, F. X., E. G. Stanley, et al. (2002). "Bone morphogenetic proteins promote development of fetal pancreas epithelial colonies containing insulin-positive cells." *J Cell Sci* **115**(Pt 4): 753-60.

- Joglekar, M. V., V. S. Parekh, et al. (2007). "MicroRNA profiling of developing and regenerating pancreas reveal post-transcriptional regulation of neurogenin3." Dev Biol **311**(2): 603-12.
- Johansson, C. B., S. Momma, et al. (1999). "Identification of a neural stem cell in the adult mammalian central nervous system." Cell **96**(1): 25-34.
- Johansson, K. A., U. Dursun, et al. (2007). "Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types." Dev Cell **12**(3): 457-65.
- Johnson, A. D. (1995). "The price of repression." Cell **81**(5): 655-8.
- Johnson, J. H., A. Ogawa, et al. (1990). "Underexpression of beta cell high Km glucose transporters in noninsulin-dependent diabetes." Science **250**(4980): 546-9.
- Jonsson, J., L. Carlsson, et al. (1994). "Insulin-promoter-factor 1 is required for pancreas development in mice." Nature **371**(6498): 606-9.
- Jun, H. S. and J. W. Yoon (2003). "A new look at viruses in type 1 diabetes." Diabetes Metab Res Rev **19**(1): 8-31.
- Kaestner, K. H., J. Katz, et al. (1999). "Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression in vivo." Genes Dev **13**(4): 495-504.
- Kageyama, R. and T. Ohtsuka (1999). "The Notch-Hes pathway in mammalian neural development." Cell Res **9**(3): 179-88.
- Kahn, S. E., S. Andrikopoulos, et al. (1999). "Islet amyloid: a long-recognized but underappreciated pathological feature of type 2 diabetes." Diabetes **48**(2): 241-53.
- Kamachi, Y., M. Uchikawa, et al. (2001). "Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development." Genes Dev **15**(10): 1272-86.
- Kanno, R., T. Ogihara, et al. (2006). "Activin A-induced expression of PAX4 in AR42J-B13 cells involves the increase in transactivation of E47/E12." Biochim Biophys Acta **1759**(1-2): 44-50.
- Karlsson, O., T. Edlund, et al. (1987). "A mutational analysis of the insulin gene transcription control region: expression in beta cells is dependent on two related sequences within the enhancer." Proc Natl Acad Sci U S A **84**(24): 8819-23.
- Karlsson, O., S. Thor, et al. (1990). "Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain." Nature **344**(6269): 879-82.
- Karvonen, M., J. Tuomilehto, et al. (1993). "A review of the recent epidemiological data on the worldwide incidence of type 1 (insulin-dependent) diabetes mellitus. World Health Organization DIAMOND Project Group." Diabetologia **36**(10): 883-92.
- Karvonen, M., M. Viik-Kajander, et al. (2000). "Incidence of childhood type 1 diabetes worldwide. Diabetes Mondiale (DiaMond) Project Group." Diabetes Care **23**(10): 1516-26.
- Kataoka, K., M. Noda, et al. (1994). "Maf nuclear oncoprotein recognizes sequences related to an AP-1 site and forms heterodimers with both Fos and Jun." Mol Cell Biol **14**(1): 700-12.
- Kawaguchi, Y., B. Cooper, et al. (2002). "The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors." Nat Genet **32**(1): 128-34.

- Kerppola, T. K. and T. Curran (1994). "A conserved region adjacent to the basic domain is required for recognition of an extended DNA binding site by Maf/Nrl family proteins." Oncogene **9**(11): 3149-58.
- Kerppola, T. K. and T. Curran (1994). "Maf and Nrl can bind to AP-1 sites and form heterodimers with Fos and Jun." Oncogene **9**(3): 675-84.
- Khoo, S., S. C. Griffen, et al. (2003). "Regulation of insulin gene transcription by ERK1 and ERK2 in pancreatic beta cells." J Biol Chem **278**(35): 32969-77.
- Kim, S. K., M. Hebrok, et al. (2000). "Activin receptor patterning of foregut organogenesis." Genes Dev **14**(15): 1866-71.
- Kim, S. K., M. Hebrok, et al. (1997). "Pancreas development in the chick embryo." Cold Spring Harb Symp Quant Biol **62**: 377-83.
- King, G. L. and M. R. Loeken (2004). "Hyperglycemia-induced oxidative stress in diabetic complications." Histochem Cell Biol **122**(4): 333-8.
- Kioussi, C., P. Briata, et al. (2002). "Identification of a Wnt/Dvl/beta-Catenin --> Pitx2 pathway mediating cell-type-specific proliferation during development." Cell **111**(5): 673-85.
- Kirmizis, A. and P. J. Farnham (2004). "Genomic approaches that aid in the identification of transcription factor target genes." Exp Biol Med (Maywood) **229**(8): 705-21.
- Kishi, A., T. Nakamura, et al. (2003). "Sumoylation of Pdx1 is associated with its nuclear localization and insulin gene activation." Am J Physiol Endocrinol Metab **284**(4): E830-40.
- Knight, B. C. and S. High (1998). "Membrane integration of Sec61alpha: a core component of the endoplasmic reticulum translocation complex." Biochem J **331** (Pt 1): 161-7.
- Kojima, I. and K. Umezawa (2006). "Conophylline: a novel differentiation inducer for pancreatic beta cells." Int J Biochem Cell Biol **38**(5-6): 923-30.
- Koranyi, L., M. A. Permutt, et al. (1989). "Proinsulin I and II gene expression in inbred mouse strains." Mol Endocrinol **3**(11): 1895-902.
- Kowluru, R. A. and A. Kennedy (2001). "Therapeutic potential of anti-oxidants and diabetic retinopathy." Expert Opin Investig Drugs **10**(9): 1665-76.
- Krapp, A., M. Knofler, et al. (1996). "The p48 DNA-binding subunit of transcription factor PTF1 is a new exocrine pancreas-specific basic helix-loop-helix protein." Embo J **15**(16): 4317-29.
- Krapp, A., M. Knofler, et al. (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**(23): 3752-63.
- Krumlauf, R. (1994). "Hox genes in vertebrate development." Cell **78**(2): 191-201.
- Kuo, M. H. and C. D. Allis (1999). "In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment." Methods **19**(3): 425-33.
- Lammert, E., J. Brown, et al. (2000). "Notch gene expression during pancreatic organogenesis." Mech Dev **94**(1-2): 199-203.
- Lane, J. T. (2004). "Microalbuminuria as a marker of cardiovascular and renal risk in type 2 diabetes mellitus: a temporal perspective." Am J Physiol Renal Physiol **286**(3): F442-50.

