PhD degree in Molecular Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Settore disciplinare: MED/04

FUNCTIONAL DISSECTION OF THE CONTRIBUTION OF BAZ1B TO NEURAL CREST DYSREGULATION IN NEURODEVELOPMENTAL DISORDERS CAUSED BY SYMMETRICAL CNVs AT 7q11.23

Matteo Zanella

European Institute of Oncology, Milan

Matricola n. R10748

Supervisor: Prof. Giuseppe Testa

European Institute of Oncology, Milan

University of Milan, Milan

Anno accademico 2016-2017

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

αSMA	α Smooth Muscle Actin
7dup	7q11.23 duplication syndrome
Acf1	ATP-utilizing Chromatin assembly and remodeling Factor 1
ADHD	Attention-Deficit/Hyperactivity Disorder
ALP	Alkaline Phosphatase
ASD	Autism Spectrum Disorders
BAV	Bicuspid Aortic Valve
BAZ1B	Bromodomain Adjacent to Zinc Finger Domain 1B
BMP	Bone Morphogenetic Protein
bp	Base pair
BPTF	Bromodomain and PHD domain Transcription Factor
CdLS	Cornelia de Lange syndrome
CHD7	Chromodomain helicase DNA binding protein 7
ChIP	Chromatin ImmunoPrecipitation
CMA	Chromosomal MicroArray
CNS	Central Nervous System
CNV	Copy Number Variation
CSB	Cockayne Syndrome protein B
DD	Distal Deletion
DEG	Differentially Expressed Gene
dpc	Days post coitum
EIF4H	Eukaryotic Initiation Factor 4H
ELN	Elastin
EMT	Epithelial to Mesenchymal Transition
ENC	Enteric Nervous System
ENCC	Enteric Neural Crest Cell

ESC	Embryonic Stem Cell
FACS	Fluorescence-Activated Cell Sorting
FD	Familial Dysautonomia
FD	Fold Change
FDR	False Discovery Rate
FGF	Fibroblast Growth Factor
FISH	Fluorescent In Situ Hybridization
GO	Gene Ontology
GTF2I	General Transcription Factor 2I
hNECs	human NeuroEctodermal cell
iPSC	Induced Pluripotent Stem Cell
IQ	Intelligence Quotient
ISWI	Imitation SWItch
KD	Knock Down
LC-MS/MS	Liquid chromatography–tandem MS
LCR	Low Copy Repeats
LH	Loop-Helix
LIMK1	LIM-Kinase1
Mb	Million base pairs
miRNA	Micro RNA
MLPA	Multiplex Ligation-dependent Probe Amplification
mRNA	Messenger RNA
MSC	Mesenchymal Stem Cell
Myb-bp1a	Myb-binding protein 1a
NAHR	Non-Allelic Homologous Recombination
NC	Neural Crest
NCSC	Neural Crest Stem Cell

NM1	Nuclear Myosin 1
NPC	Neural Precursor Cell
NR2F1/2	Nuclear Receptor subfamily 2 group F members 1/2
PCA	Principal Component Analysis
PCNA	Proliferating Cell Nuclear Antigen
PD	Proximal Deletion
PDGF	Platelet-Derived Growth Factor
PHD	Plant HomeoDomain
PiRNA	Piwi-interacting RNA
RCPS	Richieri-Costa-Pereira syndrome
RFC2	Replication Factor C subunit 2
SCNT	Somatic Cell Nuclear Transfer
scr	scrambled
sh	short hairpin
Shh	Sonic hedgehog
SMAD	Small Mother Against Decapentaplegic
SMARCA5	SWI/SNF related, Matrix associated, Actin dependent Regulator of Chromatin, subfamily A, Member 5
SMC	Smooth Muscle Cell
SNP	Single Nucleotide Polymorphism
STEMCCA	STEM Cell Cassette
STX1A	Syntaxin 1A
SVAS	SupraValvular Aortic Stenosis
TF	Transcription Factor
TFAP2A	Transcription Factor AP-2-Alpha
TGFβ1	Transforming Growth Factor β1
TSS	Transcription Start Site

VEE Venezuelan Equine Encephalitis

WAC WSTF/Acf1/cbp146

WAKZ WSTF/Acf1/KIAA0314/ZK783.4

WBS Williams-Beuren Syndrome

WBSCR Williams-Beuren Syndrome Critical Region

WSTF Williams Syndrome Transcription Factor

INDEX OF FIGURES AND TABLES

1. ABSTRACT

The hemizygous deletion and duplication of 28 genes in the 7q11.23 region is responsible for two neurodevelopmental disorders: Williams-Beuren syndrome (WBS) and 7q11.23 duplication syndrome (7dup), respectively. These two syndromes exhibit both shared and symmetrically opposite clinical traits, thus pointing to a dosage-sensitive impact of a small subset of genes, mostly affecting cognition, sociality and Neural Crest (NC)-mediated craniofacial and cardiovascular development.

Despite some studies on mouse models and patients with smaller deletions have established possible associations between specific genes in the 7q11.23 region and specific features of the disease, the vast majority of the genes remains without a clear connection.

BAZ1B, a transcriptional regulator and chromatin remodeler included in the region, is a prime candidate to study disease-associated NC alterations, because of its critical role in the migration of Neural Crest Stem Cells (NCSCs) in *Xenopus* and in the acquisition of craniofacial defects in mice lacking this protein. Intriguingly, WBS patients bearing a partial deletion of the region that spares few genes, including *BAZ1B*, display milder craniofacial dysmorphisms, further pointing to its involvement in this specific phenotype.

In order to define the BAZ1B-dependent alterations responsible for NC-related defects in the two genetic conditions, we integrated transcriptomic analysis with enhancer profiling, given the critical role of their regulatory architecture in NCSC development and function.

First, we derived *bona fide* NCSCs from a large and uniquely informative cohort of patient-specific induced pluripotent stem cells (iPSCs), including some derived from atypical *BAZ1B*-sparing deletions. Our results suggest that gene dosage alterations at the 7q11.23 region do not affect the cell-specific differentiation potential and the differentiation process itself. In addition, RNA profiling of NCSCs indicates dysregulation of genes specifically involved in NC-related biological processes and functions.

Next, I selectively downregulated *BAZ1B* levels by RNA interference in both control- and patient- NCSC lines and I performed RNA-sequencing analysis to dissect its contribution to transcriptional dysregulation. I uncovered a subset of genes whose levels followed *BAZ1B* levels either in a direct or inverse fashion, indicating clear *BAZ1B* dosage-dependent transcriptional alterations. These genes were mostly involved in pathways related to bone and cardiovascular system development and cell migration and motility.

Moreover, the observed impact of *BAZ1B* knock-down on human NCSC migration confirmed the evidence in animal models and shed new lights into BAZ1B-dependent mechanisms responsible for the acquisition of NC-related phenotypes.

Finally, preliminary analyses indicate a limited, but potentially crucial, BAZ1B-dependent remodeling on NCSC putative active enhancer regions.

Taken together, these results provide new inroads into a functional partitioning of BAZ1B transcriptional and epigenetic roles in NCSCs and they increase our understanding of these two debilitating neurodevelopmental disorders which will eventually pave the way for the development of novel effective therapies.

2. INTRODUCTION

The specification of neural tissues and the establishment of their architecture constitute a critical and extremely complex stage of embryo development. In particular, in humans, it begins at day 18 of development with the formation of the notochord starting from condensed mesodermal cells. This process is followed by the generation of the neural plate and the neural tube, the ancestors of the central nervous system, in a process called neurulation. Specifically, the neural tube gives rise to three main populations: i) neural stem cells that proliferate and produce neural precursors, ii) neuroblasts, non-dividing cells that differentiate into neurons and iii) neural crest cells that migrate in the developing embryo and generate peripheral sensory neurons, melanocytes and bone and cartilage craniofacial structures (Purves et al., 2008).

During neurodevelopment, a relevant number of signalling pathways has to intervene in a cooperative and coordinate fashion in order to allow the correct cell specification and organization in structures. Hence, genetic alterations, such as single mutations, deletions or duplications, affecting any step of these pathways can be responsible for the onset of several neurodevelopmental disorders.

Neurodevelopmental disorders are a class of diseases that typically manifest in the early phases of development, usually before school age, characterized by an impairment of social and personal skills because of developmental deficiencies of variable entity (Mitchell, 2014).

Neurodevelopmental disorders include:

1. intellectual disability, characterized by deficits in reasoning, learning and other mental abilities;
2. communication disorders that involve language and speech capacities;
3. attention-deficit/hyperactivity disorder (ADHD), marked by disorganization and inattention;

4. Autism Spectrum Disorder (ASD), that affects social interaction and communication;
5. motor disorders (movement and coordination);
6. specific learning disorders.

Comorbidity frequently occurs between different neurodevelopmental disorders; for example, children with autism often have intellectual disabilities, while learning disorders are commonly diagnosed in ADHD patients.

In this study, I focused on two peculiar Copy Number Variation (CNV)-dependent neurodevelopmental disorders: Williams-Beuren syndrome (or Williams syndrome, OMIM 194050) and 7q11.23 duplication syndrome (OMIM 609757), respectively caused by the deletion or duplication of 26-28 genes in the 11.23 region on chromosome 7q (Pober, 2010; Somerville et al., 2005).

2.1. Williams-Beuren syndrome and 7q11.23 duplication syndrome

2.1.1. Williams-Beuren syndrome

In 1956 Schlesinger and colleagues observed the first cases of Williams-Beuren syndrome (WBS) with the description of patients with hypercalcemia in association with cognitive alterations and a particular facial appearance (Schlesinger et al., 1956). Few years later Williams (Williams et al., 1961) and Beuren (Beuren et al., 1962) reported on Circulation the clinical description of a child having a shrinkage of the aorta (supravalvular aortic stenosis, SVAS) and peculiar facial features, that also helped in the diagnosis of three other patients. WBS was then recognized as a new syndrome only in 1964 when idiopathic hypercalcemia and SVAS were described as common alterations of the same disease (Garcia et al., 1964).

Epidemiological studies by Stromme and co-workers reported a prevalence of 1 case in 7500 (Strømme et al., 2002).

2.1.1.1. WBS clinical features

The significant number of genes involved in WBS deletion results in the complex clinical picture of WBS patients, in which numerous systems and organs are affected.

Cardiovascular alterations. Cardiovascular abnormalities and complications are considered the major cause of death in WBS patients, with a higher cardiovascular-associated mortality when compared to controls (Wessel et al., 2004). 70% of WBS patients present SVAS, a narrowing of medium and large arteries that follows the thickening of the vascular tunica media due to Smooth Muscle Cells (SMCs) overgrowth (Pober et al., 2008). Hypertension develops in 50% of WBS patients, probably as a physiological answer to the altered vessel structure (Faury et al., 2003).

Cognitive and social alterations. WBS developmental and cognitive profile is often described as a pattern of peaks (strengths) and valleys (weaknesses). WBS patients present strengths in facial processing and recognition as well as higher musical abilities when compared to controls (Deruelle et al., 2005; Gagliardi et al., 2003). On the contrary visuospatial construction is compromised in all WBS patients and the impairment in saccadic eye movements might contribute to these deficits (van der Geest et al., 2004). Intellectual disability ranges from mild to moderate with an average Intelligence Quotient (IQ) of 50-60, but appears to be more variable when the entire spectrum of young and adult WBS patients are considered (Martens et al., 2008). Moreover, some WBS patients were shown to have a higher verbal IQ than performance IQ. Language skills are preserved but with a delay in the main language achievements in young children (Karmiloff-Smith et al., 1997; Martens et al., 2008).

Because of their high empathy, friendliness and social abilities WBS patients often exhibit a “cocktail party” type personality (Pober, 2010). At the same time, more than 80% of

WBS adult display anxiety, frequent obsessions and fears and irritability, as well as hyperactivity (*e.g.,* ADHD) (Davies et al., 1998; Leyfer et al., 2006).

Metabolism alterations. Calcium abnormalities, in particular hypercalcemia and hypercalciuria, are often associated with WBS, especially in young children (Cagle et al., 2004). 1,25-dihydroxyvitamin D increased levels and alterations in calcitonin production and release were suggested as possible explanations for hypercalcemia (Culler et al., 1985).

In addition, impaired glucose tolerance and diabetes mellitus were diagnosed in several adult WBS patients (Morris et al., 1988).

Craniofacial alterations. According to many reports WBS patient facial characteristics (initially described as 'elfin facies') can be a useful tool for diagnosis, because of their peculiarity and reproducibility over patients (Beuren et al., 1962; Williams et al., 1961). These abnormalities include few hard tissue defects such as a broad forehead, a small mandible and a deficient chin button (Mass and Belostoky, 1993) and many soft tissue alterations, that comprise low-set ears, periorbital fullness, strabismus, epicanthal folds, a stellate iris pattern, full cheeks and lips, a wide mouth and a long smooth philtrum (Burn, 1986; Mass and Belostoky, 1993). Dental anomalies including late eruption, malocclusion, hypodontia and microdontia are also frequent in WBS individuals (Friedman and Mills, 1969).

2.1.2. 7q11.23 duplication syndrome

WBS was already well defined and characterized when the first patient bearing a duplication of the 7q11.23 region was described by Somerville and colleagues in 2005 (Somerville et al., 2005).

Since then, 7q11.23 duplication syndrome (also known as Somerville–van der Aa syndrome or Williams-Beuren region duplication syndrome, hereafter called 7dup) was diagnosed in around sixty patients, including both children and adults (Sanders et al.,

2011; Torniero et al., 2008). The prevalence was estimated to be between 1:7.500 and 1:20.000 (Van der Aa et al., 2009).

2.1.2.1. 7dup clinical features

The later description of 7dup patients was probably due to the fact that clinicians were more prone to expect duplication of the region resulting in a clinical phenotype overlapping with WBS. Despite some features are indeed shared between WBS and 7dup (*e.g.*, growth delay, ADHD and some facies dysmorphisms), many characters are completely opposite (Somerville et al., 2005).

Cognitive and social alterations. Recent studies reported a wide range of intellectual abilities across 7dup patients during development with an overall IQ level considerably higher than the one described for WBS patients (*e.g.*, 92 for adult 7dup *vs* 64 for adult WBS patients) (Mervis et al., 2015). This observation is in line with previous findings supporting a more consistent impact of the 7q11.23 deletion on ability levels that the respective duplication (Van der Aa et al., 2009). The severe delay in expressive and receptive language already described by Somerville is another key feature of 7dup patients, with more than 75% of children having Speech Sound Disorder (Mervis et al., 2015). The weakness in verbal abilities and the relative strength in visuospatial skills of 7dup children are clearly opposite to what is observed in WBS patients. Anxiety disorders are present in both patients but while WBS children have specific phobia, the high majority of 7dup children display social phobia and selective mutism, that is in contrast with WBS patient hypersociality (Mervis et al., 2015). Moreover, Sanders and collaborators reported a strong association of the 7q11.23 region duplication with ASD in around 30% of patients (Sanders et al., 2011).

Cardiovascular alterations. Recent findings reported aortic dilation, which is symmetrically opposite to WBS aortic stenosis, to be present in all the nine 7dup patients of the study and to mostly affect the ascending aorta (Parrott et al., 2015).

Craniofacial alterations. These patients display brachycephaly, straight eyebrows, long eyelashes, posteriorly rotated ears and a high palate (Morris et al., 2015; Somerville et al., 2005).

Antagonistic alterations are also present at the craniofacial level. Contrary to WBS, 7dup children have a high, broad nose, a short philtrum, thin lips and generous-sized teeth. The broad forehead and the dental malocclusion due to a smaller jaw are instead found in both syndromes.

A comparison of WBS and 7q11.23 duplication syndrome clinical features can be found in **Table 1**.

	Williams Beuren syndrome	**7q11.23 duplication syndrome**
Cognitive and social alterations	Developmental delay/ intellectual disability	Developmental delay/ intellectual disability
	Weakness in spatial skills	Relative strength in spatial skills
	Preserved language skills	Severe language impairment
	Hypersociability	Antisocial behaviour
	ADHD	ADHD
	Hyperacusis	Normal hearing
	Anxiety (specific phobia)	Anxiety (social phobia)
Cardiovascular alterations	SVAS	Aortic dilation
Craniofacial alterations	Broad forehead	Broad forehead
	Low nasal root	High and broad nose
	Long philtrum	Short philtrum
	Normal palate	High palate
	Full lips	Thin lips
	Small jaw	Small jaw
	Dental malocclusion	Dental malocclusion
Other anomalies	Growth retardation	Growth retardation
	Low birth weight	Low birth weight

Table 1. Summary of common (green) and opposite (orange) features of WBS and 7dup patients

2.1.3. The 7q11.23 region

2.1.3.1. WBS and 7dup are contiguous gene disorders

At the beginning, the exposure to high doses of vitamin D during pregnancy was considered as the main cause of WBS (Friedman and Mills, 1969). Craniofacial alterations and lesions similar to SVAS were indeed observed in rabbit offspring whose mothers received excessive amounts of vitamin D (Friedman and Roberts, 1966).

In the early 90's, genetic was shown to play a major role in WBS development with the discovery of the autosomal dominant inheritance of the disease in few cases (Sadler et al., 1993) and the description of a mutation in the *ELN* gene as responsible for familial SVAS (Curran et al., 1993). WBS patients were screened for *ELN* mutations and one *ELN* allele was found to be entirely absent, while no *ELN* point-mutations were detected, introducing the hypothesis of WBS being a contiguous gene disorder (Ewart et al., 1993). Contiguous gene disorders are indeed caused by the deletion/duplication of multiple adjacent genes on a specific chromosome.

Thanks to further analyses based on microsatellite DNA markers, the entire WBSCR and of the genes affected by the deletion (*e.g.*, LIM-Kinase1, *LIMK1* (Tassabehji et al., 1996); Replication Factor C subunit 2, *RFC2* (Peoples et al., 1996); Syntaxin 1A, *STX1A* (Osborne et al., 1997) and Bromodomain Adjacent to Zinc Finger Domain 1B, *BAZ1B* (Lu et al., 1998)) were identified. Moreover, the high homology of a DNA marker close to the *ELN* gene with other proximal *loci* helped to pinpoint repetitive sequences flanking the region, that started to be considered potentially responsible for the mutational event in WBS patients (Robinson et al., 1996). The 28 genes composing the Williams-Beuren syndrome critical region (WBSCR) were later confirmed to be flanked by low-copy-repeat (LCR) blocks (also known as duplicons) named A, B and C, containing genes and pseudogenes with high sequence homology (**Figure 1**). These blocks are organised in

three different complexes, centromeric, medial and telomeric, according to their position in the region (Hockenhull et al., 1999).

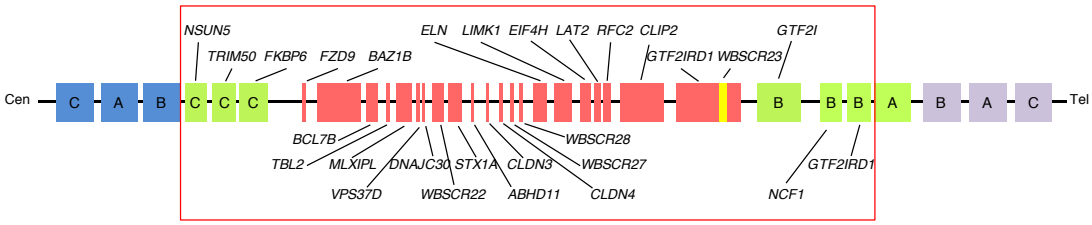


Figure 1. Williams-Beuren syndrome critical region on chromosome 7q11.23

The 28 genes forming the WBSCR are represented inside the red frame. A, B and C boxes represent centromeric (blue), medial (green) and telomeric (violet) duplicons. Redrawn from (Pober, 2010).

The high similarity of duplicon sequences facilitates an unequal crossing over event named Non-Allelic Homologous Recombination (NAHR) that then predisposes to an interchromosomal or, less frequently, interchromatid misalignment during meiosis, thus leading to the deletion and to the reciprocal duplication of the region (Schubert, 2009). Intrachromatid recombination generating a deletion and a reciprocal circular acentric chromosome, but not a duplication, is also possible and occurs with a higher rate than interchromatid recombination. This might explain the higher frequency of deletion over duplication events, with a deletion-duplication rate of 2:1 observed on sperm samples (Turner et al., 2008).

In most cases, breakpoints occur in the B blocks contained in the medial and centromeric duplicons and drive to the deletion of ~ 1.5 million base pairs (Mb) in 95% of WBS patients and of ~ 1.8 Mb in the remaining 5% of cases (Bayés et al., 2003).

2.1.3.2. Genotype-phenotype correlation

The WBSCR contains a total number of 28 genes coding for transcription (*i.e.*, General Transcription Factor 2I, *GTF2I*) and translation (*i.e.,* Eukaryotic Initiation Factor 4H, *EIF4H*) factors, chromatin regulators (*i.e., BAZ1B*) and proteins involved in other relevant cellular functions such as cytoskeletal organization (*i.e., LIMK1*) and intracellular transport (*i.e., STX1A*).

Table 2 recapitulates the main functions of proteins coded by WBSCR genes (Tassabehji, 2003).

Mouse models for single WBS genes as well as patients bearing a smaller or larger deletion/duplication of WBSCR represent valuable tools to investigate molecular mechanisms at the root of 7q11.23 CNVs and the specific contribution of a WBS gene to the diseases.

Mouse models. Mice bearing a hemizygous deletion/duplication better mirror the diseases, but in many cases a more definite phenotype was visible in a homozygous background (Crackower et al., 2003; Fujiwara et al., 2016).

The vast majority of mouse models were obtained following the selective targeting of single WBS genes. However in 2008, Li and colleagues, thanks to the high similarity of the WBSCR with a region on the mouse chromosome 5G, were able to obtain mice carrying two deletions spanning both the proximal PD (from *Gtf2i* to *Limk1*) and the distal DD (from *Limk1* to *Fkbp6*) part of the human region through Cre-LoxP recombination (Li et al., 2009). PD mice recapitulate increased sociability, while DD mice display cognitive defects. Mice bearing the two deletions (D/P) have craniofacial and motor skill alterations, together with growth delay.

The main phenotypic features of single gene mouse model are reported in **Table 3**.

Atypical individuals. Patients with shorter deletions or duplications are extremely valuable for genotype–phenotype correlation, but they are rare.

To date, among all patient clinical traits, only SVAS has been univocally associated with a specific WBSCR gene, namely *ELN* (Curran et al., 1993).

In the recent years, other studies on atypical patients tried to link disease phenotypes to a specific gene or to a limited subset of genes, but not always there was consistency among results. The current hypothesis is that the interval that goes from *GTF2I* to *ELN* is most probably responsible for the neurocognitive profile, with *GTF2I* and *GTF2IRD1* being the top candidate genes (Antonell et al.; Morris et al., 2003). *GTF2IRD1* was also suggested as a key regulator of craniofacial development according to the phenotype of a knockout mouse generated by Tassabehji and collaborators (Tassabehji et al., 2005). However, a second *Gtf2ird1* knockout mouse on a different strain did not displayed craniofacial alterations, thus questioning the strength of previous findings (Palmer et al., 2007).

Longer deletions of the region that include Huntingtin-interacting protein 1 (*HIP1*) and tyrosine 3-monooxygenase/tryptophan 5-monooxigenase activation protein gamma (*YWHAG*) genes resulted in more severe neurocognitive deficits (Fusco et al., 2013).

On the contrary, the few patients showing atypical duplications (either smaller or longer) did not seem to differ from patients with typical duplication, but comparisons with larger group of individuals are needed (Zarate et al., 2014).