- Lasorella, A., T. Uo, et al. (2001). "Id proteins at the cross-road of development and cancer." *Oncogene* **20**(58): 8326-33.
- Lee, C. S., D. D. De Leon, et al. (2006). "Regeneration of pancreatic islets after partial pancreatectomy in mice does not involve the reactivation of neurogenin-3." *Diabetes* **55**: 4.
- Lee, C. S., N. J. Sund, et al. (2005). "Foxa2 is required for the differentiation of pancreatic alpha-cells." *Dev Biol* **278**(2): 484-95.
- Lee, J. C., S. B. Smith, et al. (2001). "Regulation of the pancreatic pro-endocrine gene neurogenin3." *Diabetes* **50**(5): 928-36.
- Lee, J. E. (1997). "Basic helix-loop-helix genes in neural development." *Curr Opin Neurobiol* **7**(1): 13-20.
- Lee, J. E. (1997). "NeuroD and neurogenesis." *Dev Neurosci* **19**(1): 27-32.
- Lee, Y. H. and J. P. Saint-Jeannet (2003). "Sox9, a novel pancreatic marker in Xenopus." *Int J Dev Biol* **47**(6): 459-62.
- Leonard, J., B. Peers, et al. (1993). "Characterization of somatostatin transactivating factor-1, a novel homeobox factor that stimulates somatostatin expression in pancreatic islet cells." *Mol Endocrinol* **7**(10): 1275-83.
- Li, B., Q. Dai, et al. (2002). "Ovol2, a mammalian homolog of Drosophila ovo: gene structure, chromosomal mapping, and aberrant expression in blind-sterile mice." *Genomics* **80**(3): 319-25.
- Li, B., D. R. Mackay, et al. (2002). "The LEF1/beta -catenin complex activates movo1, a mouse homolog of Drosophila ovo required for epidermal appendage differentiation." *Proc Natl Acad Sci U S A* **99**(9): 6064-9.
- Li, B., M. Nair, et al. (2005). "Ovol1 regulates meiotic pachytene progression during spermatogenesis by repressing Id2 expression." *Development* **132**(6): 1463-73.
- Li, H., S. Arber, et al. (1999). "Selective agenesis of the dorsal pancreas in mice lacking homeobox gene Hlxb9." *Nat Genet* **23**(1): 67-70.
- Li, H. and H. Edlund (2001). "Persistent expression of Hlxb9 in the pancreatic epithelium impairs pancreatic development." *Dev Biol* **240**(1): 247-53.
- Lin, J. W., A. V. Biankin, et al. (2004). "Differential requirement for ptf1a in endocrine and exocrine lineages of developing zebrafish pancreas." *Dev Biol* **274**(2): 491-503.
- Lioubinski, O., M. Muller, et al. (2003). "Expression of Sox transcription factors in the developing mouse pancreas." *Dev Dyn* **227**(3): 402-8.
- Logan, C. Y. and R. Nusse (2004). "The Wnt signaling pathway in development and disease." *Annu Rev Cell Dev Biol* **20**: 781-810.
- Lu, J., J. Andrews, et al. (1998). "Drosophila OVO zinc-finger protein regulates ovo and ovarian tumor target promoters." *Dev Genes Evol* **208**(4): 213-22.
- Lu, M., J. Seufert, et al. (1997). "Pancreatic beta-cell-specific repression of insulin gene transcription by CCAAT/enhancer-binding protein beta. Inhibitory interactions with basic helix-loop-helix transcription factor E47." *J Biol Chem* **272**(45): 28349-59.
- Lynn, F. C., S. B. Smith, et al. (2007). "Sox9 coordinates a transcriptional network in pancreatic progenitor cells." *Proc Natl Acad Sci U S A* **104**(25): 10500-5.

- MacDonald, M. J., L. A. Fahien, et al. (2005). "Perspective: emerging evidence for signaling roles of mitochondrial anaplerotic products in insulin secretion." Am J Physiol Endocrinol Metab **288**(1): E1-15.
- Macfarlane, W. M., C. M. McKinnon, et al. (1999). "Glucose stimulates translocation of the homeodomain transcription factor PDX1 from the cytoplasm to the nucleus in pancreatic beta-cells." J Biol Chem **274**(2): 1011-6.
- Malecki, M. T., U. S. Jhala, et al. (1999). "Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus." Nat Genet **23**(3): 323-8.
- Mandrup-Poulsen, T. (2003). "Apoptotic signal trasduction pathways in diabetes." Biochem Pharmacol **66**: 8.
- Manova, K., V. De Leon, et al. (1995). "mRNAs for activin receptors II and IIB are expressed in mouse oocytes and in the epiblast of pregastrula and gastrula stage mouse embryos." Mech Dev **49**(1-2): 3-11.
- Mansouri, A., G. Goudreau, et al. (1999). "Pax genes and their role in organogenesis." Cancer Res **59**(7 Suppl): 1707s-1709s; discussion 1709s-1710s.
- Mansouri, A., L. St-Onge, et al. (1999). "Role of Genes in Endoderm-derived Organs." Trends Endocrinol Metab **10**(4): 164-167.
- Marini, M. G., K. Chan, et al. (1997). "hMAF, a small human transcription factor that heterodimerizes specifically with Nrf1 and Nrf2." J Biol Chem **272**(26): 16490-7.
- Marsich, E., A. Vetere, et al. (2003). "The PAX6 gene is activated by the basic helix-loop-helix transcription factor NeuroD/BETA2." Biochem J **376**(Pt 3): 707-15.
- Martin, B. C., J. H. Warram, et al. (1992). "Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study." Lancet **340**(8825): 925-9.
- Marzban, L., K. Park, et al. (2003). "Islet amyloid polypeptide and type 2 diabetes." Exp Gerontol **38**(4): 347-51.
- Mashima, H., H. Ohnishi, et al. (1996). "Betacellulin and activin A coordinately convert amylase-secreting pancreatic AR42J cells into insulin-secreting cells." J Clin Invest **97**(7): 1647-54.
- Mashima, H., H. Shibata, et al. (1996). "Formation of insulin-producing cells from pancreatic acinar AR42J cells by hepatocyte growth factor." Endocrinology **137**(9): 3969-76.
- Massague, J. and Y. G. Chen (2000). "Controlling TGF-beta signaling." Genes Dev **14**(6): 627-44.
- Masu, Y., S. Ikeda, et al. (1998). "Expression of murine novel zinc finger proteins highly homologous to Drosophila ovo gene product in testis." FEBS Lett **421**(3): 224-8.
- Matschinsky, F. M., B. Glaser, et al. (1998). "Pancreatic beta-cell glucokinase: closing the gap between theoretical concepts and experimental realities." Diabetes **47**(3): 307-15.
- McCormick, P. J., Y. Miao, et al. (2003). "Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins." Mol Cell **12**(2): 329-41.
- McMahon, A. P. (2000). "More surprises in the Hedgehog signaling pathway." Cell **100**(2): 185-8.
- Meglasson, M. D. and F. M. Matschinsky (1986). "Pancreatic islet glucose metabolism and regulation of insulin secretion." Diabetes Metab Rev **2**(3-4): 163-214.

- Mellitzer, G., S. Bonne, et al. (2006). "IA1 is NGN3-dependent and essential for differentiation of the endocrine pancreas." Embo J **25**(6): 1344-52.
- Melloul, D., Y. Ben-Neriah, et al. (1993). "Glucose modulates the binding of an islet-specific factor to a conserved sequence within the rat I and the human insulin promoters." Proc Natl Acad Sci U S A **90**(9): 3865-9.
- Merezak, S., A. A. Hardikar, et al. (2001). "Intrauterine low protein diet increases fetal beta-cell sensitivity to NO and IL-1 beta: the protective role of taurine." J Endocrinol **171**(2): 299-308.
- Merrill, B. J., U. Gat, et al. (2001). "Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin." Genes Dev **15**(13): 1688-705.
- Mével-Ninio, M., E. Fouilloux, et al. (1996). "The three dominant female-sterile mutations of the *Drosophila* ovo gene are point mutations that create new translation-initiator AUG codons." Development **122**(12): 4131-8.
- Mével-Ninio, M., R. Terracol, et al. (1991). "The ovo gene of *Drosophila* encodes a zinc finger protein required for female germ line development." Embo J **10**(8): 2259-66.
- Mével-Ninio, M., R. Terracol, et al. (1995). "ovo, a *Drosophila* gene required for ovarian development, is specifically expressed in the germline and shares most of its coding sequences with shavenbaby, a gene involved in embryo patterning." Mech Dev **49**(1-2): 83-95.
- Miller, C. P., R. E. McGehee, Jr., et al. (1994). "IDX-1: a new homeodomain transcription factor expressed in rat pancreatic islets and duodenum that transactivates the somatostatin gene." Embo J **13**(5): 1145-56.
- Miralles, F., T. Battelino, et al. (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**(3): 827-36.
- Miralles, F., P. Czernichow, et al. (1999). "Signaling through fibroblast growth factor receptor 2b plays a key role in the development of the exocrine pancreas." Proc Natl Acad Sci U S A **96**(11): 6267-72.
- Miralles, F., P. Czernichow, et al. (1998). "Follistatin regulates the relative proportions of endocrine versus exocrine tissue during pancreatic development." Development **125**(6): 1017-24.
- Miralles, F., P. Czernichow, et al. (1999). "Pancreatic acinar AR42J cells express functional nerve growth factor receptors." J Endocrinol **160**(3): 433-42.
- Miralles, F., L. Lamotte, et al. (2006). "Interplay between FGF10 and Notch signalling is required for the self-renewal of pancreatic progenitors." Int J Dev Biol **50**(1): 17-26.
- Monaghan, A. P., K. H. Kaestner, et al. (1993). "Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm." Development **119**(3): 567-78.
- Mosley, A. L. and S. Ozcan (2004). "The pancreatic duodenal homeobox-1 protein (Pdx-1) interacts with histone deacetylases Hdac-1 and Hdac-2 on low levels of glucose." J Biol Chem **279**(52): 54241-7.

- Moustakas, A. K. and G. K. Papadopoulos (2002). "Molecular properties of HLA-DQ alleles conferring susceptibility to or protection from insulin-dependent diabetes mellitus: keys to the fate of islet beta-cells." *Am J Med Genet* **115**(1): 37-47.
- Muller, H., A. P. Bracken, et al. (2001). "E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis." *Genes Dev* **15**(3): 267-85.
- Murtaugh, L. C., A. C. Law, et al. (2005). "Beta-catenin is essential for pancreatic acinar but not islet development." *Development* **132**(21): 4663-74.
- Nair, M., V. Bilanchone, et al. (2007). "Ov011 represses its own transcription by competing with transcription activator c-Myb and by recruiting histone deacetylase activity." *Nucleic Acids Res* **35**(5): 1687-97.
- Nair, M., A. Teng, et al. (2006). "Ov011 regulates the growth arrest of embryonic epidermal progenitor cells and represses c-myc transcription." *J Cell Biol* **173**(2): 253-64.
- Nakagoshi, H., C. Kanei-Ishii, et al. (1992). "Transcriptional activation of the c-myc gene by the c-myb and B-myb gene products." *Oncogene* **7**(6): 1233-40.
- Nakamura, T., K. Takio, et al. (1990). "Activin-binding protein from rat ovary is follistatin." *Science* **247**(4944): 836-8.
- Narendran, P., E. Estella, et al. (2005). "Immunology of type 1 diabetes." *Qjm* **98**(8): 547-56.
- Naya, F. J., H. P. Huang, et al. (1997). "Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice." *Genes Dev* **11**(18): 2323-34.
- Naya, F. J., C. M. Stellrecht, et al. (1995). "Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor." *Genes Dev* **9**(8): 1009-19.
- Nerup, J., T. Mandrup-Poulsen, et al. (1988). "Mechanisms of pancreatic beta-cell destruction in type I diabetes." *Diabetes Care* **11 Suppl 1**: 16-23.
- Nishimura, W., T. Salameh, et al. (2005). "Regulation of insulin gene expression by overlapping DNA-binding elements." *Biochem J* **392**(Pt 1): 181-9.
- Nonogaki, K. (2000). "New insights into sympathetic regulation of glucose and fat metabolism." *Diabetologia* **43**(5): 533-49.
- Norgaard, G. A., J. N. Jensen, et al. (2003). "FGF10 signaling maintains the pancreatic progenitor cell state revealing a novel role of Notch in organ development." *Dev Biol* **264**(2): 323-38.
- Norton, J. D. (2000). "ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis." *J Cell Sci* **113 (Pt 22)**: 3897-905.
- Nygaard Jensen, J. and J. Jensen (2004). "Cell therapy of diabetes." *Adv Exp Med Biol* **552**: 16-38.
- O'Driscoll, L., P. Gammell, et al. (2004). "Expression in murine teratocarcinoma F9 cells of transcription factors involved in pancreas development." *Transplant Proc* **36**(4): 1151-8.
- Odagiri, H., J. Wang, et al. (1996). "Function of the human insulin promoter in primary cultured islet cells." *J Biol Chem* **271**(4): 1909-15.
- Offield, M. F., T. L. Jetton, et al. (1996). "PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum." *Development* **122**(3): 983-95.