Gene	Protein Name	Protein Function	References
NSUN5	NOP2/Sun RNA Methyltransferase Family Member 5	Involved in ribosomal RNA methylation.	(Schosserer et al., 2015)
TRIM50	Tripartite Motif Containing 50	E3 ubiquitin ligase. Regulates vesicular trafficking in gastric cells.	(Micale et al., 2008; Nishi et al., 2012)
FKBP6	FK506 Binding Protein 6	Acts as co-chaperone during spermatogenesis. Required for PiRNA biogenesis	(Ichiyanagi et al., 2014; Meng et al., 1998)
FZD9	Frizzled Class Receptor 9	Receptor for WNT2. Controls new bone formation.	(Chalangkarn et al., 2016; Heilmann et al., 2013)
BAZ1B	Bromodomain Adjacent To Zinc Finger Domain 1B	Chromatin remodeler, part of WICH complex (see chapter 2.4).	(Lu et al., 1998; Poot et al., 2004)
BCL7B	BCL Tumor Suppressor 7B	Tumor suppressor. Involved in apoptosis and Wnt signalling pathway.	(Uehara et al., 2015)
TBL2	Transducin Beta Like 2	Beta transducin protein. Plays a role in endoplasmic reticulum stress.	(Pérez Jurado et al., 1999; Tsukumo et al., 2016)
MLXIPL	MLX Interacting Protein Like	Transcriptional repressor. Regulates triglyceride synthesis.	(Cairo et al., 2001)
VPS37D	VPS37D, ESCRT-I Subunit	Part of the endosomal sorting complex. Regulates vesicular trafficking.	(Katzmann et al., 2001)
DNAJC30	DnaJ Heat Shock Protein Family Member C30	Molecular chaperone involved in protein binding and folding.	(Merla et al., 2002)
WBSCR22	Williams-Beuren Syndrome Chromosome Region 22	Involved in ribosomal RNA processing.	(Haag et al., 2015)
STX1A	Syntaxin 1A	Nervous system-specific protein involved in hormone and neurotransmitter exocytosis.	(Bennett et al., 1993)
ABHD11	Abhydrolase Domain Containing 11	Important for lipid metabolism.	(Arya et al., 2017)
CLDN3	Claudin 3	Integral membrane protein part of tight junctions in kidney epithelial cells.	(Günzel and Yu, 2013)
CLDN4	Claudin 4	Integral membrane protein part of tight junctions in liver epithelial cells.	(Günzel and Yu, 2013)
WBSCR27	Williams Beuren Syndrome Chromosome Region 27	Belongs to the ubiE/COQ5 methyltransferase family.	(Hahn and Lee, 2005)
WBSCR28	Williams Beuren Syndrome Chromosome Region 28	Unknown function.	
ELN	Elastin	Component of elastic fibers. Responsible for SVAS in WBS patients.	(Curran et al., 1993; Ge et al., 2012)
LIMK1	LIM Domain Kinase 1	Regulates actin polymerization. Involved in cell motility and cell cycle.	(Gorovoy et al., 2005; Kawano et al., 2013)
EIF4H	Eukaryotic Translation Initiation Factor 4H	Part of the translation initiation complex. Promotes protein synthesis.	(Marintchev et al., 2009)
LAT2	Linker For Activation Of T-Cells Family Member 2	Important for T lymphocyte development and activation.	(Fuller et al., 2011)
RFC2	Replication Factor C Subunit 2	Activates DNA polymerases during replication.	(Peoples et al., 1996)
CLIP2	CAP-Gly Domain Containing Linker Protein 2	Modulates the interaction between microtubules and dendritic lamellar bodies.	(Hoogenraad et al., 1998)
GTF2IRD1	GTF2I Repeat Domain Containing 1	Potential role in cell cycle progression and chromatin regulation.	(Corley et al., 2016; Howard et al., 2012)
WBSCR23	Williams Beuren Syndrome Chromosome Region 23	Unknown function.	
GTF2I	General Transcription Factor 2I	Promotes DNA translesion synthesis and genomic stability.	(Fattah et al., 2014)
NCF1	Neutrophil Cytosolic Factor 1	Subunit of neutrophil NADPH oxidase. Responsible for chronic granulomatous disease.	(Francke et al., 1990; Sareila et al., 2013)
GTF2IRD2	GTF2I Repeat Domain Containing 2	Transcription factor of the TFII-I family.	(Makeyev et al., 2004)

Table 2. List of the 28 genes included in the WBSCR region

Gene	Mouse model	Disease-relevant phenotypes	References
FKBP6	*Fkbp6 +/-*	No phenotype.	(Crackower et al., 2003)
	Fkbp6 -/-	Aspermia and infertility.	
FZD9	*Fzd9 +/-*	Reduced seizure threshold and hippocampus development defects.	(Chailangkarn et al., 2016; Ranheim et al., 2005)
	Fzd9 -/-	Reduced seizure threshold, hippocampus development and visuospatial learning defects and abnormal B-cell development.	
BAZ1B	*Baz1b +/-*	Mild craniofacial dysmorphisms.	(Ashe et al., 2008)
	Baz1b -/-	Craniofacial dysmorphisms and growth retardation.	
MLXIPL	*Mlxipl +/-*	No data available.	(Iizuka et al., 2004)
	Mlxipl -/-	Reduced lipogenesis and glycolysis.	
STX1A	*Stx1a +/-*	No relevant phenotypes.	(Fujiwara et al., 2016)
	Stx1a -/-	Neuropsychological abnormalities and unusual social behavior.	
ELN	*Eln +/-*	Thickening of the aorta wall.	(Li et al., 1998a) (Li et al., 1998b)
	Eln -/-	Obstructive arterial disease leading to embryonic lethality.	
LIMK1	*Limk1 +/-*	No data available.	(Kawano et al., 2013; Meng et al., 2003)
	Limk1 -/-	Altered spine morphology and reduced bone mass.	
EIF4H	*Eif4h +/-*	No phenotype.	(Capossela et al., 2012)
	Eif4h -/-	Reduced fertility, growth retardation and impaired cognitive/behavioral abilities.	
LAT2	*Lat2 +/-*	No data available.	(Zhu et al., 2006)
	Lat2 -/-	Autoimmune syndrome.	
CLIP2	*Clip2 +/-*	Mild growth and motor coordination defects, hippocampal dysfunction.	(Hoogenraad et al., 2002)
	Clip2 -/-	Growth and motor coordination defects, hippocampal dysfunction.	
GTF2IRD1	*Gtf2ird1 +/-*	Learning and memory deficits, hypersociability.	(Tassabehji et al., 2005; Young et al., 2008)
	Gtf2ird1 -/-	Learning and memory deficits, hypersociability, craniofacial dysmorphisms and motor dysfunction.	
GTF2I	*Gtf2i +/-*	Microcephaly, growth retardation, craniofacial defects and hypersociability,	(Enkhmandakh et al., 2009; Mervis et al., 2012; Sakurai et al., 2011)
	Gtf2i -/-	Embryonic lethality with neural tube closure defects.	
	Gtf2i +/dup and Gtf2i dup/dup	Separation anxiety.	

Table 3. Mouse models for the genes included in the WBSCR

Single gene mouse models recapitulate patient cognitive defects (green), craniofacial defects (blue), cardiovascular defects (orange) or other disease-related defects (white).

2.1.4. WBS and 7dup diagnosis and treatment

WBS and 7dup are nowadays diagnosed through the detection of the 1.5-1.8 Mb heterozygous deletion or duplication at chromosome 7q11.23. Currently, the most used molecular techniques for the identification of copy number variations of sequences are:

1. Chromosomal MicroArray (CMA) based on oligonucleotide arrays that recognize duplicated or deleted chromosomal segments (Beaudet, 2013);
2. targeted analyses, such as Fluorescence In Situ Hybridization (FISH) that uses sequence specific-fluorescent probes (Elçioglu et al., 1998), Multiplex Ligation-dependent Probe Amplification (MLPA) based on a multiplex Polymerase Chain Reaction (PCR) (van Hagen et al., 2007) and quantitative PCR (qPCR) (Schubert and Laccone, 2006).

Unfortunately, none of the currently available treatments for WBS and 7dup patient defects are curative. Most of endocrine alterations of WBS patients such as subclinical hypothyroidism, glucose intolerance and hypercalcemia are constantly monitored with routine screening. SVAS and aortic dilation represent the most critical manifestations of the two diseases and, often, monitoring is not enough but surgery or less invasive procedures are required. Cognitive and social disabilities are addressed through early intervention and education programs, including counselling, speech and behavioural therapies (Pober, 2010).

2.2. The neural crest

To date, the vast majority of reports on neurodevelopmental disorders are specifically focused on brain and on neurological alterations, while much less studies considered Neural Crest Stem Cells (NCSCs) and their derivatives as a worthy tool to target, even though their functions are clearly compromised in these diseases.

The Neural Crest (NC), also known as the 'fourth germ layer' (Hall, 2000) is a transient and multipotent cell population unique to vertebrates. It originates during gastrulation and gives rise to a large variety of cell types, including adipocytes, osteocytes and chondrocytes, which composed craniofacial structures, SMCs in the vessels, melanocytes, and Schwann cells and sensory neurons of the peripheral nervous system (**Figure 2**) (Muñoz and Trainor, 2015).

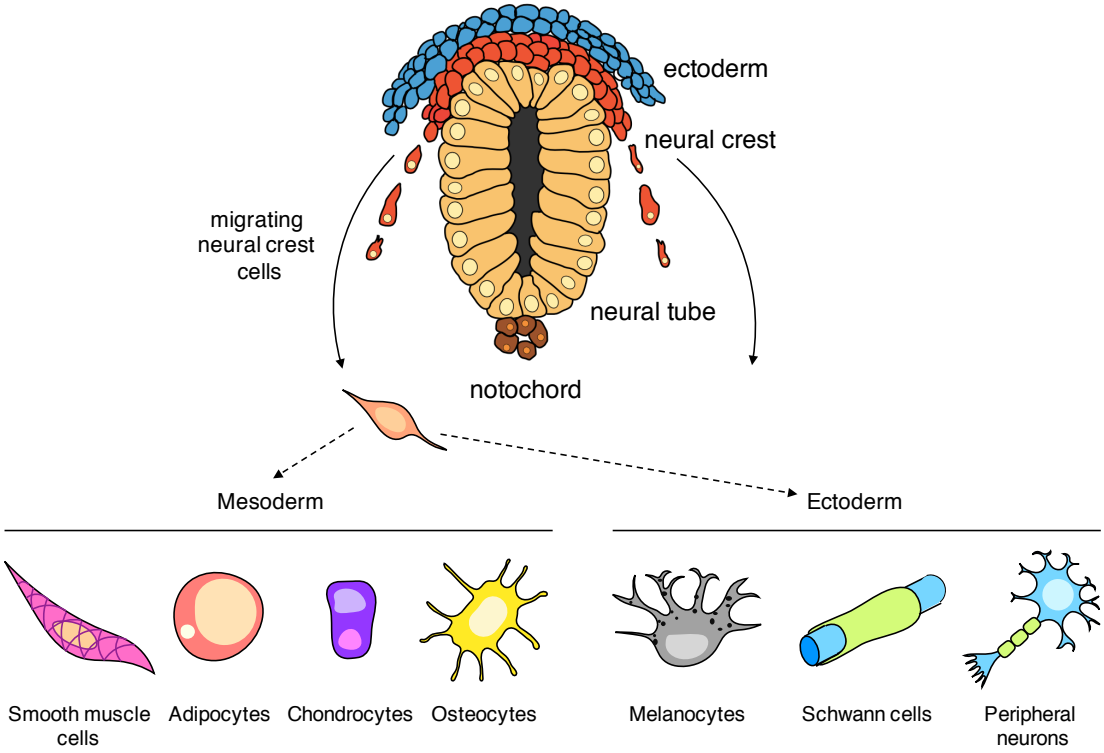


Figure 2. Schematic representation of NC derivatives

NCSCs originate from the neural plate border region and then migrate giving rise to cells with both mesodermic (*i.e.,* smooth muscle cells, adipocytes, chondrocytes and osteocytes) and ectodermic (*i.e.,* melanocytes, Schwann cells and peripheral neurons) origins. Redrawn from 'http://web.biologie.uni-bielefeld.de/cellbiology/index.php/research/neural-crest-derived-stem-cells'.

2.2.1. Neural crest induction and migration

NC was originally named *Zwischenstrang (*intermediate cord) by Wihelm His, that discovered it in 1868 (His, 1868). The origin of the NC was highly debated in the past and most of the current knowledge derives from studies on *Xenopus laevis* and avian embryos (Sadaghiani and Thiébaud, 1987; Selleck and Bronner-Fraser, 1995). Tissue grafting experiments performed in the above mentioned animal models indicated that interactions occurring between the neural plate and the non-neural ectoderm were fundamental for NC formation (Mancilla and Mayor, 1996; Selleck and Bronner-Fraser, 1995). In 2000, the identification of NCSCs in mice provided valuable data on facial structures (Chai et al., 2000) and heart development (Jiang et al., 2000).

NC induction starts at early stages of gastrulation (*i.e.*, around day 20 in humans) at the neural plate border region and continues until neural tube closure, in a process guided by a series of environmental signals acting at specific time points. First, extracellular molecules such as Wnt, fibroblast growth factor (FGF) and bone morphogenetic proteins (BMPs), are released within the neural plate border and promote the activation in neural plate border cells of specific transcription factors (TFs) named 'neural plate border specifiers', that include *Msx1*, *Pax3*, *Pax7*, *Zic1* and *Tfap2* (Basch et al., 2006; Garnett et al., 2012; Knight et al., 2003; Monsoro-Burq et al., 2005).

Next, extracellular signals and neural plate border specifiers cooperate and activate 'NC specifiers genes' encoding *Snail*, *Slug*, *Foxd3*, *Sox9* and *Sox10* that form a regulatory loop fundamental for the maintenance of multipotency in NCSCs (Cheung et al., 2005). In addition, NC specifiers promote NCSC epithelial to mesenchymal transition (EMT) and

motility through the regulation of adhesion molecules (specifically cadherins) and Rho GTPase signalling pathways (Cheung et al., 2005).

Cell lineage tracing studies indicate that NCSCs migrate along specific routes throughout the embryo and the subsequent NCSC fates are specifically mediated both by their axial origin within the neural tube and by environmental signals (Liu and Cheung, 2016). In most vertebrates, NCSCs initiate migration shortly after neural tube closure; however, in mammals and amphibians a specific subtype of NCSCs, named cranial NCSCs, starts migrating prior to closure (Bronner and LeDouarin, 2012).

2.2.2. Neural crest subtypes

NCSC multipotency was first suggested by Stemple and Anderson in 1992 and then supported by further isolation and transplantation experiments in both rat and chick embryo in the following years (Stemple and Anderson, 1992).

NCSCs can be divided in 5 main subtypes, according to both their axial origin and the cell types they generate (**Figure 3**).

CRANIAL NCSC. In mammals, cranial NCSCs originate in the regions encompassing the forebrain, the midbrain and the anterior hindbrain and give rise to the main facial structures. In particular, cells derived from the forebrain and the midbrain form the palate, the frontonasal process and the mesenchyme of the first pharyngeal arch that will eventually give rise to jawbones and ear bones malleus and incus. Cells from the anterior hindbrain generate the mesenchyme of the second pharyngeal arch from which facial cartilage and the stapes bone of the ear are derived. Finally, cranial NCSCs also generate the mesenchyme of the third, fourth and sixth pharyngeal arches that produce muscles and bones of the neck; while the fifth pharyngeal arch degenerates in humans (Bhatt et al., 2013).

The subdivision of *Hox*-gene-negative- and *Hox*-gene-positive cranial NCSCs is conserved across vertebrates and required for their specification (Bhatt et al., 2013).

CARDIAC NCSCs. They derive from somite 1 to 3 and partially overlap with the anterior portion of the vagal region. They form the endothelium and the muscle cells in the aortic arch arteries and the septum between the aorta and the pulmonary artery (Waldo et al., 1998).

VAGAL NCSCs. They arise from posterior hindbrain in a region that includes somite 1 to 7 and give rise to the Enteric Nervous System (ENS) along the gastrointestinal tract (M Hao, 2016).

TRUNK NCSCs. They derive from somite 8 to 28 and follow two different pathways while migrating: i) the trunk NC cells that migrate **dorsolaterally** enter the dermis, reach the skin and hair follicles and differentiate into melanocytes; ii) the cells following a **ventral** route migrate through the anterior portion of sclerotome and generate sensory neurons of the dorsal root ganglia. Some of these cells further migrate to form sympathetic neurons, Schwann cells along the spinal nerves and secretory cells of the endocrine system (Vega-Lopez et al., 2017).

SACRAL NCSCs. They lie posterior to somite 28 and are responsible for ENS generation together with vagal NCSCs (M Hao, 2016).

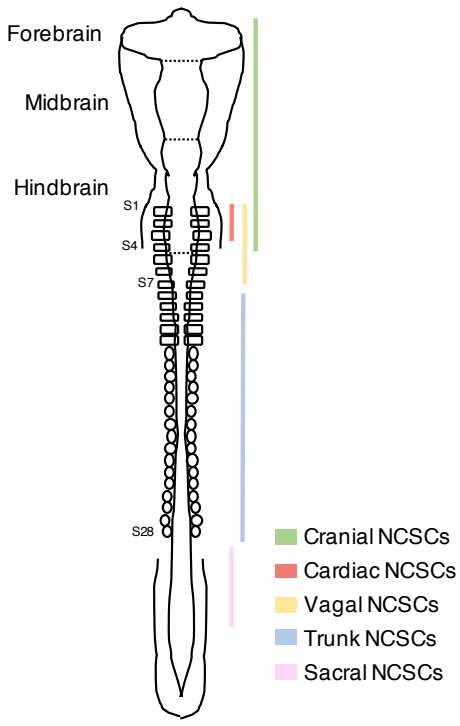


Figure 3. Schematic representation of NCSC subtype site of origin in the developing chick embryo

Cranial NCSCs originate from the forebrain, the midbrain and the hindbrain and give rise to the ectomesenchyme; cardiac NCSCs originate from somite 1-3 and produce SMCs in the aorta; trunk NCSCs migrate from somite 8-28 and form both melanocytes and sensory neurons; vagal NCSCs derive from somite 1-7 and together with sacral NCSCs generate the ENS. Redrawn from (Begbie, 2013).

2.2.3. Neural crest-associated disorders

Organ and tissue dysplasias due to NCSC defects in migration, differentiation, division or survival are referred as 'neurocristopathies' (Benish, 1975).

Neurocristopathies include two main components, tumors and malformations, that often coexist.

Neurocristopathic tumors include neuroblastomas affecting the sympathetic components, cutaneous cancers such as melanoma and Merkel cell carcinoma and few multi-system disorders that predispose to cancer, in particular neurofibromatosis type I (Etchevers et al., 2006).

The most common neurocristopathic malformations comprise congenital disorders that affect NC and NC-derived cells exclusively or, in most cases, in combination with other NC-unrelated tissues.

Hirschprung's disease, also known as aganlionic megacolon, is an intestinal mobility disorder that affects 1 in 5000 newborns. It is caused by the migration failure of enteric neural crest cells (ENCCs) in the distal intestine that leads to intestinal obstruction and dilation of the bowel (Parisi, 1993).

Familial dysautonomia is a neurodegenerative disorder of the autonomic nervous system that leads to a progressive loss of sensory, sympathetic and some parasympathetic neurons. It is characterized by insensitivity to pain and to temperature

and unbalanced blood pressure and it is mainly caused by a point mutation in *IKBKAP*, a subunit of the Elongator complex involved in transcription (Slaugenhaupt et al., 2001).

Bicuspid aortic valve (BAV) is a congenital disease caused by the fusion of two of the aortic valvular leaflets and to the formation of a bicuspid instead of a tricuspid valve (Wang et al., 2016). It is often associated with aneurysms in the ascending aorta, that is composed by SMCs derived from cardiac NCSCs (Majesky, 2007).

2.3. Induced pluripotent stem cells and disease modeling

Despite the undeniable advantages provided by animal models and atypical patients in reconstructing the molecular dysregulation of the diseases (see chapter 2.1.3.2.), more reliable models are necessary.

First, mice cannot fully recapitulate human neurodevelopment, especially in terms of cerebral cortex complexity.

On the other side, there are three main issues to consider regarding atypical patients:

1. usually, they are not diagnosed by the same subset of clinicians and thus subjected to the same types of cognitive and social test;
2. detailed mapping of atypical deletion or duplication are often missing;
3. they are single case report, thus making difficult to distinguish disease-related phenotypes from phenotypes due to genome variability between individuals.

The possibility to directly investigate the patient genome *in vitro* through the derivation of patient-specific cell types is a step forward into a trustworthy molecular characterization of many diseases (Antonell et al., 2010). However, in most cases, the actual diseased cells reside in poorly accessible tissues, such as the nervous, the cardiovascular and the musculoskeletal systems.

For example, multipotent NCSCs can be isolated not only from rodent and chick embryos but also from rodent and human adult tissues, including skin (Toma et al., 2001), heart (Tomita et al., 2005), bone marrow (Shi et al., 2016), gut (Kruger et al., 2002), cornea (Brandl et al., 2009) and dental pulp (Janebodin et al., 2011). Thus, adult NCSCs might represent an attractive tool for replacement therapies. However, unfortunately, they are very rare, they often have a reduced accessibility and they show limited self-renewal capacity and multipotency when compared to their embryonic counterparts (Achilleos and Trainor, 2012).

Cellular reprogramming techniques based on the conversion of accessible differentiated cell types, such as fibroblasts and blood cells, into induced Pluripotent Stem Cells (iPSCs) represent a breakthrough in disease modelling (Takahashi and Yamanaka, 2006). Indeed, iPSCs are pluripotent and can be potentially differentiated into all disease-relevant cell types, thus solving the aforementioned limit of tissue accessibility.

2.3.1. The route to iPSCs

The first attempts of cellular reprogramming involved the generation of embryos following the transplantation of somatic cell nuclei into enucleated oocytes with a technique known as somatic cell nuclear transfer (SCNT). In 1962, Gurdon was able to transfer nuclei of intestinal cells from adult frogs into unfertilized eggs and to obtain vital tadpoles (Gurdon, 1962). More than 30 years later, these initial findings were confirmed with the generation of mammals upon somatic cloning of mammary epithelial cells (Wilmut et al., 1997).

Another step forward has been the isolation and the establishment in culture of pluripotent Embryonic Stem Cells (ESCs) from mouse (Evans and Kaufman, 1981) and human blastocysts (Thomson et al., 1998) that were able to form derivatives of all the three embryonic germ layers when transplanted into immunodeficient mice. However, the employment, for research purposes, of ESCs derived from human embryos has opened many debates regarding the ethical implications of their usage. This has reduced both their availability and the feasibility of exploiting this cell type for disease studies.

Thanks to Takahashi and Yamanaka achievements in 2007, we can now bypass the main ethical questions raised in the past years by generating pluripotent cells directly from patient somatic cells (Takahashi et al., 2007). The two Nobel prizes showed that the forced expression of four TFs involved in pluripotency (*i.e.*, OCT3/4, SOX2, c-MYC and KLF4, also known as Yamanaka factors) was enough to induce iPSCs from both mouse (Takahashi and Yamanaka, 2006) and human fibroblasts (Takahashi et al., 2007).

iPSCs are very similar to ESCs in terms of morphology, differentiation potential and ability to form embryoid bodies, *in vitro*, and teratomas containing cells from the three embryonic germ layers, *in vivo* (Takahashi and Yamanaka, 2006; Takahashi et al., 2007).

2.3.2. Reprogramming protocols

Since 2006, the reprogramming protocol proposed by Takahashi and Yamanaka was widely modified, especially through the usage of both integrating and non-integrating approaches for the expression of the four Yamanaka factors.

Today, we can distinguish five different techniques for the delivery of pluripotency genes.

Retroviral/lentiviral vector. The original and still most used protocol is based on retroviral vectors (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). In 2009, Sommer and colleagues produced a single polycistronic lentiviral vector named “STEM Cell CAssette” (STEMCCA) to improve the delivery efficiency (Sommer et al., 2009).

However, the stable integration of these viral vectors is responsible for a significative risk of insertional mutagenesis that, together with the residual transgene expression, might affect the reprogramming efficiency as well as the iPSC differentiation capability (Warren et al., 2010). These problems were partially solved with the introduction of a Cre-excisable polycistronic vector in 2010 (Sommer et al., 2010).