- Ogawa, K., K. Abe, et al. (1993). "Expression of alpha, beta A and beta B subunits of inhibin or activin and follistatin in rat pancreatic islets." FEBS Lett **319**(3): 217-20.
- Ogihara, T., H. Watada, et al. (2003). "p38 MAPK is involved in activin A- and hepatocyte growth factor-mediated expression of pro-endocrine gene neurogenin 3 in AR42J-B13 cells." J Biol Chem **278**(24): 21693-700.
- Ohlsson, H., K. Karlsson, et al. (1993). "IPF1, a homeodomain-containing transactivator of the insulin gene." Embo J **12**(11): 4251-9.
- Ohneda, K., H. Ee, et al. (2000). "Regulation of insulin gene transcription." Semin Cell Dev Biol **11**(4): 227-33.
- Ohneda, K., R. G. Mirmira, et al. (2000). "The homeodomain of PDX-1 mediates multiple protein-protein interactions in the formation of a transcriptional activation complex on the insulin promoter." Mol Cell Biol **20**(3): 900-11.
- Ohnishi, H., N. Ohgushi, et al. (1995). "Conversion of amylase-secreting rat pancreatic AR42J cells to neuronlike cells by activin A." J Clin Invest **95**(5): 2304-14.
- Olbrot, M., J. Rud, et al. (2002). "Identification of beta-cell-specific insulin gene transcription factor RIPE3b1 as mammalian MafA." Proc Natl Acad Sci U S A **99**(10): 6737-42.
- Oliver, B., D. Pauli, et al. (1990). "Genetic evidence that the ovo locus is involved in Drosophila germ line sex determination." Genetics **125**(3): 535-50.
- Oliver, B., J. Singer, et al. (1994). "Function of Drosophila ovo+ in germ-line sex determination depends on X-chromosome number." Development **120**(11): 3185-95.
- Oliver, N. A. (1987). "Altered production of fibronectin and collagen in hypercortisolism may inhibit tissue repair." Arch Dermatol **123**(5): 570-1.
- Onodera, K., J. A. Shavit, et al. (1999). "Characterization of the murine mafF gene." J Biol Chem **274**(30): 21162-9.
- Orlando, V. (2000). "Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation." Trends Biochem Sci **25**(3): 99-104.
- Orr-Urtreger, A., M. T. Bedford, et al. (1993). "Developmental localization of the splicing alternatives of fibroblast growth factor receptor-2 (FGFR2)." Dev Biol **158**(2): 475-86.
- Oster, A., J. Jensen, et al. (1998). "Rat endocrine pancreatic development in relation to two homeobox gene products (Pdx-1 and Nkx 6.1)." J Histochem Cytochem **46**(6): 707-15.
- Owerbach, D., G. I. Bell, et al. (1980). "The insulin gene is located on chromosome 11 in humans." Nature **286**(5768): 82-4.
- Oyake, T., K. Itoh, et al. (1996). "Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site." Mol Cell Biol **16**(11): 6083-95.
- Palgi, J., E. Stumpf, et al. (2000). "Transcription factor expression and hormone production in pancreatic AR42J cells." Mol Cell Endocrinol **165**(1-2): 41-9.
- Papadopoulou, S. and H. Edlund (2005). "Attenuated Wnt signaling perturbs pancreatic growth but not pancreatic function." Diabetes **54**(10): 2844-51.

- Pedersen, A. H. and R. S. Heller (2005). "A possible role for the canonical Wnt pathway in endocrine cell development in chicks." Biochem Biophys Res Commun **333**(3): 961-8.
- Peers, B., J. Leonard, et al. (1994). "Insulin expression in pancreatic islet cells relies on cooperative interactions between the helix loop helix factor E47 and the homeobox factor STF-1." Mol Endocrinol **8**(12): 1798-806.
- Petersen, H. V., P. Serup, et al. (1994). "Transcriptional regulation of the human insulin gene is dependent on the homeodomain protein STF1/IPF1 acting through the CT boxes." Proc Natl Acad Sci U S A **91**(22): 10465-9.
- Petrik, J., B. Reusens, et al. (1999). "A low protein diet alters the balance of islet cell replication and apoptosis in the fetal and neonatal rat and is associated with a reduced pancreatic expression of insulin-like growth factor-II." Endocrinology **140**(10): 4861-73.
- Pictet, R. L., W. R. Clark, et al. (1972). "An ultrastructural analysis of the developing embryonic pancreas." Dev Biol **29**(4): 436-67.
- Piper, K., S. G. Ball, et al. (2002). "Novel SOX9 expression during human pancreas development correlates to abnormalities in Campomelic dysplasia." Mech Dev **116**(1-2): 223-6.
- Pociot, F. and M. F. McDermott (2002). "Genetics of type 1 diabetes mellitus." Genes Immun **3**(5): 235-49.
- Poitout, V., R. Stein, et al. (2004). "Insulin gene expression and biosynthesis." In: International textbook of diabetes mellitus. 3rd edition DeFronzo RA, Ferrarini, E, Keen H, Zimmet P, eds Chichester: John Wiley & Sons: 26.
- Polak, M., L. Bouchareb-Banaei, et al. (2000). "Early pattern of differentiation in the human pancreas." Diabetes **49**(2): 225-32.
- Porte, D., Jr. (2001). "Clinical importance of insulin secretion and its interaction with insulin resistance in the treatment of type 2 diabetes mellitus and its complications." Diabetes Metab Res Rev **17**(3): 181-8.
- Porte, D., Jr. and M. W. Schwartz (1996). "Diabetes complications: why is glucose potentially toxic?" Science **272**(5262): 699-700.
- Prado, C. L., A. E. Pugh-Bernard, et al. (2004). "Ghrelin cells replace insulin-producing beta cells in two mouse models of pancreas development." Proc Natl Acad Sci U S A **101**(9): 2924-9.
- Prendergast, G. C., D. Lawe, et al. (1991). "Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation." Cell **65**(3): 395-407.
- Qiu, Y., M. Guo, et al. (2002). "Insulin gene transcription is mediated by interactions between the p300 coactivator and PDX-1, BETA2, and E47." Mol Cell Biol **22**(2): 412-20.
- Rafiq, I., H. J. Kennedy, et al. (1998). "Glucose-dependent translocation of insulin promoter factor-1 (IPF-1) between the nuclear periphery and the nucleoplasm of single MIN6 beta-cells." J Biol Chem **273**(36): 23241-7.
- Ramalho-Santos, M., D. A. Melton, et al. (2000). "Hedgehog signals regulate multiple aspects of gastrointestinal development." Development **127**(12): 2763-72.

- Rausa, F., U. Samadani, et al. (1997). "The cut-homeodomain transcriptional activator HNF-6 is coexpressed with its target gene HNF-3 beta in the developing murine liver and pancreas." Dev Biol **192**(2): 228-46.
- Reaven, G. M. (1988). "Banting lecture 1988. Role of insulin resistance in human disease." Diabetes **37**(12): 1595-607.
- Ritvos, O., T. Tuuri, et al. (1995). "Activin disrupts epithelial branching morphogenesis in developing glandular organs of the mouse." Mech Dev **50**(2-3): 229-45.
- Roberts, D. J., R. L. Johnson, et al. (1995). "Sonic hedgehog is an endodermal signal inducing Bmp-4 and Hox genes during induction and regionalization of the chick hindgut." Development **121**(10): 3163-74.
- Roberts, D. J., D. M. Smith, et al. (1998). "Epithelial-mesenchymal signaling during the regionalization of the chick gut." Development **125**(15): 2791-801.
- Rossetti, L., A. Giaccari, et al. (1990). "Glucose toxicity." Diabetes Care **13**(6): 610-30.
- Rudnick, A., T. Y. Ling, et al. (1994). "Pancreatic beta cells express a diverse set of homeobox genes." Proc Natl Acad Sci U S A **91**(25): 12203-7.
- Rulifson, I. C., S. K. Karnik, et al. (2007). "Wnt signaling regulates pancreatic beta cell proliferation." Proc Natl Acad Sci U S A **104**(15): 6247-52.
- Ruppert, C., R. Kroschewski, et al. (1993). "Identification, characterization and cloning of myr 1, a mammalian myosin-I." J Cell Biol **120**(6): 1393-403.
- Rutter, W. J., R. L. Pictet, et al. (1978). "An analysis of pancreatic development: role of mesenchymal factor and other extracellular factors." Symp Soc Dev Biol(35): 205-27.
- Sander, M., S. C. Griffen, et al. (1998). "A novel glucose-responsive element in the human insulin gene functions uniquely in primary cultured islets." Proc Natl Acad Sci U S A **95**(20): 11572-7.
- Sander, M., L. Sussel, et al. (2000). "Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas." Development **127**(24): 5533-40.
- Santerre, R. F., R. A. Cook, et al. (1981). "Insulin synthesis in a clonal cell line of simian virus 40-transformed hamster pancreatic beta cells." Proc Natl Acad Sci U S A **78**(7): 4339-43.
- Sanvito, F., P. L. Herrera, et al. (1994). "TGF-beta 1 influences the relative development of the exocrine and endocrine pancreas in vitro." Development **120**(12): 3451-62.
- Sasaki, H. and B. L. Hogan (1993). "Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo." Development **118**(1): 47-59.
- Scharfmann, R. (1997). "Neurotrophin and neurotrophin receptors in islet cells." Horm Metab Res **29**(6): 294-6.
- Scharfmann, R. (2000). "Control of early development of the pancreas in rodents and humans: implications of signals from the mesenchyme." Diabetologia **43**(9): 1083-92.
- Schena, M., D. Shalon, et al. (1995). "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science **270**(5235): 467-70.
- Schepers, G. E., R. D. Teasdale, et al. (2002). "Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families." Dev Cell **3**(2): 167-70.