Sendai viral vectors. A good alternative to retroviral and lentiviral vector is Sendai vector, derived from a single-stranded RNA virus that does not integrate in the cell genome, thus avoiding the aforementioned problems (Fusaki et al., 2009). Since this viral vectors

replicates in iPSCs, many passages might be required to completely get rid of it in host cells (González et al., 2011).

Synthetic modified messenger RNAs (mRNAs). This method is based on the daily transfection, for a total of 17 days, of five individual mRNAs coding for the classic four pluripotency factors plus *LIN28A* (Warren et al., 2010). It is a vector- and -integration free, and thus safer, approach and it was shown to increase the reprogramming efficiency (Adamo et al., 2015).

The combination of mRNAs and specific microRNAs (miRNAs) reduced the transfection period to less than two weeks and further increased the number of iPSC clones derived from patient fibroblasts (Adamo et al., 2015).

Self-replicating RNA. More recently, Yoshioka and collaborators developed a reprogramming approach that uses a single, synthetic self-replicating Venezuelan Equine Encephalitis (VEE) virus RNA replicon expressing the four Yamanaka factors (*OCT4*, *KLF4*, *SOX2* and *c-MYC* or *GLIS1*) (Yoshioka et al., 2013). This system is integration-free and highly simplifies the reprogramming procedure because it does not require the daily mRNA transfection.

2.3.3. iPSC and iPSC-derived cell applications

iPSCs have three main advantages:

1. They can be used to derive all disease-relevant cell types;
2. They carry the patients' unique genetic information and they can recapitulate the specific pathological defects;
3. They are renewable and, as such, an unlimited cell resource.

Based on all these strengths, we can distinguish two main fields of application for human iPSCs (Bellin et al., 2012; Wen, 2017):

1. Disease modelling. The direct investigation of the diseased cell can help in understanding the molecular mechanisms responsible for the patients' phenotypes and thus developing appropriate treatments.

2. Drug discovery. iPSCs can be used in drug screening to identify the effects of new candidate compounds and to assess whether cellular toxic responses are associated to them. This application represents a new type of human preclinical trial that could eventually limit animal use.

To address the first purpose, it is necessary to develop appropriate assays able to detect and discriminate specific phenotypes associated to patient derived cells, while determining how to intervene to rescue them would help to pinpoint the second iPSC usage aim.

To date, various diseases have been studied using patient-specific iPSCs and iPSC-derived cell types and several disease phenotypes have been reported. iPSCs are widely used to model neurodevelopmental disorders, while some challenges have to be considered in their applications in the study of age-related neurodegenerative disorders. Indeed, the resetting of the cell epigenetic state during reprogramming, together with the competitive clonal growth and the frequent cell divisions, are responsible for cell rejuvenation, that represents a limit for the modeling of all those diseases where ageing is a crucial risk factor (Rando and Chang, 2012).

In the vast majority of cases, iPSCs are differentiated towards somatic cells that are relevant for the specific disease defects, such as neuronal cells in neurodevelopmental disorders or cardiomyocytes in cardiovascular diseases.

However, iPSCs cannot be considered only as a transitional population needed for the generation of disease-relevant cell types. Indeed, we recently showed that the pluripotent state is already informative about the disease alterations (Adamo et al., 2015; Urban and Purmann, 2015). RNA sequencing (RNA-seq) analysis performed on a large cohort of WBS and 7dup patient- as well as healthy controls-derived iPSCs revealed a clear transcriptomic dysregulation of specific neurodevelopmental features of the two syndromes. According to Gene Ontology (GO) enrichment analysis, the top-ranking dysregulated biological processes included: i) neural crest cell development and

migration, which is relevant for the craniofacial alterations; ii) synapse assembly and axon guidance, consistent with the neurological defects; iii) cardiovascular system development, in line with SVAS and aortic dilation described in WBS and 7dup patients, respectively (Adamo et al., 2015).

2.3.3.1. iPSC-derived NCSC protocols

The vast majority of NCSC-related disorders primarily affect pediatric patients because of an abnormal migration, proliferation or differentiation of embryonic NCSCs. While NCSCs can be derived from adult tissues (see chapter 2.3), the isolation of human fetal NCSCs encounters both ethical and scientific challenges. For this reason, disease-specific NCSCs have been obtained following differentiation of iPSCs, mainly reprogrammed from patient skin fibroblasts, and constitute an extremely valuable tool for disease modelling of neurocristopaties and other NC-related disorders. Moreover, the use of patient-specific NCSCs would allow autologous transplantations without the risk of immune-rejection.

In the last 15 years several approaches for the derivation of NCSCs from ESCs and iPSCs have been developed.

Stromal feeder-layers. In 2003, Mizuseki and colleagues have been the first ones to obtain NCSCs from mouse and primate ESCs through a co-culture with the mouse stromal line PA6 in the presence of BMP4 (Mizuseki et al., 2003). Moreover, different concentrations of BMP4 were able to further differentiate ESC-derived NCSCs towards sensory or autonomic neurons. Improvements of the protocol through the formation of ESC-derived non adhesive neurospheres increased the multipotent differentiation potential with the generation of sensory neurons and myofibroblasts both *in vitro* and *in vivo* (Jiang et al., 2009).

Neural rosettes. Co-culture of human ESCs with another stromal cell line (MS5) gives rise to neural rosettes able to self-renewal and to differentiate into bona fide $NGFR^+$ and $HNK1^+$ NCSCs (Perrier et al., 2004). As in the previous protocol, NCSCs are multipotent both i*n vitro* and *in vivo* when transplanted either into chick embryos or adult mice.

The main issues of these two approaches are the use of stromal feeders, whose secreted factors are undefined, and the time required for differentiation (20 to 40 days).

Dual-SMAD inhibition. Feeder cells can be avoided by using Small Mother Against Decapentaplegic (SMAD) signaling inhibitors, such as SB431542 and Noggin that block Activin/Nodal and BMP signaling pathways, respectively and prompt ESCs to differentiate toward neuroectoderm (Chambers et al., 2009). However, this protocol generates a culture that mostly consists of $Pax6^+$ neural precursor cells (NPCs) and only few $Pax6^-$ neural crest–like cells.

Given the relevant role played by Wnt signaling in NC induction at the neural plate border, Menendez and colleagues decided to combine SMAD inhibition to Wnt pathway activation using SB431542 and GSK3β inhibitor, respectively (Menendez et al., 2011, 2013). On average, this approach gives rise to 90-95% $NGFR^+$ and $HNK1^+$ NCSCs in 15 days, thus avoiding FACS sorting for population enrichment. iPSC/ESC-derived NCSCs can be propagated for more than 30 passages and can differentiate into SMCs, adipocytes, chondrocytes, osteocytes and peripheral neurons. However, it has not been demonstrated yet whether early stages of NC development are faithfully recapitulated through this experimental approach.

2.3.3.2. iPSC-derived NCSC applications

Applications of iPSC-derived NCSCs in disease modelling are still in its infancy and few reports are available now, some of which involve Familial Dysautonomia (FD). In particular, NCSCs derived from FD patient-iPSCs showed a migration defect that, in turn, impaired the survival rate of sensory, sympathetic and parasympathetic neurons (Lee et al., 2009; Zeltner et al., 2016). Kinetin, a plant hormone, was able to rescue the NCSC defect and defect and gave encouraging results when tested in patients (Axelrod et al., 2011).

iPSC-derived NCSCs were also obtained from patients with Hirschsprung's disease, for which standard treatment based on the surgical removal of the defective tract of the bowel

are often inefficient. Genetic analyses identified key mutations responsible for an impaired migration and differentiation of patient-specific ENCCs, that were restored through a CRISPR/Cas9-based approach (Lai et al., 2017).

Further potential applications involve corneal blindness, tooth regeneration and heart repair given the development of protocols for the differentiation of ESC- or iPSC-derived NCSCs respectively towards corneal endothelial cells (Chen et al., 2014), odontoblast cells (Seki et al., 2015) and SMCs or cardiac progenitors (Cheung et al., 2014; Tamura et al., 2016). NCSC-derived odontoblast and SMCs are particularly relevant in the context of WBS and 7dup syndrome given the dental and the aorta vasculature problems described in the patients (Pober, 2010). Jiao and collaborators recently obtained cardiac NC cells from BAV patient-iPSCs that showed defects in SMC formation and contraction and increased mTOR signalling pathway (Jiao et al., 2016). SMC defective differentiation was restored using rapamycin, a known mTOR signalling pathway inhibitor.

2.4. BAZ1B

A relevant number of observations in the past years suggest one of the genes contained in the WBSCR, known as BAZ1B or Williams syndrome transcription factor (WSTF), as a critical player in NCSC biology and, more specifically, in craniofacial development.

2.4.1. BAZ1B expression and role in development

BAZ1B was first identified and characterized in human tissues thanks to Xiaojun Lu and colleagues' studies in 1998 (Lu et al., 1998). Northern blot analysis revealed a 7.5 kb-transcript that was expressed in all the human (*i.e.*, skeletal muscle, brain, heart, ovary and placenta) and fetal tissues (*i.e.,* lung, brain, kidney and liver) tested (Lu et al., 1998).

Most of the knowledge on BAZ1B expression and role during development comes from studies performed in *Xenopus laevis* and mouse embryos.

2.4.1.1. BAZ1B role in Xenopus laevis

Baz1b can be already detected in *Xenopus laevis* unfertilized oocytes and is then expressed ubiquitously until 24 hours post-fertilization (stage 16) (Cus et al., 2006). During neurulation, it starts localizing to the closing neural tube and, in particular, at the level of the anterior brain, the optic cup and the migrating head NC cells (stage 20/21). The strong *baz1b* expression detected within the branchial arches is most probably a consequence of NC cells migration in these specific areas during development (see chapter 2.2.1).

Baz1b knock-down in *Xenopus laevis* embryos was shown to affect neural crest function, especially at the later stages of NC cell migration and maintenance (Barnett et al., 2012). Indeed, the expression of two important NC specifier genes, *Snail* and *Slug*, is not affected by *Baz1b* knock-down during the first stages of NC formation, but is clearly reduced within the branchial arches where NC cells migrate prior to the development of NC-related tissues. *Baz1b* knock-down was also responsible for an increase in *Bmp4* expression and a subsequent reduction of Sonic Hedgehog (Shh), both crucial in NC development and specification (Barnett et al., 2012; Liu and Cheung, 2016). A similar effect was observed when Imitation switch (*Iswi*), BAZ1B partner in the WICH complex (see chapter 2.4.3.1.), was knocked-down, thus supporting a potential repressive activity on BMP4 transcriptome by this complex (Dirscherl et al., 2005).

2.4.1.2. BAZ1B role in mouse

In mouse embryos, *Baz1b* expression is localized in the presumptive hindbrain, in the headfolds and in the caudal tail bud starting from 8.25 days postcoitum (dpc) and, later on, in the somites and in the forelimb bud (9.5 dpc) (Ashe et al., 2008). It is also detected

in the frontonasal process and in the mesenchyme of branchial arches 1 and 2, that give rise to the mandibular and maxillary components of the face.

Ashe and colleagues produced a mutant mouse bearing a point mutation in a highly conserved region in exon 7 that results in intermediate or reduced *Baz1b* protein levels in heterozygous or homozygous mice, respectively (Ashe et al., 2008). Homozygotes are smaller than wild type littermates and display craniofacial alterations resembling those of WBS patients such as bulbous and protruding foreheads, flattened nasal bone and shorter snouts. Heterozygotes also present skull defects including a decreased cranium width/height and length/height ratio and mandibular hypoplasia.

2.4.2. BAZ1B protein organization

BAZ1B gene consists of 20 exons spanning 80 kb and encodes a protein of 1425 amino-acids. BAZ1B protein is part of the BAZ/WAL family, that includes proteins with a specific subset of highly conserved motifs (**Figure 4a**).

Moving from N-terminus to C-terminus we can distinguish:

- a WSTF/Acf1/cbp146 (WAC) domain, that is not present in all BAZ proteins and has an intrinsic tyrosine kinase activity described in detail in chapter 2.4.3.2.;
- a DDT domain, possibly involved in DNA interaction since it has been found in different proteins known to bind DNA but lacking a known DNA-binding domain, such as Bromodomain and PHD domain Transcription Factor (BPTF) (Doerks et al., 2001). Moreover, a truncated version of BPTF that lacks the N-terminal region, including the DDT domain, displays a strong reduction in DNA-binding activity (Jordan-Sciutto et al., 1999);
- a Loop-Helix (LH) motif, shown to be relevant for transcriptional activation (Quong et al., 1993);

- BAZ 1 and BAZ 2 motifs, involved in the interaction with ISWI in B-WICH and WICH chromatin remodelling complexes (see chapter 2.4.3.1.);
- a WSTF/Acf1/KIAA0314/ZK783.4 (WAKZ) motif;
- a Plant HomeoDomain (PHD) finger. Proteins with PHD fingers domains were shown to play a role in transcription: for example, in BPTF, this domain specifically binds trimethylated lysine 4 on histone 3 (H3K4me3) at Transcription Start Sites (TSS) of active genes (Li et al., 2006). WSTF PHD finger domain specific targets have not been identified yet, but the presence of this domain in different subunits of other known chromatin remodelling complexes suggests a role in transcription;
- a bromodomain, composed by 110 amino acids with a central motif sequence (Jones et al., 2000). Bromodomains are implicated in protein-protein interactions and proteins containing this domain are often transcription cofactors and key components of chromatin remodelling complexes. For example, the ACF chromatin remodelling complex in Drosophila is composed by ISWI (also associated with BAZ1B in the WICH complex) and *BAZ1B* orthologous gene ATP-utilizing Chromatin assembly and remodeling Factor 1 (Acf1) that contains a bromodomain. ISWI itself is able to remodel chromatin but this activity is significantly increased by the presence of Acf1 in the complex (Ito et al., 1999). In addition, bromodomains are able to recognize acetylated lysine residues and, therefore, they serve to anchor bromodomain-containing proteins to acetylated chromatin (Jones et al., 2000).

2.4.3. BAZ1B functions

The conserved domains mentioned in the previous chapter support an important role for BAZ1B in chromatin remodeling, transcriptional regulation and in DNA damage response.

2.4.3.1. BAZ1B role in transcription and chromatin remodeling

BAZ1B is a common subunit of two well characterized ATP-dependent chromatin remodelers: WICH and B-WICH (**Figure 4b**).

ATP-dependent chromatin remodelers use the energy derived from ATP hydrolysis to assemble, reposition or displace nucleosomes in order to expose specific DNA sequences, therefore regulating fundamental processes including replication, transcription and repair (Vignali et al., 2000).

In particular, BAZ1B-containing complexes are part of the ISWI complexes that are able to expose longer regions of linker DNA by mobilizing nucleosome along the genome without displacing them (Vignali et al., 2000).

WICH complex. It was first isolated from *Xenopus laevis* egg extract and then shown to be conserved across vertebrates. In mammals, it is composed by BAZ1B and the ISWI isoform SNF2H, also known as SWI/SNF related, Matrix associated, Actin dependent Regulator of Chromatin, subfamily A, Member 5 (SMARCA5) (Bozhenok et al., 2002). This complex plays a crucial role in chromatin replication by mediating heterochromatin assembly through the establishment of a regular nucleosomal array from an irregular chromatin structure. In mouse cells, BAZ1B accumulates in heterochromatin around the centromere during replication and prevents aberrant heterochromatin spreading (Bozhenok et al., 2002; Poot et al., 2004). Indeed, BAZ1B depletion in different human cells is responsible for an increase in heterochromatin marks that leads to a global chromatin condensation, reduced accessibility and thus transcriptional impairment (Culver-Cochran and Chadwick, 2013; Poot et al., 2004, 2005). The same group observed that Proliferating Cell Nuclear Antigen (PCNA), a key player in DNA and chromatin replication, act as a binding platform that facilitates the targeting of WICH complex at replication foci (Poot et al., 2004).

B-WICH. BAZ1B and SNF2H/ISWI also interact with other nuclear proteins to form the B-WICH complex, that includes: nuclear myosin 1 (NM1); the proto-oncogene Dek; the Cockayne Syndrome protein B (CSB), also part of the SWI/SNF complex; the RNA

helicase Gu alpha; the nucleolar protein Myb-binding protein 1a (Myb-bp1a) and SAP155, involved in splicing (Cavellán et al., 2006).

Most of these proteins are implicated in transcription *per se* and chromatin remodeling-dependent transcriptional regulation. In particular, ChIP analysis in HeLa cells show that the core proteins (*i.e.*, BAZ1B, SNF2H and NM1) interact with *7SL RNA*, *5S rRNA* and *45S rRNA* genes, transcribed by RNA polymerase I and III, and that their levels are reduced upon BAZ1B silencing (Cavellán et al., 2006). Further studies from the same authors support the role of B-WICH complex in facilitating active transcription by RNA polymerase I and III (Sadeghifar et al., 2015; Vintermist et al., 2011). In particular, B-WICH allows the binding of specific regulatory factors to DNA through the maintenance of an open chromatin structure.

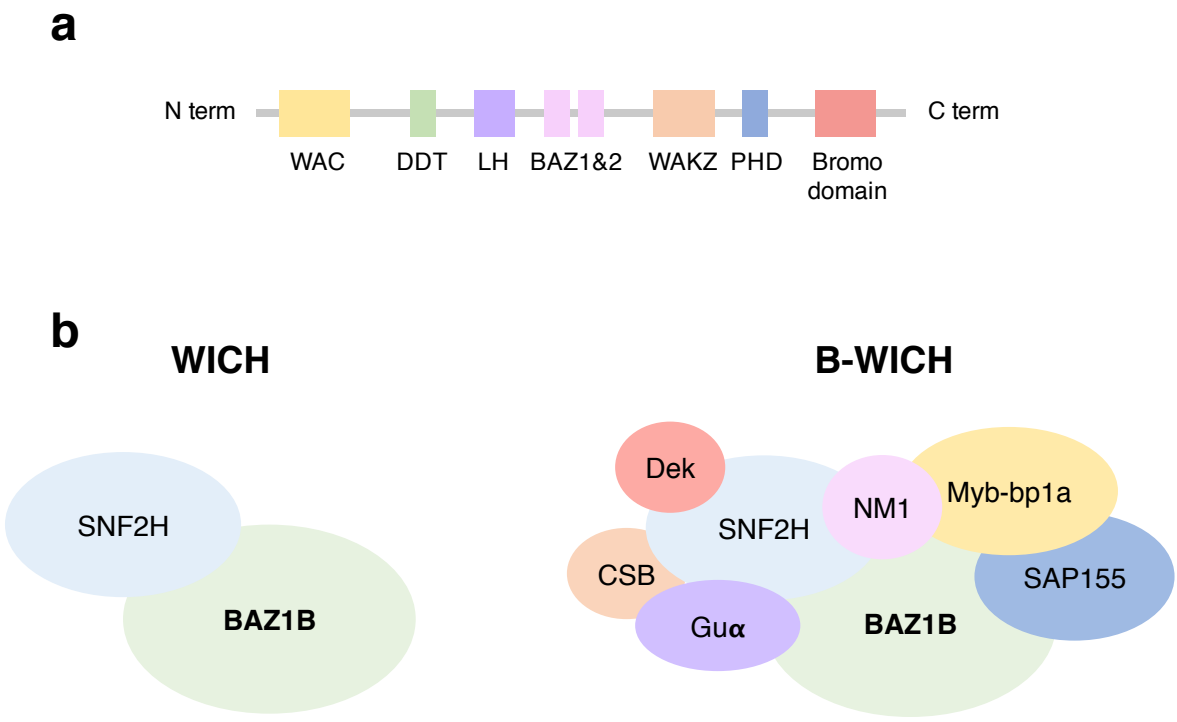


Figure 4. BAZ1B protein structure and BAZ1B-containing complexes

a) Schematic representation of conserved motifs in WSTF**; b**) BAZ1B can form a complex with SNF2H alone or in combination with other proteins involved in transcription. Redrawn from (Barnett and Krebs, 2011)

2.4.3.2. BAZ1B role in DNA damage response

In 2009, Xiao and colleagues identified in the N-terminal portion of BAZ1B an active kinase domain encompassing the WAC domain and a previously unknown C-motif, required for the phosphorylation of tyrosine 142 (Tyr142) on the C terminus of H2AX (Xiao et al., 2009). Tyr142 is constitutively phosphorylated in normal conditions and progressively dephosphorylated during DNA damage response, when instead phosphorylation on serine 139 (γ-H2AX) increases. *BAZ1B* knock-down is responsible for an impairment in γ-H2AX foci formation as a consequence of defects in the recruitment of Mdc1 and ATM, indicating a crucial role for BAZ1B in this process (Xiao et al., 2009). In particular, it has been hypothesised that BAZ1B acts during DNA damage by enhancing the recruitment of γ-H2AX kinase ATM. Thus, through its kinase activity on Tyr142, it should maintain the γ-H2AX phosphorylation response.

BAZ1B and other two members of the B-WICH complex, SNF2H and CSB, were later shown to be involved in another DNA repair process, namely transcription-coupled nucleotide excision repair that resolves DNA lesions affecting transcription (Aydin et al., 2014).

Few studies reported an impairment of ATR-dependent DNA damage response in WBS patient-derived cells (fibroblasts and lymphocytes), but a direct link with a specific WBSCR gene is still missing (Guenat et al., 2017; Savina et al., 2011).

2.4.4. BAZ1B in diseases

Little is known about the implication of BAZ1B in diseases such as cancer and neurodevelopmental disorders.

Recently, Meng and colleagues observed that *BAZ1B* overexpression promotes both proliferation and invasion of human lung cancer cells in a mouse xenograft model, thus accelerating tumor aggressiveness (Meng et al., 2016). One year later, Lundqvist and collaborators reported that *BAZ1B* expression is higher in human breast cancer cells than

in normal tissues and this was associated with a BAZ1B-dependent activation of estrogen receptor signaling (Lundqvist et al., 2017).

In 2016, it was shown that *BAZ1B* haploinsufficiency is responsible for transcriptional alterations and differentiation defects in neurons derived from WBS patient-iPSCs, letting the authors conclude that BAZ1B might contribute to the neurological phenotype (Lalli et al., 2016).

Other indications on BAZ1B contribution to WBS derive from a peculiar patient carrying a smaller deletion of the 7q11.23 region (around 1.2 Mb instead of 1.5 Mb) that does not involve *BAZ1B* and few other genes (Fusco et al., 2013). Language and communication skills are well preserved in the patient, in line with what is observed in WBS patients with the typical deletion. However, this patient does not present most of the typical WBS facial alterations, including the stellate iris, the short, upturned nose, the long philtrum, the wide mouth, the full lips, the dental problems and the malocclusion, suggesting that *BAZ1B* and/or the few other spared genes might be involved in the craniofacial phenotype. Notably, this patient was included in my study, as shown in the results chapter.

2.4.5. Chromatin remodeling in neural crest development

Many evidences support the relevant and direct involvement of BAZ1B in NCSCs formation and migration and in the consequent development of craniofacial structures (as demonstrated in *Xenopus laevis*, in a mouse model and in a peculiar case of WBS patient). However, the molecular mechanisms at the bases of BAZ1B action in NCSCs have not yet been elucidated. Several publications support its function as chromatin remodeler and transcriptional regulator in different cell types, including human cells. Thus, BAZ1B-dependent regulation of neural crest development might occur at the chromatin level.

Chromatin remodeling aimed at generating transcriptionally permissive chromatin state for the binding of specific TFs was shown to be extremely relevant for the development of many cell types including NC cells.

For example, the homozygous loss of the SWI/SNF chromatin remodeling complex BAF-A, induces severe heart defects and consequent embryonic lethality in mice (Chandler and Magnuson, 2016). In addition, the absence of ARID1A, a subunit of BAF-A complex, affects mouse early cardiac NC cell migration and is associated with craniofacial defects (Chandler and Magnuson, 2016).