- Schilham, M. W., M. A. Oosterwegel, et al. (1996). "Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4." Nature **380**(6576): 711-4.
- Schmid, P., W. A. Schulz, et al. (1989). "Dynamic expression pattern of the myc protooncogene in midgestation mouse embryos." Science **243**(4888): 226-9.
- Schroeter, E. H., J. A. Kisslinger, et al. (1998). "Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain." Nature **393**(6683): 382-6.
- Schuit, F. C. (1997). "Is GLUT2 required for glucose sensing?" Diabetologia **40**(1): 104-11.
- Schwitzgebel, V. M. (2001). "Programming of the pancreas." Mol Cell Endocrinol **185**(1-2): 99-108.
- Schwitzgebel, V. M., D. W. Scheel, et al. (2000). "Expression of neurogenin3 reveals an islet cell precursor population in the pancreas." Development **127**(16): 3533-42.
- Sekito, A., S. Koide-Yoshida, et al. (2006). "DJ-1 interacts with HIPK1 and affects H₂O₂-induced cell death." Free Radic Res **40**(2): 155-65.
- Serup, P., H. V. Petersen, et al. (1995). "The homeodomain protein IPF-1/STF-1 is expressed in a subset of islet cells and promotes rat insulin 1 gene expression dependent on an intact E1 helix-loop-helix factor binding site." Biochem J **310** (Pt 3): 997-1003.
- Seufert, J., G. C. Weir, et al. (1998). "Differential expression of the insulin gene transcriptional repressor CCAAT/enhancer-binding protein beta and transactivator islet duodenum homeobox-1 in rat pancreatic beta cells during the development of diabetes mellitus." J Clin Invest **101**(11): 2528-39.
- Seymour, P. A., K. K. Freude, et al. (2007). "SOX9 is required for maintenance of the pancreatic progenitor cell pool." Proc Natl Acad Sci U S A **104**(6): 1865-70.
- Shapiro, A. M., J. R. Lakey, et al. (2000). "Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen." N Engl J Med **343**(4): 230-8.
- Sharma, A., D. Fusco-DeMane, et al. (1995). "The role of the insulin control element and RIPE3b1 activators in glucose-stimulated transcription of the insulin gene." Mol Endocrinol **9**(11): 1468-76.
- Sharma, A. and R. Stein (1994). "Glucose-induced transcription of the insulin gene is mediated by factors required for beta-cell-type-specific expression." Mol Cell Biol **14**(2): 871-9.
- Sharpe, P. C., W. H. Liu, et al. (1998). "Glucose-induced oxidative stress in vascular contractile cells: comparison of aortic smooth muscle cells and retinal pericytes." Diabetes **47**(5): 801-9.
- Shen, C. N., M. E. Horb, et al. (2003). "Transdifferentiation of pancreas to liver." Mech Dev **120**(1): 107-16.
- Shen, C. N., J. M. Slack, et al. (2000). "Molecular basis of transdifferentiation of pancreas to liver." Nat Cell Biol **2**(12): 879-87.
- Shieh, S. Y. and M. J. Tsai (1991). "Cell-specific and ubiquitous factors are responsible for the enhancer activity of the rat insulin II gene." J Biol Chem **266**(25): 16708-14.

- Shih, D. Q. and M. Stoffel (2001). "Dissecting the transcriptional network of pancreatic islets during development and differentiation." Proc Natl Acad Sci U S A **98**(25): 14189-91.
- Shiozaki, S., T. Tajima, et al. (1999). "Impaired differentiation of endocrine and exocrine cells of the pancreas in transgenic mouse expressing the truncated type II activin receptor." Biochim Biophys Acta **1450**(1): 1-11.
- Shtutman, M., J. Zhurinsky, et al. (1999). "The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway." Proc Natl Acad Sci U S A **96**(10): 5522-7.
- Slack, J. M. (1995). "Developmental biology of the pancreas." Development **121**(6): 1569-80.
- Smith, S. B., R. Gasa, et al. (2003). "Neurogenin3 and hepatic nuclear factor 1 cooperate in activating pancreatic expression of Pax4." J Biol Chem **278**(40): 38254-9.
- Smith, S. B., H. Watada, et al. (2004). "Neurogenin3 activates the islet differentiation program while repressing its own expression." Mol Endocrinol **18**(1): 142-9.
- Snoeck, A., C. Remacle, et al. (1990). "Effect of a low protein diet during pregnancy on the fetal rat endocrine pancreas." Biol Neonate **57**(2): 107-18.
- Snyder, S. E., B. Peng, et al. (2003). "Expression of VGF mRNA in developing neuroendocrine and endocrine tissues." J Endocrinol **179**(2): 227-35.
- Soares, M. B., D. N. Ishii, et al. (1985). "Developmental and tissue-specific expression of a family of transcripts related to rat insulin-like growth factor II mRNA." Nucleic Acids Res **13**(4): 1119-34.
- Sobngwi, E., P. Boudou, et al. (2003). "Effect of a diabetic environment in utero on predisposition to type 2 diabetes." Lancet **361**(9372): 1861-5.
- Song, Y. D., E. J. Lee, et al. (2007). "Islet cell differentiation in liver by combinatorial expression of transcription factors neurogenin-3, BETA2, and RIPE3b1." Biochem Biophys Res Commun **354**(2): 334-9.
- Soria, B. (2001). "In-vitro differentiation of pancreatic beta-cells." Differentiation **68**(4-5): 205-19.
- Sosa-Pineda, B. (2004). "The gene Pax4 is an essential regulator of pancreatic beta-cell development." Mol Cells **18**(3): 289-94.
- Sosa-Pineda, B., K. Chowdhury, et al. (1997). "The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas." Nature **386**(6623): 399-402.
- Spooner, B. S., B. T. Walther, et al. (1970). "The development of the dorsal and ventral mammalian pancreas in vivo and in vitro." J Cell Biol **47**(1): 235-46.
- St-Onge, L., B. Sosa-Pineda, et al. (1997). "Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas." Nature **387**(6631): 406-9.
- Steiner, D. F., S. J. Chan, et al. (1985). "Structure and evolution of the insulin gene." Annu Rev Genet **19**: 463-84.
- Stern, C. D., R. T. Yu, et al. (1995). "Activin and its receptors during gastrulation and the later phases of mesoderm development in the chick embryo." Dev Biol **172**(1): 192-205.
- Stoffers, D. A., J. Ferrer, et al. (1997). "Early-onset type-II diabetes mellitus (MODY4) linked to IPF1." Nat Genet **17**(2): 138-9.