Notably, the ATP-chromatin remodeler Chromodomain helicase DNA binding protein 7 (Chd7) was shown to play a crucial role in NC migration both in *Xenopus laevis* embryos and human NC-like cells (Bajpai et al., 2010). Indeed, in humans, *CHD7* mutations are responsible for CHARGE syndrome, a neurodevelopmental disorder characterized by craniofacial defects and other NC-related alterations (Pagon et al., 1981). Moreover, the authors observed that CHD7 interacts with BRG1 (the catalytic subunit of SWI/SNF chromatin remodeling complex) and that they both regulate the transcriptional activity of NC migration specifiers (*Sox9* and *Twist*) by preferentially co-occupying their enhancer elements (Bajpai et al., 2010).

Enhancers are short DNA sequences of around 200-500 base pairs (bp) that functionally interact with promoters to regulate gene expression (Buecker and Wysocka, 2012). Lineage-specific TFs recognize peculiar motifs on enhancer regions and bind them upon recruitment of coactivators, such as chromatin remodelers and histone modifiers, which, in turn, establish permissive chromatin states.

Recently, it was shown that chromatin states associated with an active enhancer signature (*i.e.*, H3K27ac and H3K4me1) are more dynamic, across different cell types, than promoter regions (Rada-Iglesias et al., 2012). In particular, around 80% of active enhancer regions identified in hESC-derived NC cells were unique to these cells, while only few were shared with hESCs or human neuroectodermal cells (hNECs). On the contrary, hESC-associated chromatin patterns such as bivalent promoters (in which both active H3K4me3 and inactive H3K27me3 marks coexist) and poised enhancers (H3K4me1 but not H3K27ac) are much less represented in human NC cells.

The investigation of the most represented DNA motifs at human NC cells enhancer regions revealed a clear association of distal region with TFs specifically involved in NC induction and differentiation and in craniofacial development including Transcription Factor AP-2-Alpha (TFAP2A) and Nuclear Receptor subfamily 2 group F members 1/2 (NR2F1/2) (Rada-Iglesias et al., 2012). In addition, these two TFs co-occupy a relevant number of sites characterized by an active enhancer signature, suggesting a role in the establishment of permissive chromatin states in NC cell enhancers.

It was also observed that single nucleotide polymorphisms (SNPs) on NR2F1/2 recognition motifs, affecting their binding to enhancer regions, are responsible for alterations in enhancer states that in most cases are subtle, but can have anyway negative effects on craniofacial development (Rada-Iglesias et al., 2013). Two years later, Prescott and collaborators identified clusters of enhancer regions bound by TFs involved in NC development and associated with facial variation (Prescott et al., 2015). Briefly, they were able to distinguish higher primate- and human-biased enhancers according to their association with genes whose higher or lower expression was reflected in primate or human typical facial traits, thus showing how critical gene dosage is for a proper species-specific craniofacial development.

2.5. Aim of the thesis

Taking advantage of WBS and 7dup patient specific-iPSC lines derived in our laboratory, in my PhD project I aimed at better understand the molecular alterations due to gene dosage imbalances at the 7q11.23 region and responsible for developmental defects in NC and NC derivatives in patients.

In order to achieve this aim:

1. I differentiated iPSCs derived from WBS and 7dup patients towards *bona fide* multipotent NCSCs;
2. I dissected BAZ1B specific contribution to NCSC dysregulation at both transcriptomic and chromatin levels in patient-derived NCSCs, given its involvement in NC development in animal models and, possibly, in a peculiar atypical WBS case study;
3. I evaluated possible BAZ1B-dependent cellular alterations (*i.e.,* migration defects) in patient-derived NCSCs in the perspective of the observed molecular dysregulation.

The main advantage of this experimental setup is the possibility to directly investigate NCSCs carrying the genetic information of WBS and 7dup patients, thus revealing a reliable scenario of patient-specific deficiencies.

3. MATERIAL AND METHODS

3.1. Human samples

Participation in this study by patients and their relatives along with skin biopsy donations and informed consent procedures were approved by the ethics committees of the Genomic and Genetic Disorder Biobank (Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy) and the University of Perugia (Azienda Ospedaliera–Universitaria 'Santa Maria della Misericordia', Perugia, Italy).

3.2. Fibroblast reprogramming and iPSC culture

WBS1, WBS2, WBS3, WBS4, 7dupASD1, AtWBS1 and CTL2 fibroblasts were reprogrammed using the mRNA Reprogramming kit (Stemgent, 00-0071), while the 7dupASD2 and CTL1R lines were reprogrammed with the microRNA Booster kit (Stemgent, 00-0073). The CTL3 line was reprogrammed by transfection with the STEMCCA polycistronic lentiviral vector followed by Cre-mediated excision of the integrated polycistron. 7dupASD3 and CTL4R were reprogrammed using the Simplicon™ RNA Reprogramming kit (Millipore, SCR550).

Prior to differentiation, iPSC lines were cultured on plates coated with hESC-qualified Matrigel (BD Biosciences, 354277) diluted 1:40 in DMEM (Lonza, BE12-614F) + F-12 (Life Technologies, 11765054) and grown in mTeSR-1 medium (StemCell Technologies, 85850). They are passaged upon treatment with Accutase (Sigma, A6964) and then plated in mTeSR-1 medium supplemented with 5 μM Y-27632 (Sigma, Y0503).

3.3. Differentiation

Differentiation into NCSCs, mesenchymal stem cells (MSCs), osteocytes and smooth muscle cells (SMCs) was performed as previously described (Cheung et al., 2014; Menendez et al., 2013).

1. NCSCs (15-20 days). 90% confluent iPSCs were detached with accutase and plated onto hESC-qualified Matrigel coated tissue culture plates in mTESR supplemented with 5 μM Y-27632 at a density of ~9.2 × 10^4 cells per cm^2. The day after neural crest medium (see below) was added and changed every day for 15-20 days. Cells were passaged every 4-5 days and plated at a high concentration (1:1 the first time and 1:2 the following ones) onto hESC-qualified Matrigel coated tissue culture plates for the entire duration of the differentiation.

   Upon differentiation, NCSCs were cultured in the same medium and passaged every 3-4 days.

NCSC medium: DMEM + F-12, 10% probumin Life Science Grade (20% (vol/vol) stock solution; Millipore, 821001), 1% Penicillin/Streptomycin (Life Technologies, 15140-122), 1% L-Glutamine (Life Technologies, 25030-024), 1% Non-essential amino acids (Lonza, BE13-114E), 0.1% 1000X trace elements complex (CA055-010, Gentaur Italy Srl), 0.2% 50 mM β-mercaptoethanol (Life Technologies, 31350-010), 10 ug/ml Transferrin, bovine (Holo form) (Life Technologies, 11107-018), 50 ug/ml (+)-Sodium L-ascorbate (Sigma, A4034), 10 ng/ml Heregulin-1 (Peprotech, 100-03), 200 ng/ml LONG®R^3 IGF-I (Sigma, 85580C), 8 ng/ml recombinant human FGF2-b (Thermo Fisher Scientific, PHG0021), 3 μM GSK3 inhibitor IX (BIO) (MedChem express, HY-10580) and 20 μm SB431542 (MedChem express, HY-10431).

2. MSCs (10 days). NCSCs were detached with accutase and plated at a density of 6.5 × 10^4 cells per cm^2 in MSC medium (see below) onto noncoated tissue culture

plates. Medium was changed every other day and cells were passaged with trypsin every 4-5 days.

MSC medium. DMEM + F-12, 1% Penicillin/Streptomycin, 1% L-Glutamine, 0.2% 50 mM β-mercaptoethanol and 10% Fetal Bovine Serum (FBS) (HyClone, SH30088.03).

3. Osteocytes (20 days). MSCs were detached with trypsin and plated at a density of 5×10^3 cells per cm^2 in STEMPRO® osteogenesis differentiation medium (Thermo Fisher Scientific, A10072-01) onto noncoated tissue culture plates and medium was changed every 3-4 days.

4. SMCs (12 days). NCSCs were detached with accutase and plated at a density of 2×10^4 cells per cm^2 in SMC differentiation medium (see below) onto noncoated tissue culture plates and half of the medium was changed every 2 days.

SMC medium: IMDM (Sigma, 13390) + F-12, 1% of Penicillin/Streptomycin, 15 ug/ml of Transferrin, bovine (Holo form), 0.05% human Insulin (Sigma, I9278), 0.004% monothioglycerol (Sigma, M6145), 10ng/ml recombinant human Platelet-Derived Growth Factor BB (PDGF-BB, PeproTech, 100-14B) and 2 ng/ml recombinant human Transforming Growth Factor β1 (TGF-β1, PeproTech, 100-21C).

All the cell types used in this study were maintained at low oxygen (3%) conditions.

3.4. Flow cytometry

NCSCs or MSCs were detached using accutase or trypsin, counted and 1×10^6 cells/experimental condition were fixed in 4% paraformaldehyde and then blocked in 10% Bovine Serum Albumin (BSA). Cells were incubated for 1 hour with primary antibodies conjugated to fluorophores (see **Table 4**). Analyses were performed on a FACSCalibur

instrument (BD Biosciences) and data were analyzed with FCS express software (Tree Star).

3.5. SMC and osteocytes staining

SMCs and osteocytes were fixed in 4% paraformaldehyde for 20 minutes and subsequently blocked in 10% FBS and 0.1% Triton X-100 for 30 minutes at room temperature. SMCs were incubated with the primary antibody overnight at 4°C and then with secondary antibodies for 1 hour at room temperature. Both primary and secondary antibodies were resuspended in PBS + 10% FBS. Cell nuclei were stained with Dapi (Sigma, D9542). Slides were mounted using Vectamount AQ Mounting Medium (Vector Laboratories, H-5501).

Osteocytes were incubated with alkaline phosphatase (AP) for 15-20 minutes in the dark at room temperature.

All images were acquired on an Olympus AX70 microscope.

The list of the antibodies used is reported in **Table 4**.

3.6. Lentiviral vector production and NCSC transfection

BAZ1B knock-down was performed using validated pLKO.1 TRC vector TRCN0000013338 (referred to as sh38) and TRCN0000013341 (referred to as sh41). A pLKO.1 TRC vector containing a scrambled short hairpin sequence was used as a negative control.

A pBOS H2B-GFP vector was used for the generation of NCSC reporter lines.

The envelope plasmid PMD2-VSV-G, the packaging plasmid psPAX2 and the plKO.1 plasmids were extracted using the Nucleobond Xtra Maxi kit (Macherey-Nagel, 700414-10). Second generation lentiviral vector were produced through calcium phosphate

transfection of human embryonic kidney (HEK293T) cells and ultracentrifugation. Briefly, $9x10^6$ HEK293T cells were plated in DMEM, 10% FBS, 1% penicillin/streptomycin and 1% L-Glutamine and incubated at 37°C, 5% O_2, 5% CO_2. The day after, medium was changed and the following mix was prepared.

Transfection mix for 2 x15 cm dishes

PMD2-VSV-G: 16 μg

psPAX2: 29,2 μg

plKO.1 sh / pBOS H2B-GFP: 45 μg

$CaCl_2$: 244 μl

H_2O: to 2 ml

The tube was put on a rotating wheel for 10 minutes and then the mix was splitted in two tubes (1 ml/tube).

DNA precipitate was obtained by drop wise addition of 1 ml 2X HBS solution (281 mM NaCl, 100 mM HEPES, 1.5 mM NA_2HPO_4 pH 7.12) to the previously prepared tube mix placed on a full speed vortex. The final mix was immediately added on HEK293T cell medium surface, the cells were incubated for 14 hours and then medium was replaced with 16 ml of fresh medium. 30 hours later the supernatant was 0.22 μm filtered and ultracentrifuged for 2 hours at 20°C and 20000 rpm in SW32Ti rotor (Optima L-60 preparative Ultracentrifuge; Beckman). Pellets were then resuspended in PBS and the vector was stored at -80°C.

$3\text{-}4x10^5$ NCSCs were infected upon splitting with 1-3 μl of lentiviral vector and plated in 12 multiwell plates. The cells that integrated the viral vector were selected by adding 1μg/ml puromycin to the medium for at least 7 days.

3.7. RNA extraction

RNA was extracted using the RNeasy Micro Plus kit (Qiagen, 74004) to optimize the extraction of messenger RNA (mRNA). Briefly, cells were resuspended in a denaturing buffer (RLT) plus β-mercaptoethanol to inactivate RNases and then lysed using a blunt 20-gauge needle. Then RNA specific binding to the spin column was obtained by adding 70% ethanol, while DNA was eliminated using DNase I. RNA was eluted in 20-30 μl of RNase free water and stored at -80°C.

Prior to RNA library preparation, the concentration and the quality of RNA samples were assessed using a NanoDrop spectrophotometer (NanoDrop Technologies) and an Agilent 2100 Bioanalyzer (with the RNA 6000 Nano kit, 5067-1511), respectively.

3.8. RNA retrotranscription and Real Time quantitative PCR (RT-qPCR)

Retrotranscribed cDNA was obtained from 0.5-1 μg of total RNA using the SuperScript VILO Retrotranscription kit (Thermo Fisher Scientific, 11754-050).

RT-qPCR was performed on a 7500 Fast Real-Time PCR system (Applied Biosystems) using SYBR Green Master Mix (Applied Biosystems, 4334973) as detecting reagent and 10 μM for each primer pair. A total cDNA amount corresponding to 15 ng of starting RNA was used for each reaction. Each sample was analyzed in triplicate and normalized to *GAPDH*. Relative mRNA quantity was calculated by the comparative cycle threshold (Ct) method using the formula $2^{-\Delta Ct}$.

qPCR steps

1. 20 seconds at 95°C
2. 40 cycles of: 3 seconds at 95°C, 30 seconds at 60°C and 15 seconds at 95°C

3. Dissociation stage: 1 minute at 60°C, 15 seconds at 95°C and 15 seconds at 60°C

Oligo pairs used for *BAZ1B*: CCTCGCAGTAAGAAAGCAAAC (forward); ACTCATCCAGCTCCTTTTGAC (reverse).

Oligo pairs used for *GAPDH*: GCACCGTCAAGGCTGAGAAC (forward); AGGGATCTCGCTCCTGGAA (reverse).

3.9. RNA-seq libraries preparation

Library preparation for RNA sequencing was performed according to TruSeq Total RNA sample preparation protocol (Illumina, RS-122-2202), starting from 250 ng - 1 μg of total RNA.

First ribosomal RNA was depleted using Ribo-Zero rRNA magnetic beads and rRNA specific biotinylated primers. Next, mRNA was fragmented using divalent cations at 94°C and retrotranscribed with random primers, while the second cDNA strand was synthetized using DNA polymerase I and RNase H. Double strand cDNA was purified using AMPure XP beads (Beckman Coulter, A63881). Overhangs resulting from fragmentation were converted into blunt ends using an End Repair mix and then a single adenine was added to the cDNA 3' ends and specific adapters were ligated using their 3' overhang thymine. In the end, ligated cDNA was enriched and amplified using adapter specific primers.

cDNA library quality was assessed at an Agilent 2100 Bioanalyzer, using the High sensitivity DNA kit (5067-4626).

Libraries were sequenced with the Illumina HiSeq machine at a read length of 50 bp paired-end and a coverage of 35 million of reads per sample.

3.10. Protein extraction and Western blot

NCSCs were lysed in RIPA buffer (10 mM Tris, pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl, 1 mM EDTA) supplemented with protease inhibitor cocktail (Sigma, P8340) and 0.5 Mm Phenylmethylsulfonyl fluoride (PMSF) (Sigma, P7626) for 1 hour at 4°C. Protein quantification was performed using the Bradford protein assay (Bio-Rad) as follows: 200 μl of dye reagent, 800 μl of ddH_2O and 1 μl of protein extract. The standard curve was derived using BSA (NEB).

Protein extracts (30-50 μg per sample) were supplemented with NuPAGE LDS sample buffer (Thermo Fisher Scientific, NP0008) and 50 mM Dithiothreitol (DTT) (Thermo Fisher Scientific, NP0007) and denaturated at 95°C for 3 minutes. Then extracts were run on a precast NuPAGE 4–12% Bis-Tris Gel (Thermo Fisher Scientific, NP0322BOX) in NuPAGE MOPS SDS Running Buffer (Thermo Fisher Scientific, NP0001) and transferred to a 0.45 μm nitrocellulose membrane (GE Healthcare, GE10600002) for one hour at 100 V in a buffer containing 20% absolute ethanol and 10% 0.25 M Tris base, 1.9 M glycine. The membranes were blocked in TBST (50 mM Tris, pH 7.5, 150 mM NaCl and 0.1% Tween-20) and 5% milk for one hour, incubated with primary antibodies overnight (O/N) at 4°C and with secondary antibodies for one hour at room temperature (RT). Primary and secondary antibodies (see **Table 4**) were also diluted in TBST and 5% milk. Blots were detected with the ECL Prime Western Blotting detection reagents (Sigma, GERPN2236) and scanned using ChemiDoc system (Biorad).

3.11. Chromatin immunoprecipitation coupled with sequencing (ChIP-seq)

I used ~2 x 10^5 cells for each immunoprecipitation. NCSCs were fixed with PBS containing 1% formaldehyde (Sigma, F8775) for 10 minutes to cross-link proteins and

DNA and the reaction was then stopped by adding 125 mM glycine for 5 mintues. Cells were lysed with SDS buffer containing 100 mM NaCl, 50 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 10% SDS and sonicated using the S220 Focused-ultrasonicator (Covaris) to generate <300 bp DNA fragments. Sonication conditions were: 105 peak power, 5.0 duty factor, 200 cycles/burst for 5 minutes. ~1/20 of sonicated chromatin was decrosslinked in 9 volumes of decrosslinking buffer (1% SDS and 100 mM $NaHCO_3$) at 65 °C for one hour to check for the correct fragment size. Sonicated chromatin was diluted fivefold with IP buffer containing 1 volume of SDS buffer and 0.5 volume of Triton dilution buffer (100 mM Tris-HCl pH 8.5, 5 mM EDTA pH 8.0, 5% Triton X-100), incubated O/N at 4°C with primary antibodies (see **Table 4**) and then 3 hours with Dynabeads Protein G (Thermo Fisher Scientific, 10004D). Beads were washed three times with low-salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0 and 150 mM NaCl) and once with high-salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0 and 500 mM NaCl). Immunocomplexes were eluted in decrosslinking buffer at 65 °C for 2 hours and DNA was purified using QIAquick PCR columns (Qiagen, 28104) and quantified with Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Q32851). DNA libraries were prepared by the sequencing facility at IFOM/IEO campus according to the protocol described by Blecher-Gonen and colleagues (Blecher-Gonen et al., 2013) and DNA was sequenced on the Illumina HiSeq 2000 platform.

3.12. Migration assays

Scratch assay. 5-7x10^4 cells were plated in each of the two matrigel-coated wells of silicone culture-inserts (Ibidi, 80209) attached to 6 well-culture plates. After 24 hours the insert was removed, medium was changed to remove dead cells and time lapse was performed for 20 hours at the rate of one image every 10 minutes at a 10x magnification. For each condition, I considered 5 independent fields.

Random migration assay. 7.5x10^4 cells were sparsely plated in 6 well-culture plates and

24 hours later the time lapse experiment started. Images were acquired every 5 minutes at a 10x magnification using the same microscope. For each condition, 10 independent fields were considered. The analysis was performed by manually tracking five cells per each field for 6 hours (for a total number of 50 cells/condition) using the 'Chemotaxis tool' plugin from Fiji and statistical significance was evaluated using Student's *t*-test.

For both migration assays, images were acquired with the BX61 upright microscope equipped with a motorized stage from Olympus.

3.13. BAZ1B complex isolation

Proteins were extracted as described above. 2-3 mg of total lysate were immunoprecipitated per each condition. Protein lysates were incubated O/N at 4°C with BAZ1B antibodies (see **Table 4**) and then 3 hours with Dynabeads Protein G. Beads were washed twice with low-salt RIPA buffer (140 mM NaCl) and twice with high-salt RIPA buffer (250 mM NaCl). Elution was performed at 95°C for 3 minutes in one volume of 4x Novex LDS sample buffer + 50 mM DTT and 9/10 of the eluate were run on a precast NuPAGE 4–12% Bis-Tris Gel and was stained with InstantBlue protein stain (Expedeon, ISB1L). 1/10 of the IP was subjected to Western Blot analysis as a control.

3.14. Liquid chromatography–tandem MS (LC–MS/MS) analysis and protein identification

IP sample processing through LC-MS/MS was performed by the Mass Spectrometry Unit at IFOM/IEO campus.

Briefly, bands of interest were cut from gels and trypsinized as described by Shevchenko and colleagues (Shevchenko et al., 1996). Peptides were desalted, dried in a Speed- Vac and resuspended in 10 μL of solvent A (2 % ACN, 0.1% formic acid). 3 µL were injected

on a quadrupole Orbitrap Q-Exactive mass spectrometer (Thermo Scientific) coupled with an UHPLC Easy-nLC 1000 (Thermo Scientific) with a 25 cm fused-silica emitter of 75 µm inner diameter. Columns were packed in-house with ReproSil-Pur C18-AQ beads (Dr. Maisch Gmbh, Ammerbuch, Germany), 1.9 µm of diameter, using a high-pressure bomb loader (Proxeon, Odense, Denmark).

Peptides separation was achieved with a linear gradient from 95% solvent A (2 % ACN, 0.1% formic acid) to 40% solvent B (80% acetonitrile, 0.1% formic acid) over 30 minutes and from 40% to 100% solvent B in 2 minutes at a constant flow rate of 0.25 µL/min, with a single run time of 33 minutes in case of single bands.

MS data were acquired using a data-dependent top 12 method, the survey full scan MS spectra (300–1750 Th) were acquired in the Orbitrap with 70000 resolution, AGC target 1e6, IT 120 ms. For HCD spectra resolution was set to 35000, AGC target 1e5, IT 120 ms; normalized collision energy 25% and isolation width of 3.0 m/z.

For protein identification, the raw data were processed using Proteome Discoverer (version 1.4.0.288, Thermo Fischer Scientific). MS^2 spectra were searched with Mascot engine against uniprot_human_201503 database (89909 sequences; 35686673 residues), with the following parameters: enzyme Trypsin, maximum missed cleavage 2, fixed modification carbamidomethylation (C), variable modification oxidation (M) and protein N-terminal acetylation, peptide tolerance 10 ppm, MS/MS tolerance 20 mmu. Peptide Spectral Matches (PSM) were filtered using percolator based on q-values at a 0.01 FDR (high confidence). Scaffold (version Scaffold_4.3.4, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications.

3.15. NCSC Microarray

NCSC microarrays were prepared by the internal facility at IFOM/IEO campus.

Briefly, biotin-labeled cDNA targets were synthesized starting from 150 ng of total RNA. Synthesis of double-stranded cDNA and related cRNA was performed with the Ambion

WT Expression kit. The same kit was used to synthesize the sense-strand cDNA to be fragmented and labeled with the Affymetrix GeneChip WT Terminal Labeling kit. Hybridization was performed using the GeneAtlas Hybridization, Wash and Stain kit, according to the manufacturer's instructions. The array strips were imaged using the GeneAtlas Imaging Station. Affymetrix GeneAtlas Command Console software was used to acquire the array strip images and generate DAT and CEL files. Microarray analysis was performed with the Bioconductor Affy package using the brainarray RefSeq annotation v18.0.0, robust multi-array average (RMA) for background normalization and quantile normalization between samples. Quantification was obtained using perfectly matching probes only with median polish summarization, averaging probe sets to obtain gene-level expression. Probe sets not assigned to known genes or having log2 fold changes of <0.5 were discarded, and differential expression was assessed by two-tailed t test. For the purpose of enrichment analyses, DEGs with FDR < 0.2 were considered.

3.16. RNA-seq analysis

RNA-seq data were quantified using our standard pipeline based on Salmon 0.80 to quantify reads and calculate read-counts and transcripts per million (TPMs) in a transcript- and gene-wise fashion, using the quasi-mapping off-line algorithm (Germain et al., 2016). EdgeR was used to test for differential gene expression, using generalized linear regression methods (GLMRT), to identify pattern of differential expression following four different schemes:

1. a factorial analysis based on the definition of one group of scrambled and one group of knock-down samples to identify genes dysregulated similarly across short-hairpins characterized by different efficiencies (chapter 4.3.3.1.);
2. a multifactorial analysis based on the definition of three groups separating the two short-hairpins and the scrambled samples (chapter 4.3.3.2.);

3. an analysis based on the definition of an independent variable recapitulating the degree of average efficiency of the two short-hairpins, following qRT-PCR measurements of BAZ1B levels in all samples ending with a coefficient of 1 identifying scrambled, 0.6 identifying sh38 and 0.3 identifying sh41 samples (chapter 4.3.3.3.);
4. a further numerical analysis in which BAZ1B levels as quantified by RNA-seq were used as the independent variable (chapter 4.3.3.4.).

All these analyses were performed correcting by individual, in order to take into account the genetic background of each individual. Genes were identified and characterized by an absolute fold-change (FC) of 1.1 and an FDR lower than 0.05.

Gene Onthology (GO) enrichments were performed using topGO R package version 2.28.0, on the list of genes derived from the fourth Differential Expression Analysis (DEA) based on the linear regression following BAZ1B levels, and characterized by a FC > 1.1 and FDR < 0.1 in order to expand the list of genes, allowing the risk of including some false positives, but then selecting only enrichments with an FDR < 0.05.

3.17. ChIP-seq analysis

ChIP-seq experiments were analyzed both qualitatively and quantitatively. Reads were trimmed with FastX toolkit (-Q33 -t 20 -l 22), aligned with Bowtie 1.0 (-v 2 -m 1) on the Human hg38 genome and peaks were called by means of MACS 2.1.1. H3K4me1 and H3K27ac peaks were called with -m 4 50 and a q-value threshold equal to 0.05 (-q 0.05). H3K27me3 peaks were called as "broad" using default parameters.

Whole genome regions

Qualitative analysis was performed in all the 21 samples using BedTools version 2.16.1, while quantitative analysis was obtained measuring differential deposition of the H3K27ac mark with edgeR, using GLMRT, testing for BAZ1B levels while blocking for individuals as previously done on the RNA-seq data.

Association of differentially H3K27 acetylated regions with proximal or distal regulatory regions was obtained using the ChIPseeker package.

Active enhancer regions

Qualitative analysis was performed as above but selecting two different set of regions in the two control non-infected lines:

1. putative active enhancer regions, characterized by H3K4me1 and H3K27ac peaks;
2. putative silenced regions, exclusively characterized by H3K27me3.

Quantitative analysis was performed as above but focusing on the putative active enhancers regions, and measuring differential deposition of each of the three marks. Association of chromosomic regions to genes was done using GREAT with the "basal plus extension" scheme (McLean et al., 2010).

Target	***Application***	***Company***	***Source***	***Catalog #***
HNK1-FITC	FACS	BD Biosciences	Mouse	555619
NGFR-ALEXA647	FACS	BD Biosciences	Mouse	560877
CD44-APC	FACS	EBIOS	Mouse	17-0441-82
CD73-PE	FACS	BD Biosciences	Mouse	550257
SMA	ICC	Sigma	Mouse	A2547
CY3 anti mouse	ICC	Jackson IR	Goat	115-165-146
BAZ1B	WB	Abcam	Rabbit	ab50632
BAZ1B	IP	SantaCruz	Mouse	sc-81426
BAZ1B	IP	Abcam	Rabbit	ab50850
GAPDH	WB	Abcam	Mouse	ab8245
H3K27ac	ChIP	Abcam	Rabbit	ab4729
H3K27me3	ChIP	Cell Signalling	Rabbit	9733BF
H3K4me1	ChIP	Abcam	Rabbit	ab8895

Table 4. List of antibodies used in this study

NCSC lines	MSC differentiation	Osteocytes differentiation	SMC differentation	sh38	sh41	KD validation	RNAseq	ChIPseq	Scratch assay	Random migration	IP and MS
WBS1-C1			x	x	x	x	x	x			x
WBS2-C2				x	x	x	x				x
WBS4-C1	x	x		x	x	x	x	x	x	x	
AtWBS1-C2	x	x	x	x	x	x	x	x			
CTL1R-C3	x	x		x	x	x	x	x	x		x
CTL2-C2				x	x	x	x		x	x	
CTL3-C1					x	x	x (only scr and sh41)				
CTL4R-C1			x	x	x	x	x	x	x		
7dupASD1-C2				x	x	x	x		x	x	x
7dupASD2-C1	x	x		x	x	x	x	x			
7dupASD3-C1			x	x	x	x	x (only scr and sh38)	x			

Table 5. List of experiments performed in each NCSC line

4. RESULTS

4.1. Generation of NCSC lines from patient specific-iPSCs

In order to model *in vitro* patient specific craniofacial alterations, we took advantage of the large cohort of WBS and 7dup iPSC lines generated in our laboratory and we differentiated them to NCSCs (Adamo et al., 2015). As mentioned before, NC is a temporary lineage that generates a relevant number of tissues whose development is clearly affected in the two syndromes.

The protocol chosen for the iPSC to NCSC differentiation (Menendez et al., 2013) allows the generation of NCSCs both from human ESCs and iPSCs in about two weeks. NCSCs are obtained through the combination of SMAD pathway inhibition (using a TGF-β inhibitor known as SB431542) to obtain a neuroectodermal population (Chambers et al., 2009) and Wnt pathway activation (using a glycogen synthase kinase 3 inhibitor, GSK3i) required for NC induction at the neural plate border (Patthey et al., 2008).

The main advantages of this protocol are: i) the generation of a highly homogenous NCSC population (more than 90% of cells in culture are *bona fide* NCSCs) and ii) the absence of feeder cell layers that might interfere with the transcriptomic and epigenomic analyses.

This protocol was adapted to our iPSC lines with minor changes in the very first steps. In particular, I observed that leaving the cells growing till confluence and splitting them starting from day 4-5 of differentiation at a very high density (a 1:1 dilution the first time and a 1:2 dilution the following ones), significantly increases differentiation efficiency and NC induction.

A total number of 11 iPSC lines was selected for differentiation, including at least three lines from different individuals per each genetic condition (**Figure 5**). In particular, the NCSC cohort consisted of:

- Four control lines (CTL2-C2, CTL3-C1, CTL1R-C3 and CTL4R-C1), represented in grey.
- Three lines carrying 7q11.23 typical deletion (WBS1-C1, WBS2-C2, WBS4-C1), in red.
- One line carrying a 7q11.23 atypical deletion (AtWBS1-C2), in yellow, that spares *BAZ1B* and eight other genes at the centromeric region (Fusco et al., 2013). This patient is particularly relevant because it presents milder craniofacial alterations when compared to typical WBS patients, supporting a role for *BAZ1B* (and/or the few other spared genes) in NC and facial development.
- Three lines carrying 7q11.23 duplication (7dupASD1-C2, 7dupASD2-C1, 7dupASD3-C1), in blue.

CTL2-C2 and CTL3-C1 were derived from unrelated unaffected individuals, while CTL1R-C3 and CTL4R-C1 were derived from the mothers of WBS1-C1 and 7dupASD3-C1, respectively and, thus, they can be considered half matched controls.

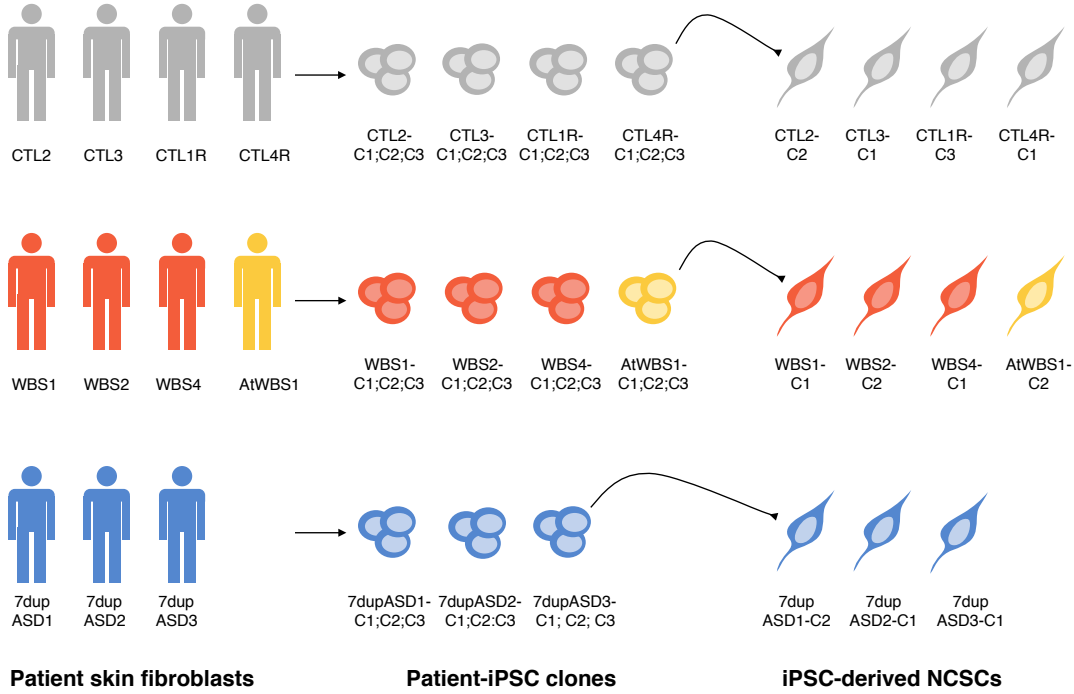


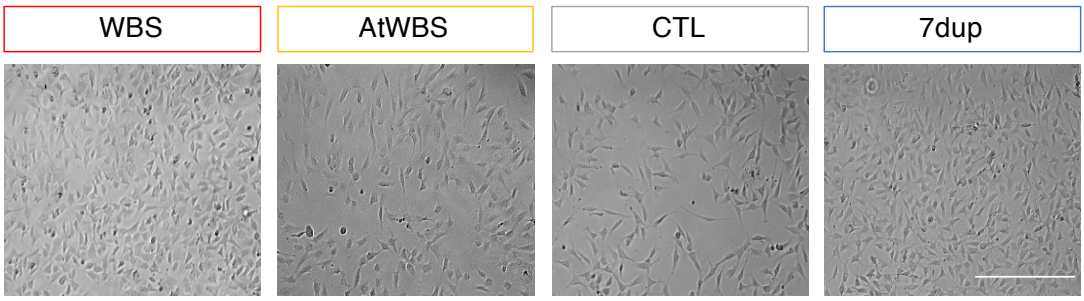
Figure 5. Strategy for the derivation of NCSC lines from patient-iPSC clones

Skin fibroblasts were collected from patient and reprogrammed to iPSCs. From each patient, at least three iPSC clones were generated and one clone per patient was chosen for the differentiation towards NCSCs.

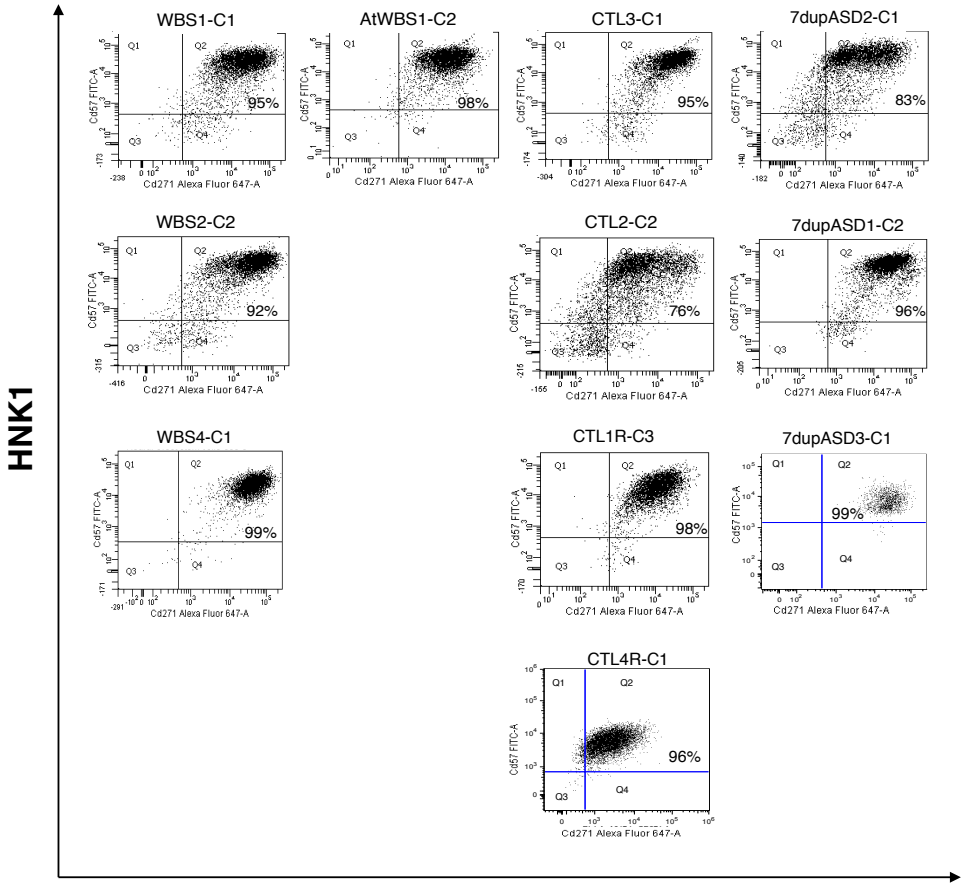
After 15 days of differentiation the cells displayed a clear morphology that was virtually identical in all four genetic conditions (controls, WBS, atypical WBS and 7dup) as shown in **Figure 6a**.

To confirm NC identity cells were stained for two markers widely used for the identification of NCSCs: HNK1 and NGFR (also known as p75). Flow cytometry analysis indicated a high percentage of double positive cells in all the four genotypes (**Figure 6b**) (Adamo et al., 2015).

a


WBS
AtWBS
CTL
7dup

b


WBS1-C1
95%
AtWBS1-C2
98%
CTL3-C1
95%
7dupASD2-C1
83%
WBS2-C2
92%
CTL2-C2
76%
7dupASD1-C2
96%
WBS4-C1
99%
CTL1R-C3
98%
7dupASD3-C1
99%
CTL4R-C1
96%
HNK1
NGFR
Cd57 FITC-A
Cd271 Alexa Fluor 647-A
Q1
Q2
Q3
Q4

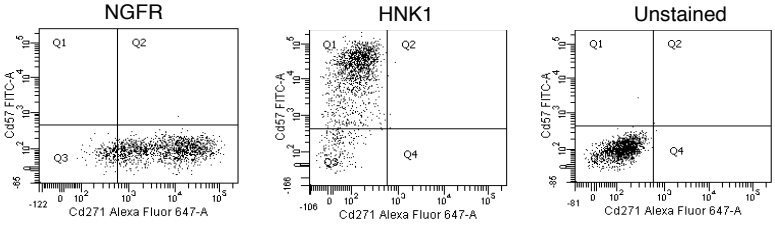

NGFR
HNK1
Unstained
Cd57 FITC-A
Cd271 Alexa Fluor 647-A
Q1
Q2
Q3
Q4

Figure 6. Characterization of iPSC-derived NCSCs

a) Phase-contrast images representing NCSCs of the four genotypes. The scale bar represents 400 µm. **b)** Flow cytometry analysis of HNK+/NGFR+ NCSCs. HNK1 and NGFR single stain and unstained cells are used as negative controls. Adapted from (Adamo et al., 2015).

Together with the validation of the expression of specific NCSC markers I have also investigated whether patient-specific NCSCs were multipotent, another key trait of this cell population. To test this NCSC feature, I further differentiated iPSC-derived NCSCs towards known NC derivatives, including Mesenchymal Stem Cells (MSCs), osteocytes and Smooth Muscle Cells (SMCs) (**Figure 7**).

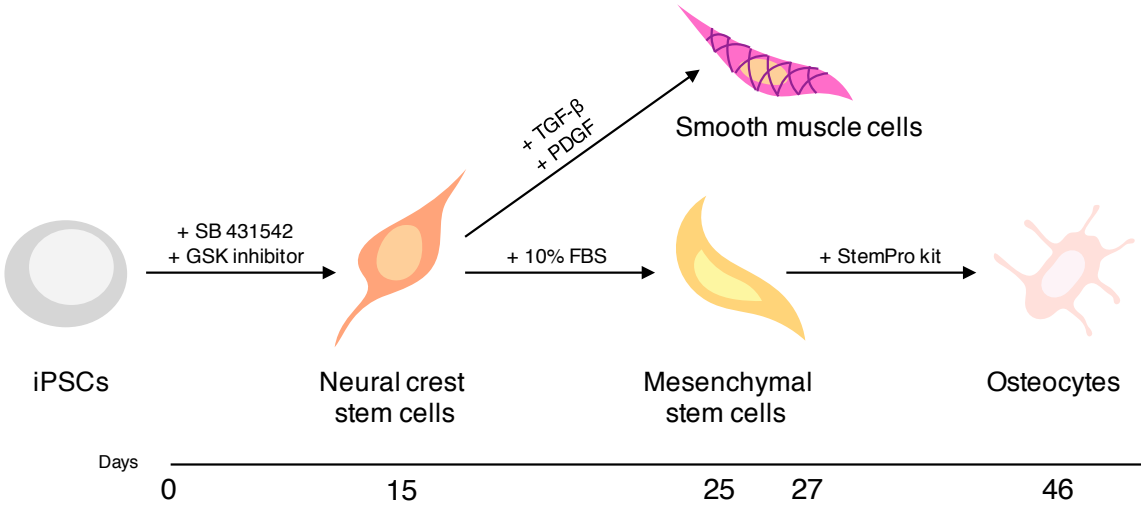


Figure 7. Schematic representation of the experimental system

First, patient-iPSCs are differentiated into NCSCs using SB 431542 and GSKi. Then NCSCs can be either differentiated towards MSCs and osteocytes or towards SMCs through TGF-β and PDGF administration.

MSCs were obtained by culturing NCSC in the presence of FBS for 12 days (**Figure 8a**) and then characterized for the expression of two mesenchymal markers, CD44 and CD73.

This protocol led to the generation of ~90-95% of CD44+/CD73+ cells on average (**Figure 8b**) (Adamo et al., 2015).

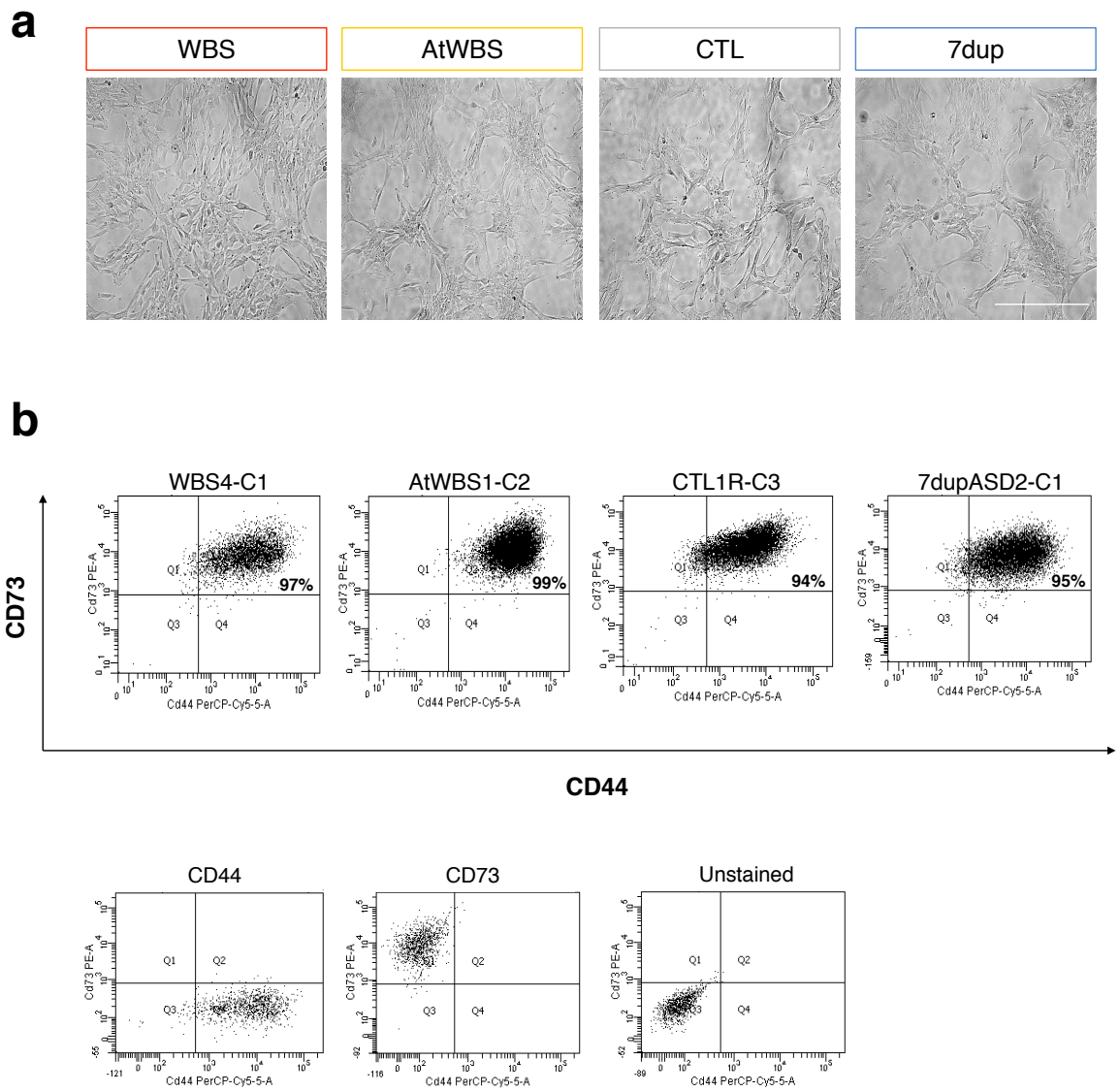


Figure 8. Characterization of NCSC-derived MSCs at day 10 of differentiation

a) Phase-contrast images representing MSCs of the four genotypes. The scale bar represents 400 µm. **b)** Flow cytometry analysis of CD73+/CD44+ MSCs. CD44 and CD73 single stain and unstained cells are used as negative controls. Adapted from (Adamo et al., 2015).

Osteocytes were then derived from MSCs **(Figure 9a**) using a ready-to-use kit from Thermo Fisher (StemPro Osteogenesis Differentiation kit) and stained homogeneously positive for alkaline phosphatase (ALP) after 20 days of differentiation (**Figure 9b**). ALP is

commonly used to stain osteocytes and other bone-related cell types together with Alizarin Red S.