- Stoffers, D. A., N. T. Zinkin, et al. (1997). "Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence." Nat Genet **15**(1): 106-10.
- Straub, S. G., G. Shanmugam, et al. (2004). "Stimulation of insulin release by glucose is associated with an increase in the number of docked granules in the beta-cells of rat pancreatic islets." Diabetes **53**(12): 3179-83.
- Strickland, S. and V. Mahdavi (1978). "The induction of differentiation in teratocarcinoma stem cells by retinoic acid." Cell **15**(2): 393-403.
- Sumazaki, R., N. Shiojiri, et al. (2004). "Conversion of biliary system to pancreatic tissue in Hes1-deficient mice." Nat Genet **36**(1): 83-7.
- Sumiya, M. and T. Mizuno (1987). "[Immunohistological study of the expression of glucagon in dorsal pancreatic endoderm in the early chick embryo]." C R Seances Soc Biol Fil **181**(6): 718-21.
- Sund, N. J., M. Z. Vatamaniuk, et al. (2001). "Tissue-specific deletion of Foxa2 in pancreatic beta cells results in hyperinsulinemic hypoglycemia." Genes Dev **15**(13): 1706-15.
- Sussel, L., J. Kalamaras, et al. (1998). "Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells." Development **125**(12): 2213-21.
- Tal, M., Y. J. Wu, et al. (1992). "[Val12] HRAS downregulates GLUT2 in beta cells of transgenic mice without affecting glucose homeostasis." Proc Natl Acad Sci U S A **89**(13): 5744-8.
- Tam, P. P. and R. S. Beddington (1992). "Establishment and organization of germ layers in the gastrulating mouse embryo." Ciba Found Symp **165**: 27-41; discussion 42-9.
- Team, R. D. C. (2007). "R a language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Thiel, G., M. Lietz, et al. (2004). "How mammalian transcriptional repressors work." Eur J Biochem **271**(14): 2855-62.
- Thomas, M. K., K. M. Yao, et al. (1999). "Bridge-1, a novel PDZ-domain coactivator of E2A-mediated regulation of insulin gene transcription." Mol Cell Biol **19**(12): 8492-504.
- Tirone, T. A. and F. C. Brunnicardi (2001). "Overview of glucose regulation." World J Surg **25**(4): 461-7.
- Tokuyama, Y., J. Sturis, et al. (1995). "Evolution of beta-cell dysfunction in the male Zucker diabetic fatty rat." Diabetes **44**(12): 1447-57.
- Tomonari, A., K. Yoshimoto, et al. (1996). "GGAAAT motifs play a major role in transcriptional activity of the human insulin gene in a pancreatic islet beta-cell line MIN6." Diabetologia **39**(12): 1462-8.
- Towle, H. C. (2005). "Glucose as a regulator of eukaryotic gene transcription." Trends Endocrinol Metab **16**(10): 489-94.
- Treff, N. R., R. K. Vincent, et al. (2006). "Differentiation of embryonic stem cells conditionally expressing neurogenin 3." Stem Cells **24**(11): 2529-37.
- Treier, M., A. S. Gleiberman, et al. (1998). "Multistep signaling requirements for pituitary organogenesis in vivo." Genes Dev **12**(11): 1691-704.

- Tremblay, K. D., P. A. Hoodless, et al. (2000). "Formation of the definitive endoderm in mouse is a Smad2-dependent process." *Development* **127**(14): 3079-90.
- Tseng, G. C., M. K. Oh, et al. (2001). "Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects." *Nucleic Acids Res* **29**(12): 2549-57.
- Tsuchiya, M., S. Taniguchi, et al. (2006). "Potential roles of large maf's in cell lineages and developing pancreas." *Pancreas* **32**(4): 408-16.
- Tsuda, H., K. Takebayashi, et al. (1998). "Structure and promoter analysis of Math3 gene, a mouse homolog of Drosophila proneural gene atonal. Neural-specific expression by dual promoter elements." *J Biol Chem* **273**(11): 6327-33.
- Tupler, R., G. Perini, et al. (2001). "Expressing the human genome." *Nature* **409**(6822): 832-3.
- Turque, N., S. Plaza, et al. (1994). "Pax-QNR/Pax-6, a paired box- and homeobox-containing gene expressed in neurons, is also expressed in pancreatic endocrine cells." *Mol Endocrinol* **8**(7): 929-38.
- Umezawa, K., A. Hiroki, et al. (2003). "Induction of insulin production in rat pancreatic acinar carcinoma cells by conophylline." *Biomed Pharmacother* **57**(8): 341-50.
- Unezaki, S., M. Nishizawa, et al. (2004). "Characterization of the isoforms of MOVO zinc finger protein, a mouse homologue of Drosophila Ovo, as transcription factors." *Gene* **336**(1): 47-58.
- Unger, R. H. (1991). "Diabetic hyperglycemia: link to impaired glucose transport in pancreatic beta cells." *Science* **251**(4998): 1200-5.
- Unger, R. H. (1995). "Lipotoxicity in the pathogenesis of obesity-dependent NIDDM. Genetic and clinical implications." *Diabetes* **44**(8): 863-70.
- van Genderen, C., R. M. Okamura, et al. (1994). "Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice." *Genes Dev* **8**(22): 2691-703.
- Verdin, E., F. Dequiedt, et al. (2003). "Class II histone deacetylases: versatile regulators." *Trends Genet* **19**(5): 286-93.
- Verschuere, K., N. Dewulf, et al. (1995). "Expression of type I and type IB receptors for activin in midgestation mouse embryos suggests distinct functions in organogenesis." *Mech Dev* **52**(1): 109-23.
- Vetere, A., E. Marsich, et al. (2003). "Neurogenin3 triggers beta-cell differentiation of retinoic acid-derived endoderm cells." *Biochem J* **371**(Pt 3): 831-41.
- Voigt, S., B. Jungnickel, et al. (1996). "Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane." *J Cell Biol* **134**(1): 25-35.
- Wada, R. and S. Yagihashi (2005). "Role of advanced glycation end products and their receptors in development of diabetic neuropathy." *Ann N Y Acad Sci* **1043**: 598-604.
- Waeber, G., N. Thompson, et al. (1996). "Transcriptional activation of the GLUT2 gene by the IPF-1/STF-1/IDX-1 homeobox factor." *Mol Endocrinol* **10**(11): 1327-34.
- Walker, M. D., T. Edlund, et al. (1983). "Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes." *Nature* **306**(5943): 557-61.