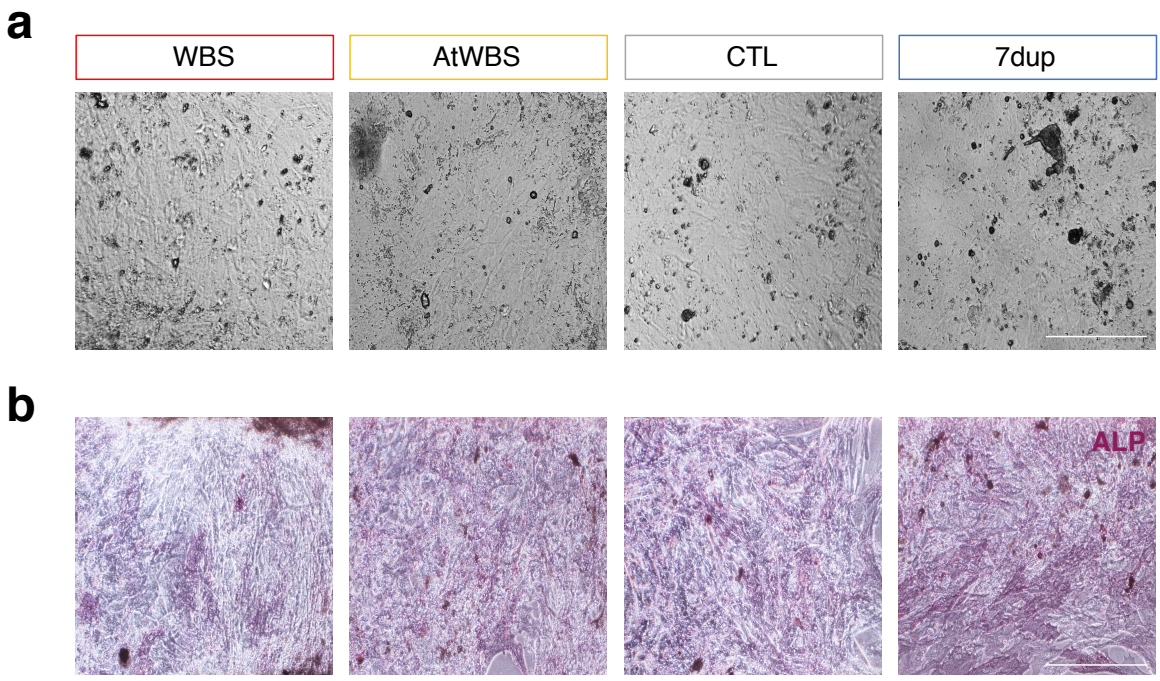


Figure 9. Characterization of MSC-derived osteocytes at day 20 of differentiation

a) Phase-contrast images representing osteocytes of the four genotypes. **b)** Alkaline phosphatase (ALP) staining of MSC-derived osteocytes. The scale bar represents 200 µm.

Finally, NCSCs were differentiated towards SMCs taking advantage of a protocol published by C. Cheung and colleagues (Cheung et al., 2014) and based on the presence in culture of PDGF and TGF-β, important for the development of SMC in vessels (Jain, 2003). The obtained SMCs were stellate shaped and expressed α-smooth muscle actin (αSMA) already 10 days after the beginning of the differentiation (**Figure 10a, b**). αSMA expression is relatively limited to vascular SMCs.

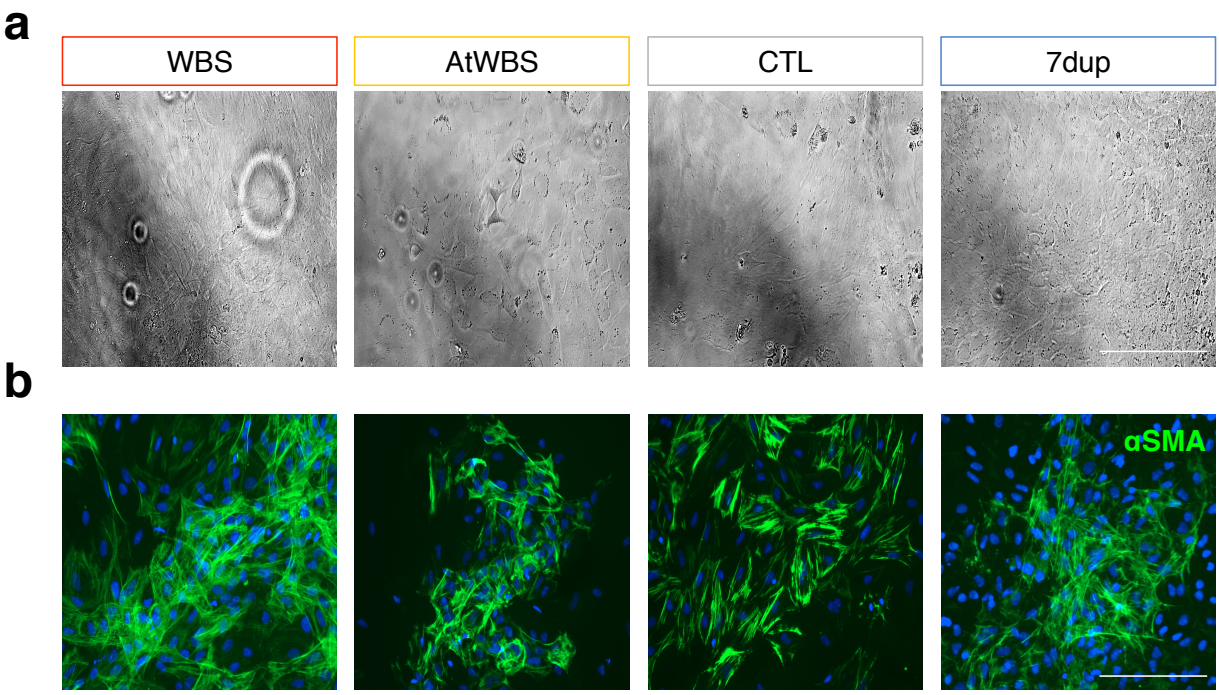


Figure 10. Characterization of NCSC-derived SMCs at day 10 of differentiation

a) Phase-contrast images representing SMCs of the four genotypes. **b)** NCSC-derived SMCs stained for αSMA. The scale bar represents 200 μm.

Given the outcome of the differentiation experiments, NCSCs derived from patients and control iPSCs can be reliably considered multipotent.

4.2. Investigation of iPSC derived-NCSC transcriptional profile

Our transcriptome analysis of 27 iPSC lines derived from healthy controls, WBS and 7dup patients, identified 757 Differentially Expressed Genes (DEGs) that, already at this very initial stage of development, were specifically enriched for GO categories encompassing a plethora of developmental pathways relevant to the specific traits defining WBS and 7dup syndromes (see chapter 2.3.3.) (Adamo et al., 2015).

Then, in order to evaluate the consequences, at later stages of development, of the transcriptional dysregulation observed at the pluripotent state, we performed microarray analysis of NCSC transcriptomes collected three days after the induction of NCSC differentiation. As observed in iPSCs, also in NCSCs RNA levels of WBSCR genes mirrored gene dosage (**Figure 11a**). Interestingly, *BAZ1B* levels in the atypical WBS sample (in which *BAZ1B* is not deleted) are comparable to the ones measured in the control samples (**Figure 11b**).

Differential expression analysis detected transcriptional alterations in genes and biological processes that were more lineage-specific that the ones identified in iPSC transcriptional profiling and potentially related to NC-dependent disease phenotypes. In particular, the 364 DEGs (with a False Discovery Rate, FDR<0.2) identified in NCSCs included genes involved in: i) craniofacial development (*ATP2C1, HHAT, LMNB1, MAPK8, PTCH1* and *SATB2*); ii) cell adhesion and membrane assembly (*APLP1, BMP7, PTPRD* and *NLGN3*), relevant for the migratory properties of NCSCs; iii) axon extension (LIMK1, SLIT3 and NLGN3) critical for the development of NC-derived peripheral neurons and iv) RhoA signaling pathway (*WASF1*, *GFAP*, *ACTR2*, *STMN1*, *MAPK8*, *ARHGEF11* and *PLXNA1)*, that was shown to play an important role in the development of SVAS (Adamo et al., 2015; Ge et al., 2012) (GO enrichment in **Figure 11c**).

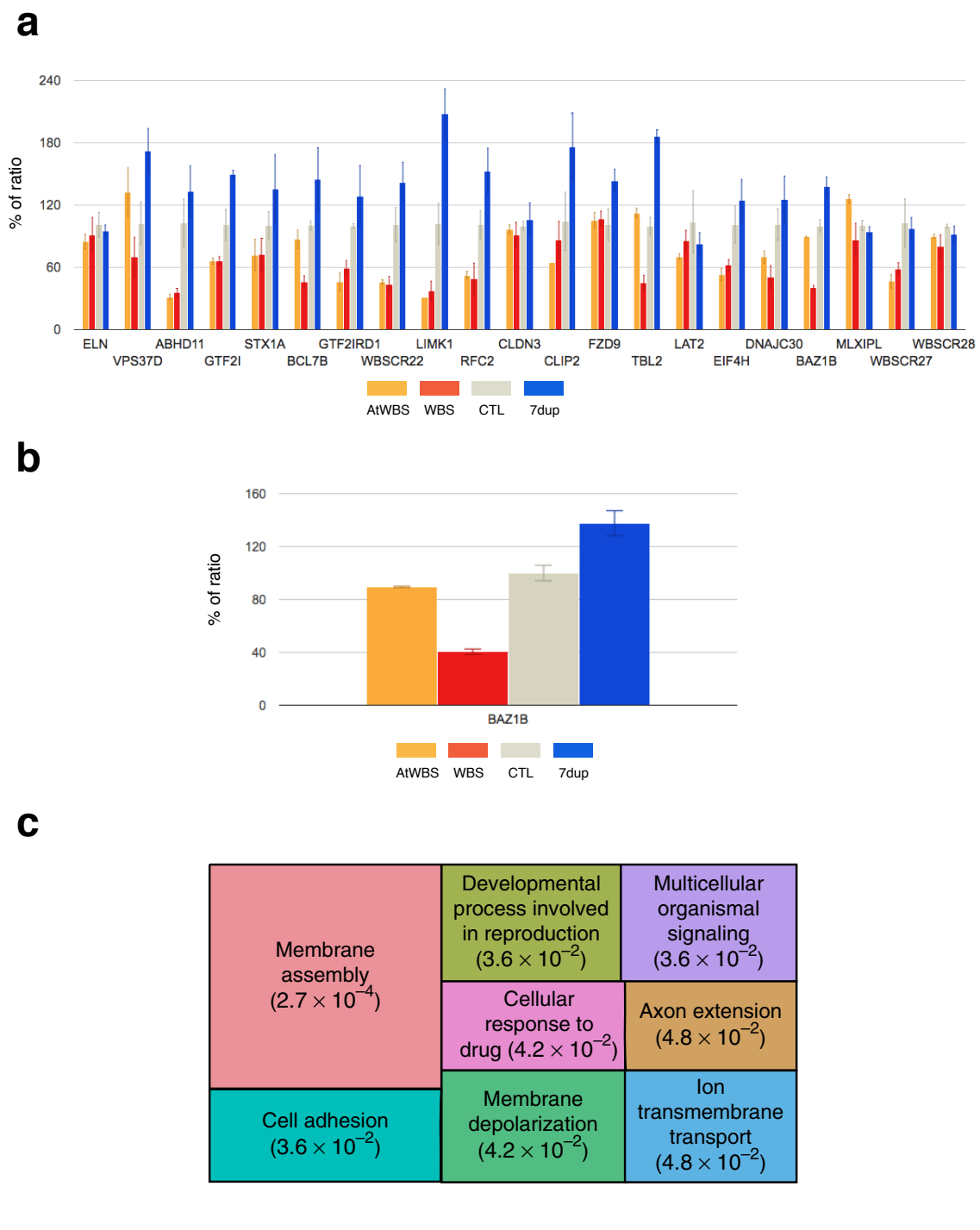


Figure 11. Transcriptional profile of iPSC-derived NCSCS

a) Expression levels of WBSCR genes in the microarray analysis. Values reported represent percentages of the ratio between the expression level of each sample and the mean expression level of controls. Samples from the same genotype are grouped and error bars represent the standard deviation for each genetic condition. **b)** *BAZ1B* expression levels in all the genetic conditions. **c)** Top enrichments for GO biological processes among NCSC DEGs (FDR < 0.2). Adapted from (Adamo et al., 2015).

4.3. Dissection of BAZ1B role in patient-specific NCSCs

4.3.1. Generation of BAZ1B-interfered NCSC lines

The deletion or duplication of almost 30 genes makes it challenging to attribute a pathological trait to a specific gene or subset of genes in the 7q11.23 region. Atypical deletions as well as mouse models have however helped to pinpoint some gene-clinical trait associations.

Among these, one gene in particular, *BAZ1B*, is likely to underline key aspects of NC-dependent craniofacial alterations typical of WBS and 7dup patients, given the evidences both in animal models and in a WBS atypical patient (see chapters 2.4.1.1, 2.4.1.2 and 2.4.4.).

Thus, in order to have better insights into BAZ1B function in WBS and 7dup patient-derived NCSCs I selectively knocked-down (KD) its expression via RNA interference in all the pathological conditions. In particular, I infected NCSCs using lentiviral vectors expressing two short hairpin (sh) sequences with a different KD efficiency against *BAZ1B* (named sh38 and sh41) plus a vector containing a scrambled sh sequence (named scr) as negative control and I generated a total of 32 interfered NCSC lines, including the one derived from the atypical WBS patient (AtWBS1-C2) where *BAZ1B* is present in two copies as in the controls (**Figure 12**).

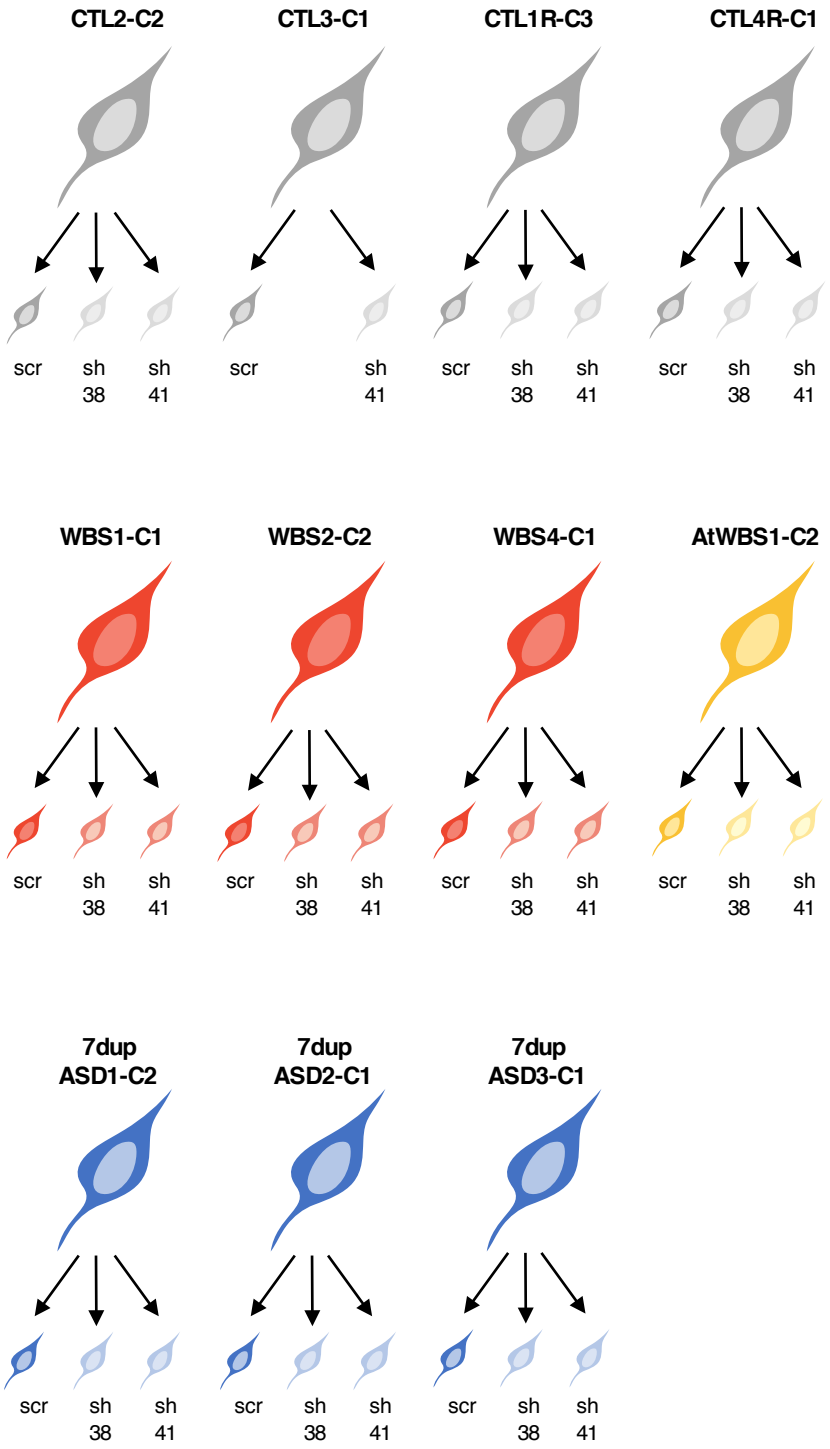


Figure 12. Scheme of the KD strategy in iPSC-derived NCSCs

Each NCSC line was interfered with two different short hairpins against *BAZ1B* (sh38 and sh41) and one sh scrambled (scr) as a negative control.

KD efficiency was first evaluated at the RNA level by quantitative PCR (qPCR) (**Figure 13a, b**), with an overall reduction of about 40% for sh38 and 70% for sh41. Then, I also confirmed these observation at the protein level by western blot analysis (**Figure 14**).

BAZ1B expression levels in the atypical WBS sample (AtWBS) were more similar to the ones of control samples than to the ones of other WBS samples, as already observed in the microarray analysis (**Figure 11b**).

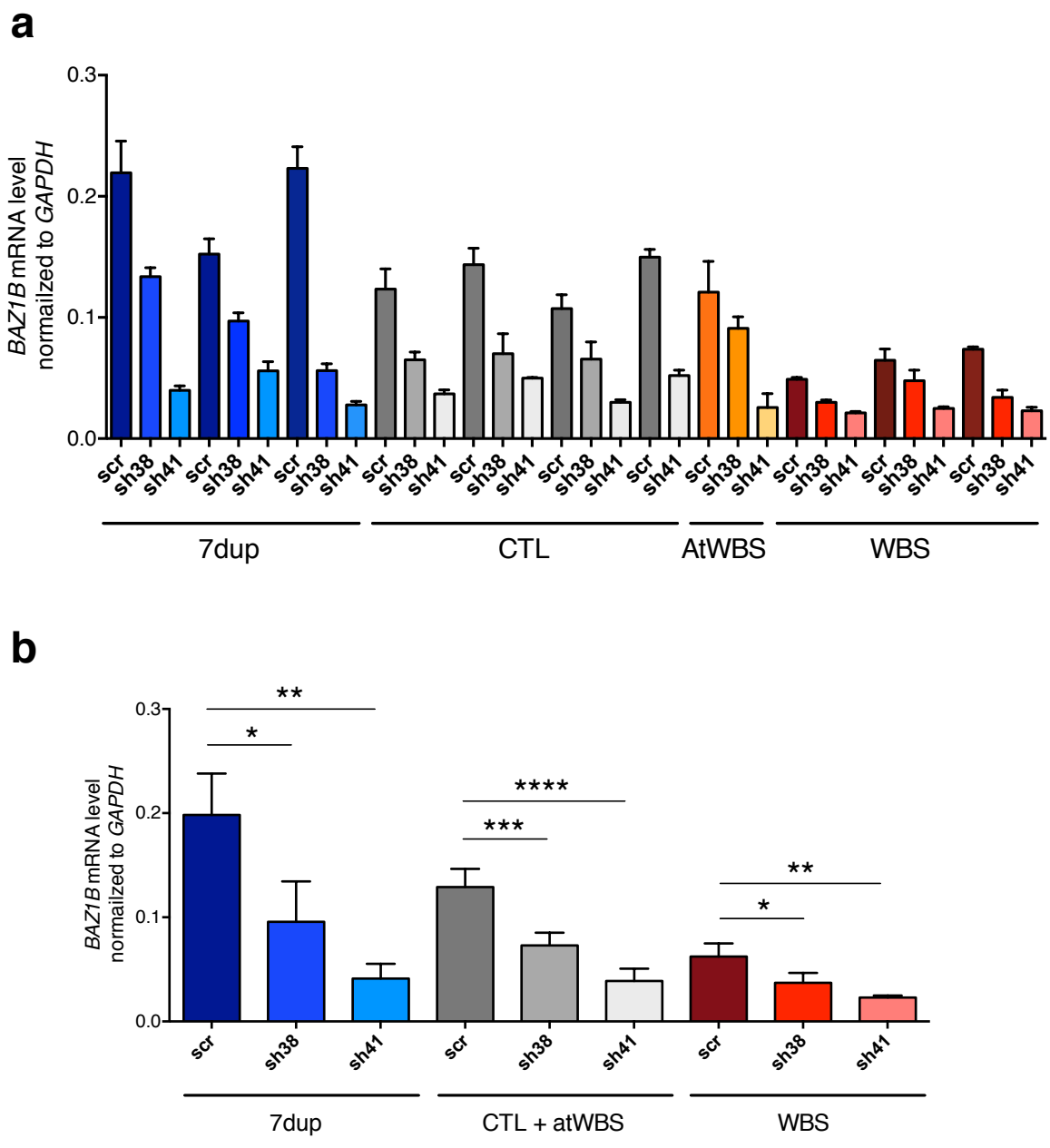


Figure 13. Validation of *BAZ1B* KD in NCSC lines through qPCR

BAZ1B mRNA levels in all the interfered lines (scr, sh38 and sh41). Data represent **a)** individual samples or **b)** aggregates of samples with the same number of *BAZ1B* copies (7dup, CTL + atWBS and WBS). GAPDH is used as normalizer. Student's *t*- test (ns: not significant; *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001 and ****: p-value < 0.0001).

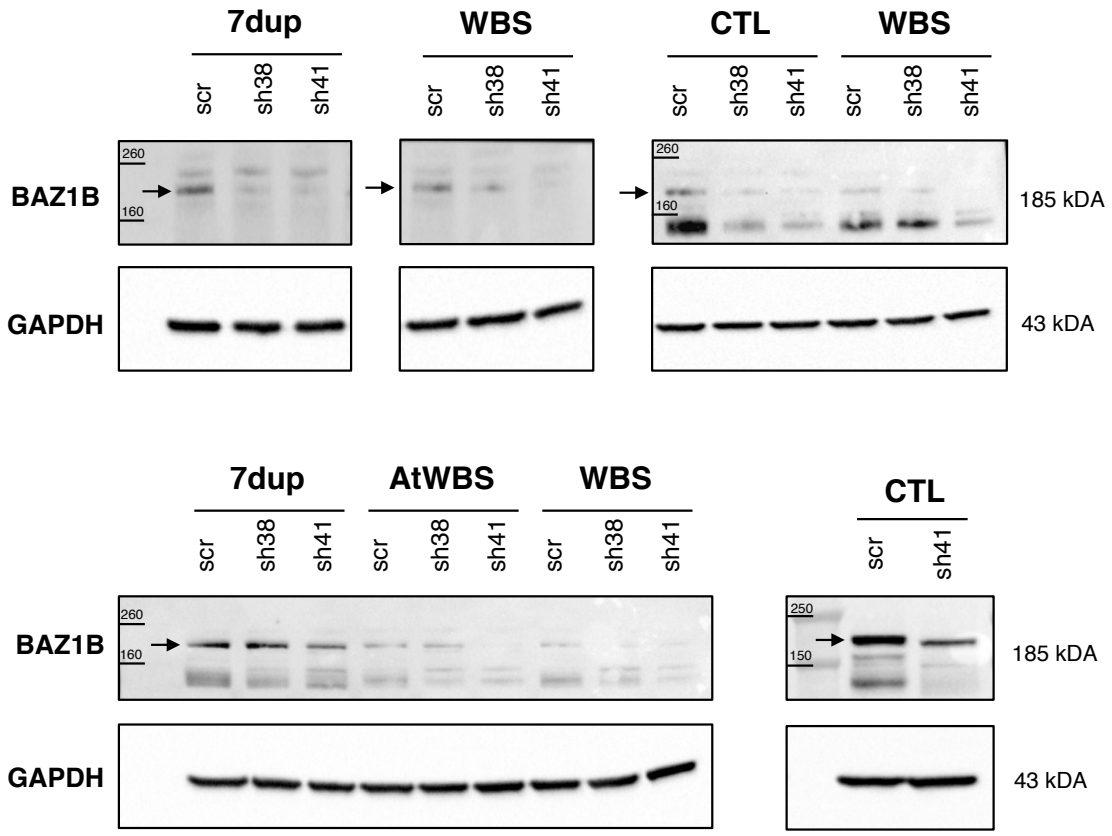


Figure 14. Validation of BAZ1B KD in NCSC lines through western blot

Protein blots showing BAZ1B levels upon KD in a subset of NCSC lines. The arrow indicates BAZ1B specific band. GAPDH is used as normalizer.

4.3.2. BAZ1B-specific impact on NCSC migration

NCSCs need to migrate in order to reach specific target regions in the developing embryo and to give rise to distinct cell types and tissues, including craniofacial structures.

Notably, the molecular pathways that were shown to be possibly dysregulated in the microarray analysis also involved categories relevant for NCSC migration. In addition, *BAZ1B* KD was shown to affect NC migration in *Xenopus laevis* and to promote cancer cell invasion in different lung cancer cell lines (Barnett et al., 2012; Meng et al., 2016).

Thus, I hypothesized that a defective regulation of NCSC migration, as a consequence of *BAZ1B* CNVs, might explain NC-related defects typical of WBS and 7dup patients.

To test this hypothesis, I compared the migration properties of *BAZ1B* KD NCSC lines to their respective control NCSC line (scr) by performing two types of migration assays:

1. a wound healing/scratch assay, to evaluate the time that cells require to fill a 500 μm cell-free gap;
2. a random migration assay, that focuses on single cell motility and, thus, better quantifies cell velocity and, subsequently, accumulated distance.

4.3.2.1. Scratch assay

For the scratch assay, NCSCs were plated at confluency in two wells of specific inserts chosen to have a more precise gap with respect to the one obtained by manually scratching the cell culture with a tip. This approach will, in turn, reduce variability between the different experimental conditions.

24 hours after plating, the insert was removed and cell migration was tracked for 20 hours.

In the first attempt, we compared NCSC lines from a 7dup, a control and a WBS patient with their *BAZ1B* KD counterparts infected with the most efficient sh vector (sh41).

Both the 7dup and the control NCSC KD lines required more time to fill the wound when compared to the respective control lines (scr), as indicated by images taken at 8 and 16 hours after the insert was removed (**Figure 15**).

No clear differences were instead observed in WBS NCSCs, where the scr line resulted as slower as the KD line in closing the gap. In particular, NCSCs from the WBS patient appeared slower than 7dup and control cells already in the 'scr' sample, further supporting BAZ1B role as a positive regulator of NC migration (**Figure 15**).

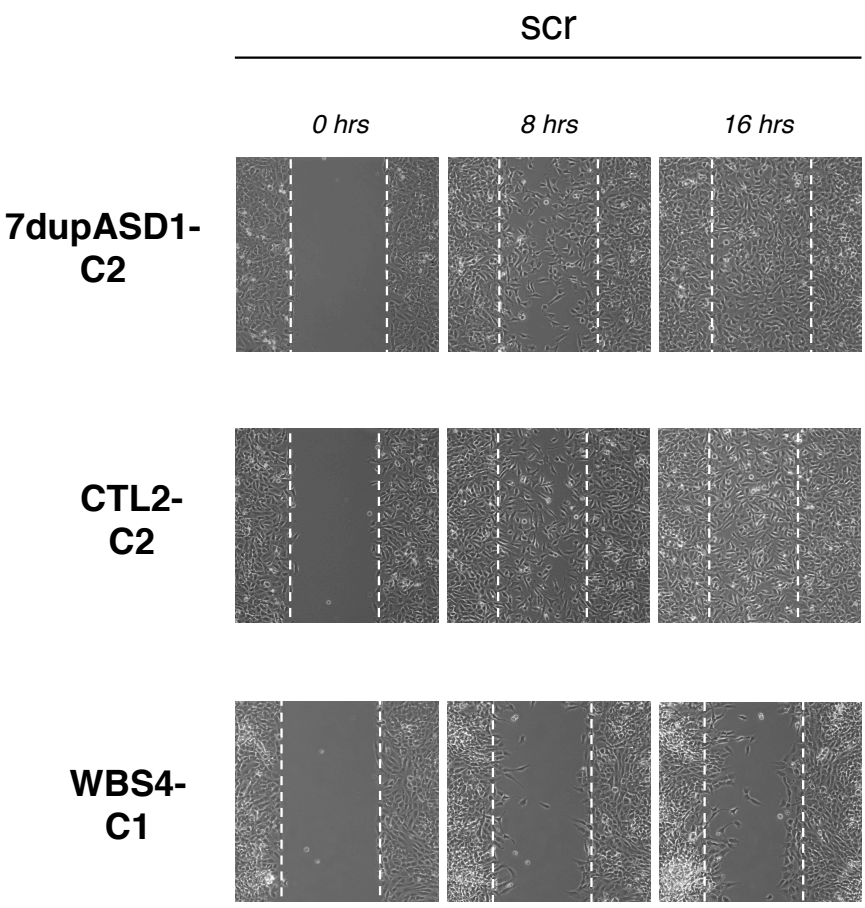

scr
0 hrs
8 hrs
16 hrs
7dupASD1-C2
CTL2-C2
WBS4-C1
shBAZ1B
0 hrs
8 hrs
16 hrs
7dupASD1-C2
CTL2-C2
WBS4-C1

Figure 15. BAZ1B KD affects NCSC migration

8 hours and 16 hours-time points from the scratch assay analysis performed on a 7dup, a control and a WBS NCSC line upon *BAZ1B* KD. Cells from the same line infected with the scrambled sh (scr) were used as references for the migration.

Notably, the delay in closing the gap was confirmed when the same analysis was performed in two other control NCSC lines (**Figure 16**).

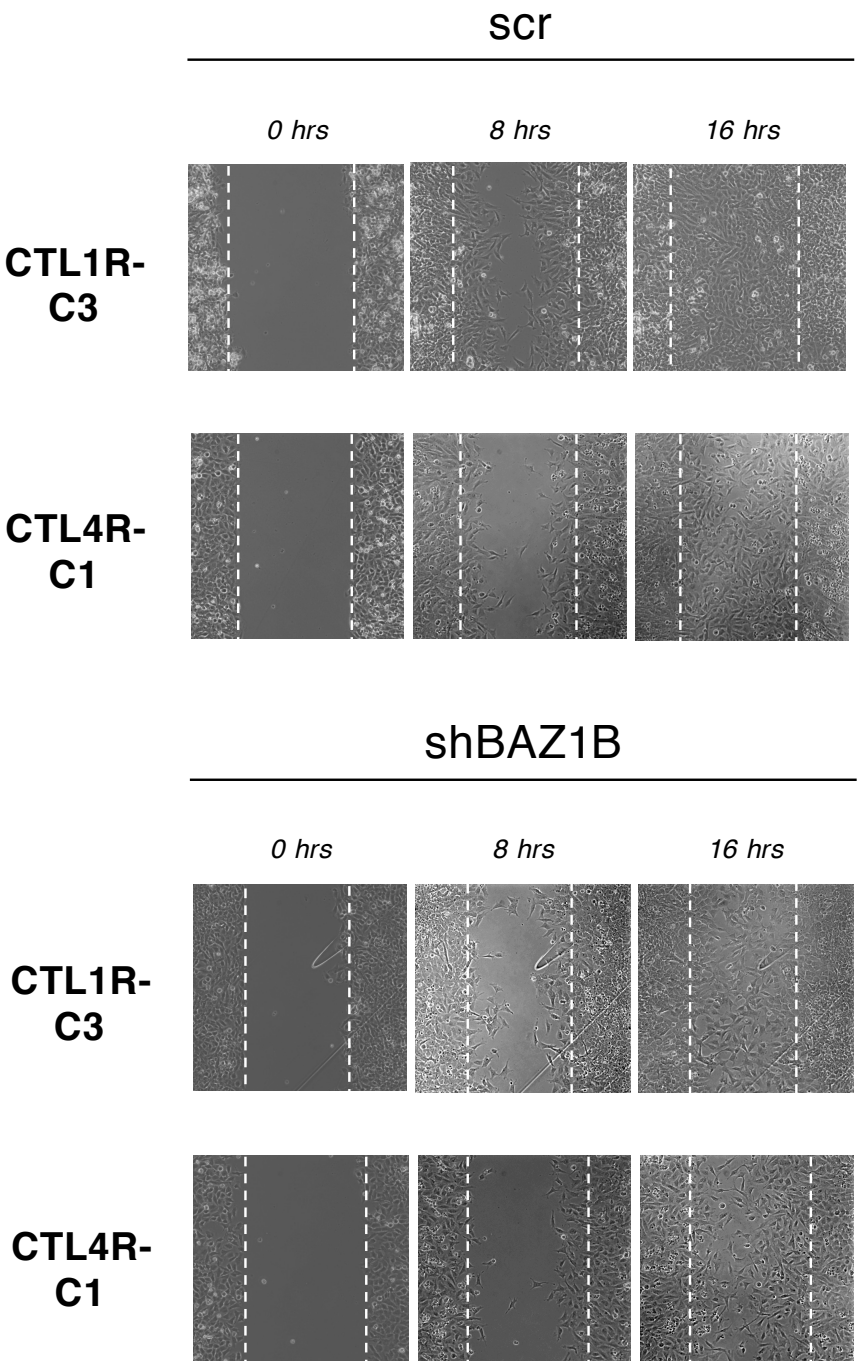


Figure 16. BAZ1B KD affects NCSC migration in two other CTL NCSC lines

8 hours and 16 hours-time points from the scratch assay analysis performed on two different control lines upon *BAZ1B* KD. Cells from the same line infected with the scrambled sh (scr) were used as references for the migration.

When looking at the entire time lapse movies, the rate of cell migration appeared similar between the two conditions (scr and shBAZ1B), while the movement of the KD NCSC lines resulted more confuse and chaotic (movies in the attached power point presentation). Thus, the apparent less organized migration upon KD could explain the delayed closure of the gap.

4.3.2.2. Random migration assay

In general, the pattern of migration of our iPSC-derived NCSCs is peculiar and different from what commonly happens in classic wound healing assays, in which cells are more packed and create a well-defined leading edge that is maintained throughout the entire migration process and until the scratch is closed. Therefore, it is difficult to have a quantitative measurement of our cell motility based on a simple scratch assay.

To overcome this problem, I performed a different assay, known as random migration assay, in which cells are plated at low confluency and, thus, can be easily manually tracked.

As noted above, I analyzed NCSCs from a 7dup, a control and a WBS patient and, per each condition, I manually tracked the migration pathways of a total number of 50 cells in 10 distinct fields (5 cells/ each field) for 6 hours.

Cell velocity and accumulated distance, strictly dependent on the former, were significantly reduced both in the 7dup and in the control line upon *BAZ1B* KD, with p values < 0.05 and < 0.01, respectively (**Figure 17a, b**).

NCSCs from the WBS patient presented a similar trend, but the reduction was not statistically significant, in line with previous results.

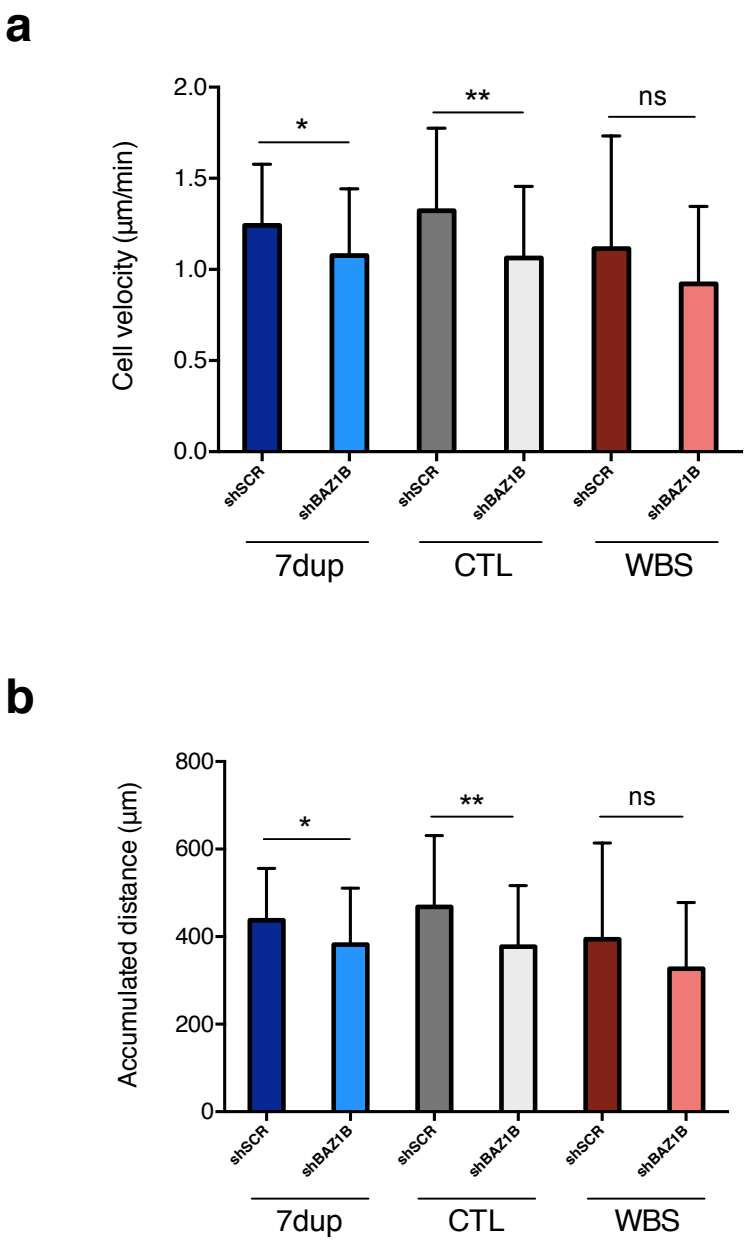


Figure 17. Quantification of NCSC migration parameters upon *BAZ1B* KD

Average cell velocity (**a**) and cell accumulated distance (**b**) before and upon *BAZ1B* KD in NCSCs from the three genetic conditions (7dup, CTL and WBS). Student's *t*- test (ns: not significant; *: p-value < 0.05; **: p-value < 0.01).

Next, in order to reduce the variability that might occur during the manual tracking of single cells, I decided to introduce a reporter gene both in scr- and shBAZ1B- interfered NCSCs and to set up the conditions for an automated cell tracking. In particular, I produced stable NCSC lines expressing a H2B-GFP fusion protein upon lentiviral infection. In addition, I FACS-sorted cells expressing high levels of GFP to further increase the homogeneity of the culture and to facilitate the automated tracking itself (**Figure 18**). Time-lapse experiments using H2B-GFP expressing NCSCs are ongoing.

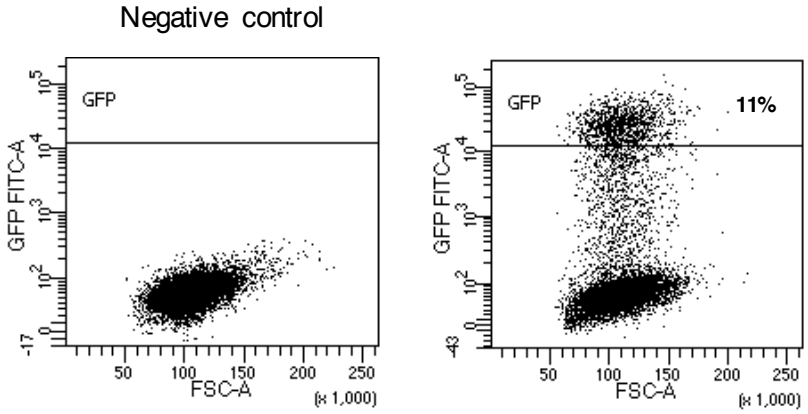


Figure 18. FACS analysis of H2BGFP-infected NCSCs

Percentage of GFP+ cells in a representative NCSC line. Only cells expressing high GFP levels were sorted**.** NCSCs not expressing GFP were used as negative control.

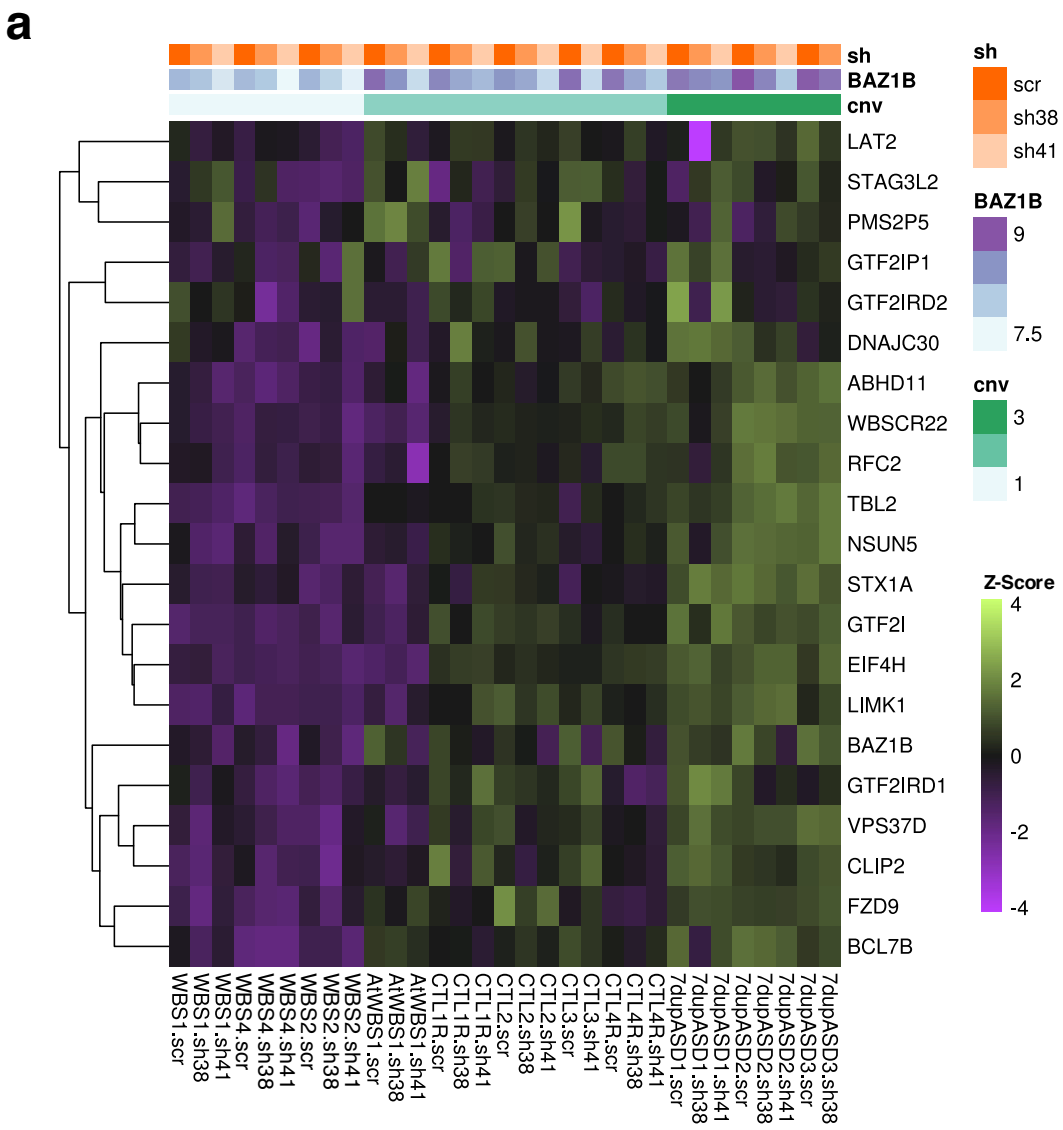
4.3.3. BAZ1B-specific impact on NCSC transcriptome

While the impact of *BAZ1B* dosage alterations in WBS and 7dup patients is still to be defined, we hypothesized that it might contribute to the above described dysregulation of NCSC transcriptome (microarray analysis on NCSCs), considering its critical role as transcriptional regulator in different cell and animal models. In particular, BAZ1B-containing complexes are able to maintain accessibility of newly synthesized DNA to transcription factors and chromatin regulators (Poot et al., 2004, 2005).

We also hypothesized that the defective migration observed upon KD might be a consequence of a dysregulation at the transcriptomic level.

To test our hypotheses, we subjected 31 interfered NCSC lines to high coverage RNA sequencing (RNA-seq) analysis. This cohort included 11 scr-, 10 sh38- and 10 sh41-interfered samples. In particular, each sample was sequenced with a coverage of 35 million of reads at a read length of 50 base pairs (bp), paired end.

Expression levels of WBSCR genes reflected the expected dosage, moving from a lower expression (in purple) in WBS samples to a gradually higher expression (in green) in control and then in the 7dup NCSC, as shown in **Figure 19a**. In particular, *BAZ1B* levels further confirmed the qPCR results on the interfered lines (**Figure 19b**).



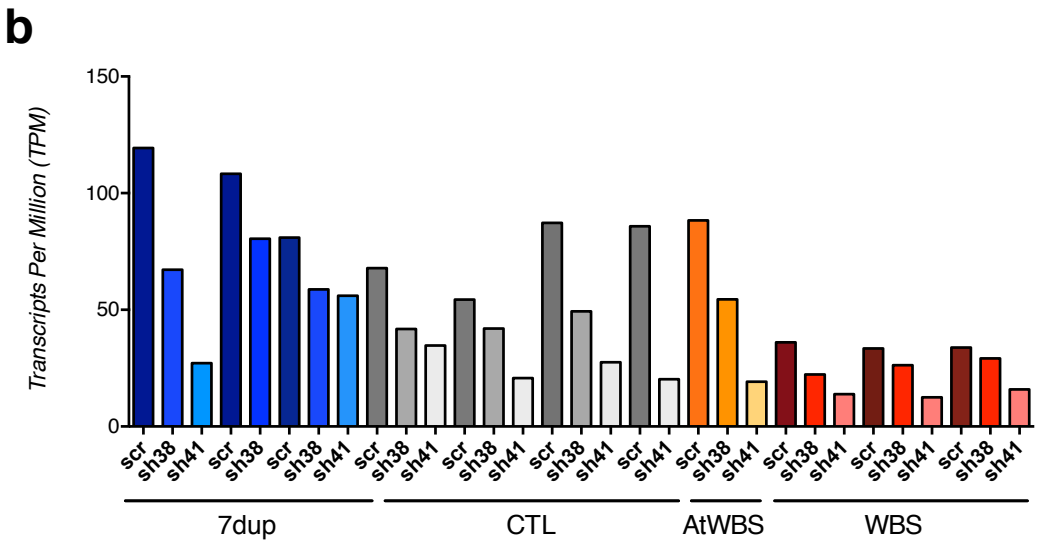


Figure 19. Expression levels of WBSCR genes in the 31 NCSC lines subjected to RNA-seq

a) Heatmap representing 21 genes included in the WBSCR and correlating with gene dosage (FDR < 0.05). Samples are ordered based on CNVs (1 for WBS, 2 for controls and 3 for 7dup) and on the type of sh (scr, sh38 and sh41). **b**) Normalized *BAZ1B* levels in the RNA-seq analysis represented as Transcripts Per Million (TPM).

Next, we performed a principal component analysis (PCA) to evaluate the distribution of transcriptomes in all the interfered samples based on an unsupervised clustering. As shown in **Figure 20**, *BAZ1B* KD did not alter the NCSC transcriptome significantly, as shown by the observation that half of the 'KD lines' tend to cluster with their respective 'scr line', supporting a strong effect of the individual genetic backgrounds. The deletion/duplication of the region appeared anyway responsible for transcriptional alterations, since the majority of both WBS and 7dup NCSC lines occupy quite distinct areas in the upper half of the graph, while control lines are more distributed in the lower half, with few exceptions.

Interestingly, the atypical WBS line shows a transcriptional profile that is closer to the control lines than to the typical WBS ones, suggesting that a significant part of the transcriptional dysregulation might be caused by genes that are not deleted in this peculiar WBS patient (including *BAZ1B*).

Given this last observation and our interest in specifically deciphering *BAZ1B* function in human NCSCs, we took advantage of the atypical WBS patient and we decided to consider this sample as a control, and not as a WBS sample, in the following analysis.