- Wang, A. Y., A. Ehrhardt, et al. (2007). "Adenovirus transduction is required for the correction of diabetes using Pdx-1 or Neurogenin-3 in the liver." Mol Ther **15**(2): 255-63.
- Wang, H., T. Brun, et al. (2007). "MAFA controls genes implicated in insulin biosynthesis and secretion." Diabetologia **50**(2): 348-58.
- Wang, J., L. Elghazi, et al. (2004). "The concerted activities of Pax4 and Nkx2.2 are essential to initiate pancreatic beta-cell differentiation." Dev Biol **266**(1): 178-89.
- Wang, J., L. Shen, et al. (1997). "Regulation of insulin preRNA splicing by glucose." Proc Natl Acad Sci U S A **94**(9): 4360-5.
- Wang, J., G. Webb, et al. (2003). "Contrasting patterns of expression of transcription factors in pancreatic alpha and beta cells." Proc Natl Acad Sci U S A **100**(22): 12660-5.
- Watada, H., Y. Kajimoto, et al. (1996). "Involvement of the homeodomain-containing transcription factor PDX-1 in islet amyloid polypeptide gene transcription." Biochem Biophys Res Commun **229**(3): 746-51.
- Watada, H., D. W. Scheel, et al. (2003). "Distinct gene expression programs function in progenitor and mature islet cells." J Biol Chem **278**(19): 17130-40.
- Wegner, M. (1999). "From head to toes: the multiple facets of Sox proteins." Nucleic Acids Res **27**(6): 1409-20.
- Weiler, S., J. M. Gruschus, et al. (1998). "Site-directed mutations in the vnd/NK-2 homeodomain. Basis of variations in structure and sequence-specific DNA binding." J Biol Chem **273**(18): 10994-1000.
- Weinstein, D. C., A. Ruiz i Altaba, et al. (1994). "The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo." Cell **78**(4): 575-88.
- Weir, G. C., A. Sharma, et al. (1997). "Transcription factor abnormalities as a cause of beta cell dysfunction in diabetes: a hypothesis." Acta Diabetol **34**(3): 177-84.
- Wells, J. M. and D. A. Melton (1999). "Vertebrate endoderm development." Annu Rev Cell Dev Biol **15**: 393-410.
- Wells, J. M. and D. A. Melton (2000). "Early mouse endoderm is patterned by soluble factors from adjacent germ layers." Development **127**(8): 1563-72.
- Wessells, N. K. and J. H. Cohen (1967). "Early pancreas organogenesis, morphogenesis, tissue interactions, and mass effects." Dev Biol **15**: 33.
- Whelan, J., D. Poon, et al. (1989). "Pancreatic beta-cell-type-specific expression of the rat insulin II gene is controlled by positive and negative cellular transcriptional elements." Mol Cell Biol **9**(8): 3253-9.
- WHO (1999). "Definitio, diagnosis and classification of diabetes mellitus and its complications." WHO, Report of a WHO Consultation Part1: 1-5.
- Winter, J., P. Klappa, et al. (2002). "Catalytic activity and chaperone function of human protein-disulfide isomerase are required for the efficient refolding of proinsulin." J Biol Chem **277**(1): 310-7.
- Wunderle, V. M., R. Critcher, et al. (1998). "Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia." Proc Natl Acad Sci U S A **95**(18): 10649-54.

- Yamagata, K., H. Furuta, et al. (1996). "Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1)." *Nature* **384**(6608): 458-60.
- Yamaoka, T., C. Idehara, et al. (1998). "Hypoplasia of pancreatic islets in transgenic mice expressing activin receptor mutants." *J Clin Invest* **102**(2): 294-301.
- Yanaka, N., Y. Nogusa, et al. (2007). "Involvement of membrane protein GDE2 in retinoic acid-induced neurite formation in Neuro2A cells." *FEBS Lett* **581**(4): 712-8.
- Yang, I. V., E. Chen, et al. (2002). "Within the fold: assessing differential expression measures and reproducibility in microarray assays." *Genome Biol* **3**(11): research0062.
- Yang, Y. H., S. Dudoit, et al. (2002). "Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation." *Nucleic Acids Res* **30**(4): e15.
- Yasuda, H., S. Tanaka, et al. (1994). "Activin A: negative regulator of amylase secretion and cell proliferation in rat pancreatic acinar AR42J cells." *Am J Physiol* **267**(2 Pt 1): G220-6.
- Zhang, Y. Q., H. Mashima, et al. (2001). "Changes in the expression of transcription factors in pancreatic AR42J cells during differentiation into insulin-producing cells." *Diabetes* **50 Suppl 1**: S10-4.
- Zhao, L., M. Guo, et al. (2005). "The islet beta cell-enriched MafA activator is a key regulator of insulin gene transcription." *J Biol Chem* **280**(12): 11887-94.
- Zhou, P., C. Byrne, et al. (1995). "Lymphoid enhancer factor 1 directs hair follicle patterning and epithelial cell fate." *Genes Dev* **9**(6): 700-13.
- Zhu, M., M. B. Breslin, et al. (2002). "Expression of a novel zinc-finger cDNA, IA-1, is associated with rat AR42J cells differentiation into insulin-positive cells." *Pancreas* **24**(2): 139-45.
- Zimmet, P. (2000). "Globalization, coca-colonization and the chronic disease epidemic: can the Doomsday scenario be averted?" *J Intern Med* **247**(3): 301-10.
- Zimmet, P., K. G. Alberti, et al. (2001). "Global and societal implications of the diabetes epidemic." *Nature* **414**(6865): 782-7.

ACKNOWLEDGEMENTS

It's a pleasure to thank the many people who helped me during my PhD period.

It is difficult to overstate my gratitude to my friend and PhD tutor Dr. Amedeo Vetere who first brought me into the world of research. Before and throughout my thesis period and thesis-writing period he has been a mentor and overall a friend providing me with encouragement, advices, good teaching, good company and lots of ideas. He shared with me a lot of his expertise, research insights, good time and even his house. He has been a cornerstone in my professional development and I would definitely be lost without him!

I'm deeply grateful to Dr. Arun Sharma to host me in his research group, without knowing my skills and experience, and to involve me in different interesting research projects. I learned a lot from his experience on pancreas development and criticism about the scientific approach and results discussion.

I would like to thank Dr. Susan Bonner-Weir that, as Dr. Sharma, hosted me in her research group and gave me the opportunity to share the stimulating scientific environment of her research group. Thanks to all these people for helping me not only during my Joslin period but also with my future professional career.

I would like to thank all my friends from Joslin Diabetes Center: Michael (Michele), Dohoon, Elena, Sean, Jeff, Julie, Jared, Kirstine, Cindy, Wats, Esther, Cristina, Masaki, Ilham, Luisa, Lorella, Jennifer for sharing with me their experience and for the very good time that we had together.

I would like to thank Dr. Claude Maina and Dr. Christopher Taron to gave me the opportunity to work at the New England Biolabs. Thanks to my "NEB friends" George, Jacopo, Dimitris, Katherin, Cathy, Ana, Mac, Chris, Brett, Paul, Pierre, Shamik.

Thanks to my Italian lab mates: Monica, Sabrina, Amelia, Luisa, Giancarlo.

Finally I would like to deeply thank two special friends for all their help and support: Angela Koh and Priscilla Hong.

Last but, definitely not least, a special thank to Tatiana for everything....I will never forget all these people!