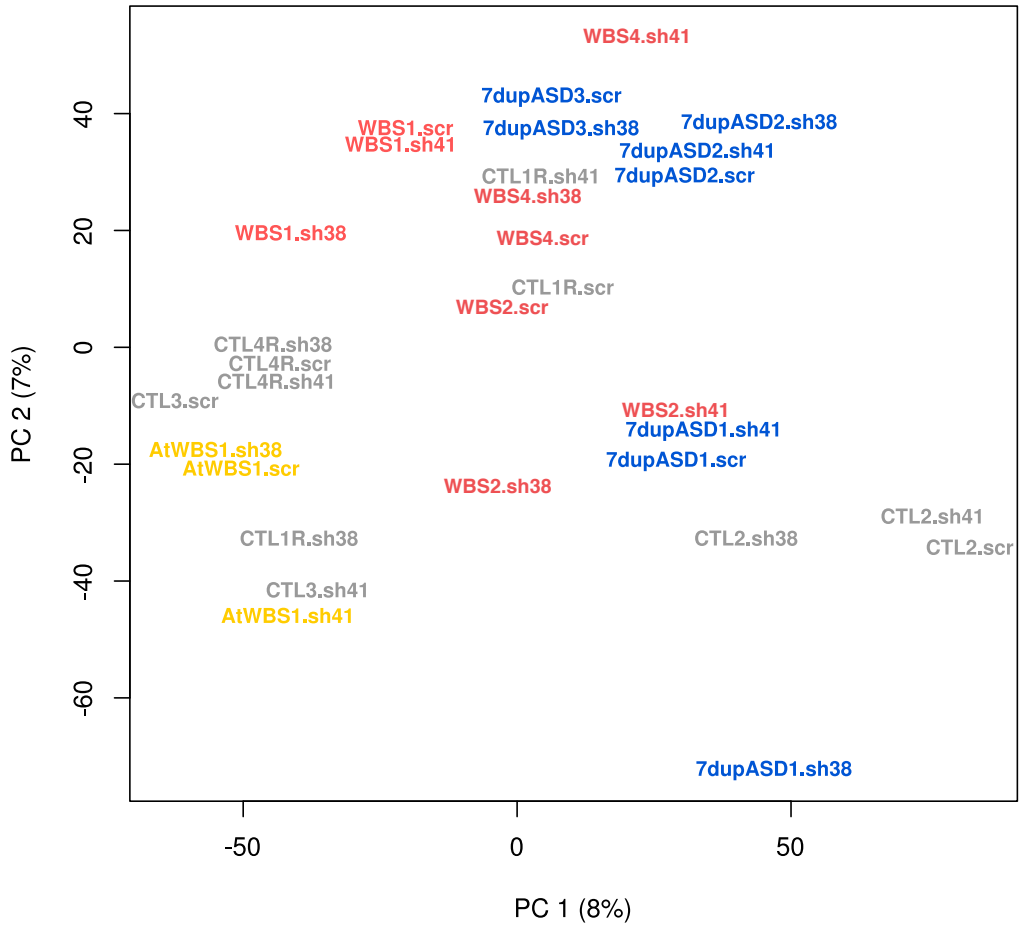


Figure 20. Principal component analysis showing the distribution of the 31 NCSC lines according to their transcriptional profiles

In particular, we resorted to four different approach in performing differential expression analysis of RNA-seq data, aiming at identifying the one that could better describe BAZ1B impact and involvement in our NCSC cohort. To reduce possible unwanted batch effects, we decided to correct all the analysis for the individual backgrounds and for the different rounds of sample processing and sequencing.

4.3.3.1. Pairwise comparative analysis: scr versus shBAZ1B (sh38+sh41)

Initially, we carried out a comparative analysis to evaluate the effects of *BAZ1B* KD. Briefly, NCSC lines generated using sh38 and sh41 were analyzed as part of a single group and compared to NCSC lines infected with the scr vector, as schematically shown in **Figure 21**.

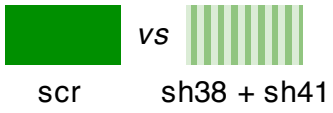


Figure 21. Schematic representation of the first strategy adopted for RNA-seq analysis

At a first sight, the identification of only 26 DEGs with FDR < 0.05 might suggest that BAZ1B impact on transcription is extremely small. However, this is in contrast with the current literature attributing to BAZ1B a crucial role as transcriptional regulator. Thus, we hypothesized that the different KD efficiency of the two sh could result in a different efficiency of dysregulation, with most of the genes being only affected with a stronger *BAZ1B* reduction (sh41) and instead not altered and/or identified as differentially expressed with a milder KD (sh38).

4.3.3.2. Multifactorial comparative analysis: scr versus sh38 versus sh41

To prove whether the two *BAZ1B* shRNA were having a different impact on NCSC transcriptome we decided to consider the lines infected with sh38 or sh41 separately and to perform a comparison between three different groups (**Figure 22**).

Figure 22. Schematic representation of the second strategy adopted for RNA-seq analysis

This second analysis identified a higher number of DEGs (151 with FDR<0.05), but only 50 of them showed a concordant dysregulation in the two groups of sh-interfered samples, with 25 being upregulated and 25 being downregulated upon *BAZ1B* KD.

In addition, most of the genes found in the previous analysis (21 out of 26) were also included in the 50 concordant genes mentioned above.

The poor overlap between the genes affected by the two different KD could also indicate an unwanted dysregulation due to possible off-targets of the shRNA. However, we excluded the involvement of genes reported as possible alternative shRNA targets in the GPP web portal (Broad Institute), since their expression levels were not affected upon infection.

Again, we concluded that this second approach was not yet sufficient to fully describe *BAZ1B* effects at the transcriptional level and that we needed a more sophisticated strategy able to integrate the notion of having two differentially effective shRNA with apparently distinguishable impacts.

4.3.3.3. Regression analysis on shRNA efficiencies

Following qPCR evidences showing residual *BAZ1B* RNA levels upon KD (around 60% for sh38-interfered lines and 30% for sh41-interfered ones when compared to their respective scr line, **Figure 13**), we decided to assign a numerical value to each of the three groups (scr, sh38 and sh41). We expected this approach to i) keep the separation between the two shRNA groups while ii) adding the notion of different efficacy, without explicitly defining what gene was supposed to follow the trend of the independent variable. Thus, we assigned values of 1 for the scr, 0.6 for the less efficient sh38 and 0.3 for the most efficient sh41 (**Figure 23**) and we identified 131 genes (FDR < 0.05) that were concordantly dysregulated in both shRNA-interfered lines.

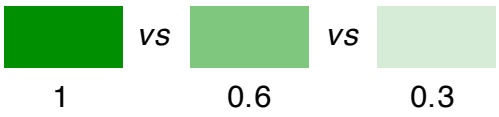


Figure 23. Schematic representation of the third strategy adopted for RNA-seq analysis

Among these genes, we identified 40 genes that were also attributed to sh38 in the multifactorial analysis and 114 genes that were similarly ascribed to sh41. Interestingly, all the 40 sh38-associated genes were included in the 114 genes associated to sh41, supporting our hypothesis of a subset of commonly dysregulated genes (40) and of another subset being affected only upon a stronger *BAZ1B* reduction (*i.e.,* sh41-mediated).

Notably, the gene with the lowest FDR was *BAZ1B* and, again, none of the known off-targets were found among the DEGs.

4.3.3.4. Regression analysis of BAZ1B-level sensitive genes

To verify the quality of the last analysis we devised a new, strongly data-driven approach in which we used *BAZ1B* expression levels from each individual sample as the independent variable (**Figure 24**). By specifically evaluating *BAZ1B* levels in each sample, we virtually increased our cohort and, thus, the statistical significance of the analysis.

Figure 24. Schematic representation of the fourth strategy adopted for RNA-seq analysis

Also in this analysis, the gene with the lowest FDR was *BAZ1B* and its FC was very similar to the one measured both in the multifactorial and in the coefficient-based analysis. Using this approach, we were able to identify a total number of 279 genes (FDR < 0.05)

whose levels followed *BAZ1B* levels, either in a direct (115) or inverse (164) way, in both sh38- and sh41-interfered samples. By ordering the samples according to the shRNA used for the infection, we could clearly distinguish a stronger transcriptional impact in the sh41 group, that, as expected, also included almost all the samples with the lowest *BAZ1B* levels (**Figure 25**). Sh38 presented instead a milder, but nevertheless distinguishable effect and resulted more transcriptionally similar to the scr-interfered lines.

In addition, the 279 identified genes included the high majority of DEGs (104/131) found in the previous analysis based on coefficients.

This last approach can be considered the most exhaustive one since it can i) efficiently combine the contributions of two shRNA with different efficiency and ii) correlate them with variations in *BAZ1B* levels.

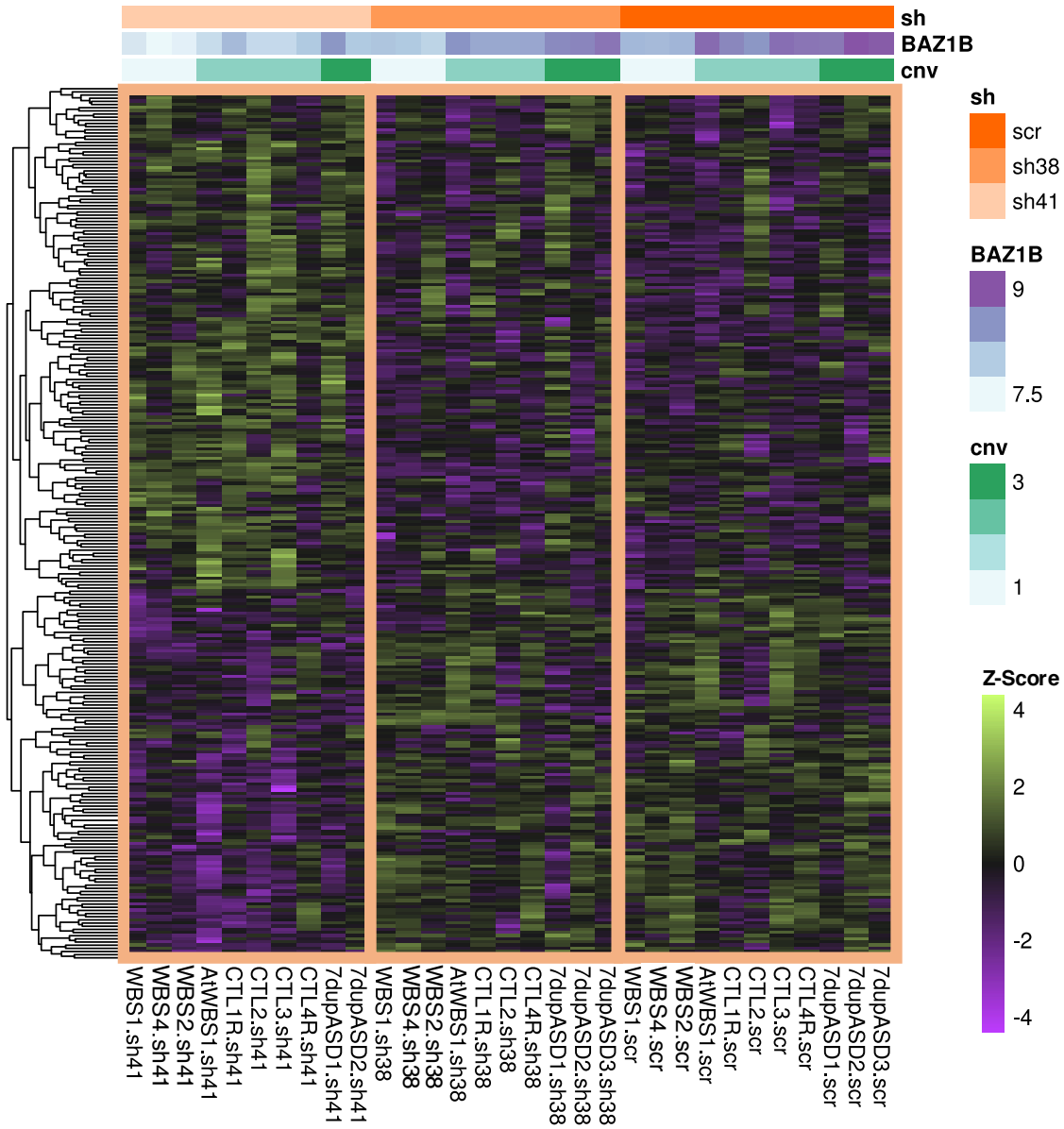


Figure 25. Expression profile of the 279 genes that follows *BAZ1B* levels in both shBAZ1B interfered lines

The samples are ordered according to shRNAs, having sh41-interferd lines on the left, sh38 in the middle and control lines on the right. In each shRNA group samples are further ordered based on *BAZ1B* copies (CNVs). FDR < 0.05.

Next, we focused our attention on some of the genes that displayed higher changes in expression levels (fold change (FC) > 2.5) and that are reported in **Figure 26**.

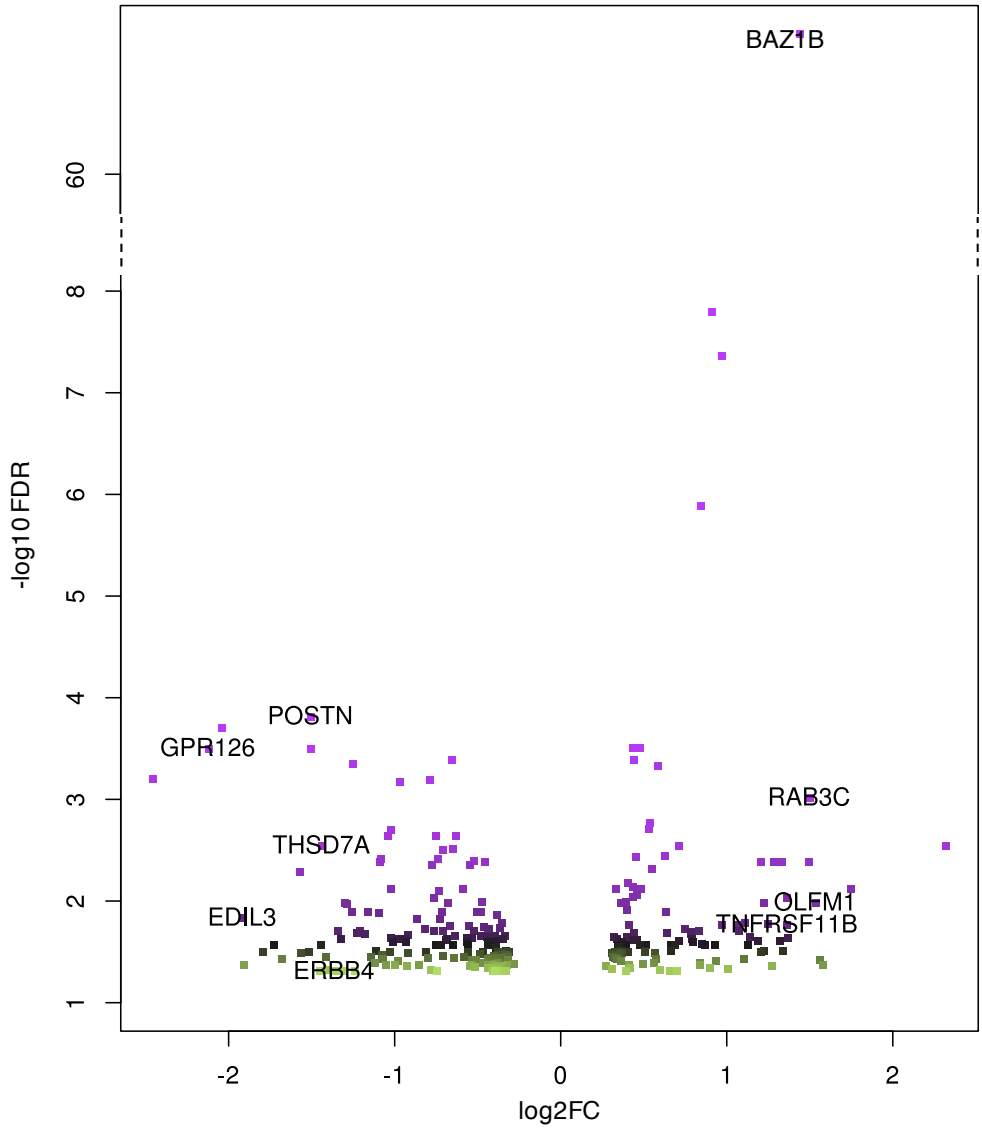


Figure 26. Genes identified in the regression analysis on *BAZ1B*-level sensitive genes

In the volcano plot, we reported the name of some of the genes with FDR < 0.05 and FC > 2.5 that followed *BAZ1B* levels in a direct (log2FC > 0) or inverse (log2FC < 0) fashion.

Within the genes that were inversely correlated with *BAZ1B* levels we identified:

- *POSTN*. It encodes for periostin (also known as osteoblast specific factor), a protein secreted by osteoblasts that is involved in the adhesion and migration of epithelial cells through its interaction with integrins and type 1 collagen (Gillan et al., 2002). It also plays a crucial role in cranial NC-mediated soft palate development (Oka et al., 2012). Interestingly, mice lacking *POSTN* present

alterations resembling the ones observed in WBS and 7dup patients such as growth retardation, heart valve defects and periodontal disease-like phenotypes (Rios et al., 2005).

- *GPR126*. It regulates differentiation and myelination in NC-derived Schwann cells (Mogha et al., 2013) and it is also involved in the development of the inner ear (Geng et al., 2013), suggesting a role in WBS patient's ear defects such as hyperacusis.
- *THSD7A* and *EDIL3* (also known as *DEL-1*). They mostly act during angiogenesis when they promote endothelial cell migration and adhesion through interactions with focal adhesion-associated proteins, including integrins (Ho et al., 2004; Wang et al., 2010).
- *ERBB4*. It functions as receptor for neuregulins, a family of proteins with crucial functions during embryogenesis, and it is important for skeletal muscle development and NC migration (Burden and Yarden, 1997). Mice lacking this protein display heart defects and aberrant cranial NC migration that results in misprojections and fusions of cranial sensory ganglia (Gassmann et al., 1995; Golding et al., 2000).

Genes that, instead, directly followed *BAZ1B* levels included:

- *RAB3C*. It is a member of the RAS oncogene family and it is involved in vescicular transport. *RAB3A*, another member of the family, was shown to interact with Rabconnectin-3a (Rbc3a) (Nagano et al., 2002), whose loss in Zebrafish NC cells led to reduced migration velocity and directionality persistence (Tuttle et al., 2014).
- *OLFM1*. It encodes for Noelin-1, a protein engaged in neural crest generation from the neural tube. In particular, overexpression of *OLFM1* promotes an increased and continued production of neural crest cells even after the time that the last cells left the neural tube (Barembaum et al., 2000).

- *TNFRSF11B*. It encodes for RANK, that negatively regulates bone resorption and mediates osteoclast activation through the interaction with his ligand (RANK-L) (Hsu et al., 1999). Interestingly, *TNFRSF11B* deficient mice display defects in bone resorption and in osteoclast development (Li et al., 2000).

Finally, using a FDR threshold < 0.1, we performed a GO analysis that suggested specific enrichments for biological processes, particularly relevant for NC and NC-derivative development and function, some of which are reported in **Figure 27**.

Figure 27. Top most-specific enrichments for GO biological processes among the DEGs in the regression analysis on *BAZ1B*-level sensitive genes

For this analysis, we selected DEGs with FDR<0.1 and only enrichments with FDR <0.05 were considered. Values in brackets represent the significance (FDR) of each enrichment.

Briefly, the dysregulated genes are known to be mostly involved in:

- cell motility, chemotaxis and migration (*e.g.*, *ANGPT1*, *CDH2*, *ERBB4*, *FN1*, *NR2F2, PDGFRA* and *POSTN*), in line with the migration defects observed upon *BAZ1B* KD (chapter 4.3.2.);
- heart and vasculature development (*e.g.*, *ANGPT1*, *CDH2, FLT4*, *MMP2*, *NR2F2, TFGB3* and *THSD7A*);
- skeletal system, bone and head development and morphogenesis (*e.g.*, *COL11A1, CDH2*, *ERBB4*, *MMP2*, *PDGFRA*, *POSTN* and *TNFRSF11B*)
- neuron and axon guidance (*e.g.*, *CDH4*, *SEMA5A*, *SEMA3C* and *TGFB2*).

4.3.4. BAZ1B-specific impact on NCSC putative enhancer regions

Many convergent lines of evidence indicate that *BAZ1B* regulates replication and transcription by directly acting on chromatin. In particular, BAZ1B is a subunit of ISWI chromatin remodelling complexes that increase DNA accessibility to TFs through an ATP-dependent mobilization of nucleosomes (chapter 2.4.3.1.).

A deeper knowledge of NC cells chromatin state, in particular through an accurate annotation of enhancer regions, was recently shown to be fundamental in understanding their development and transcriptional regulation (Rada-Iglesias et al., 2012). Indeed, the vast majority of NC cell enhancer regions are only found in this peculiar cell type and not in hESCs or hNECs.

Direct evidences of BAZ1B involvement in the specific remodeling of NCSC enhancer regions are still missing. Notably, it was already reported that another chromatin remodeler (*i.e.*, CHD7) is able to transcriptionally regulate NC migration specifiers through the binding of their enhancer regions (Bajpai et al., 2010).

Moreover, it is known that bromodomains contained in several chromatin remodelers, including BAZ1B, are required for the recognition of acetylated lysines on histones (Jones

et al., 2000). However, to date, acetylation targets of WSTF bromodomain have yet to be identified.

Thus, to evaluate whether *BAZ1B* dosage-dependent transcriptional alterations were involving specific regulatory regions, we investigated the chromatin profiles of control and patient-derived NCSCs upon *BAZ1B* KD through ChIP sequencing (ChIP-seq). In particular, we investigated the acetylation and the trimethylation of lysine 27 (H3K27ac and H3K27me3) and the monomethylation of lysine 4 (H3K4me1). H3K27ac can be found both at proximal (*e.g.*, promoter) and distal (*e.g.,* enhancer) regulatory regions. In particular, regions displaying both H3K4me1 and H3K27ac mark are typical of active enhancers, while H3K27me3 marks inactive regions (Rada-Iglesias et al., 2011).

The cohort of NCSCs subjected to ChIP-seq initially included 2 controls, 2 typical WBS, 1 atypical WBS and 2 7dup lines and was later supplemented with sh38- and sh41-interfered lines corresponding to each of the previous samples, for a total of 21 NCSC lines.

A regression analysis on BAZ1B levels performed on the entire cohort indicated that around 2000 regions were differentially acetylated on H3K27 (p value < 0.005) and 266 of them showed an FDR < 0.1 (data not shown). Within the top 250 H3K27 differentially acetylated regions, H3K27 acetylation sites were mainly found at promoter regions (~36%) and at both downstream and upstream distal regulatory sites (~47%), that include the second and the following introns (~21%) and the intergenic regions (~26%), respectively. Interestingly, 115/250 regions were associated with mesenchymal stem cell (MSC)-specific enhancer regions according to the 4DGenome. Data on NCSC-specific enhancer regions are not yet included in the above-mentioned database and MSCs represent the closest cell type for which data are currently available.

Taken together, these preliminary results suggest that BAZ1B might act at proximal but mostly at distal regulatory regions.

Given these observations, we initially focused on putative active enhancer regions (displaying the H3K4me1 and H3K27ac marks) in the 2 control samples (around 9600 regions) and we considered them a reference of the physiological distribution of marks.

Next, we looked at the distribution of each mark (H3K27ac, H3K4me1 and H3K27me3) in all the 21 NCSC lines along the putative active regions identified in the controls. Finally, we resorted to the same approach applied to RNA-seq analysis and we only selected the regions where the distribution of the histone mark directly or inversely followed *BAZ1B* levels.

In particular, when *BAZ1B* levels were average to low (the left half of the heatmap), we observed a higher presence of the H3K27me3 (**Figure 28**) in the high majority of the regions (96/114) and a lower presence of the same mark with high *BAZ1B* levels. These results indicate that, in these regions, the H3K27me3 mark is globally inversely correlated with *BAZ1B* levels.

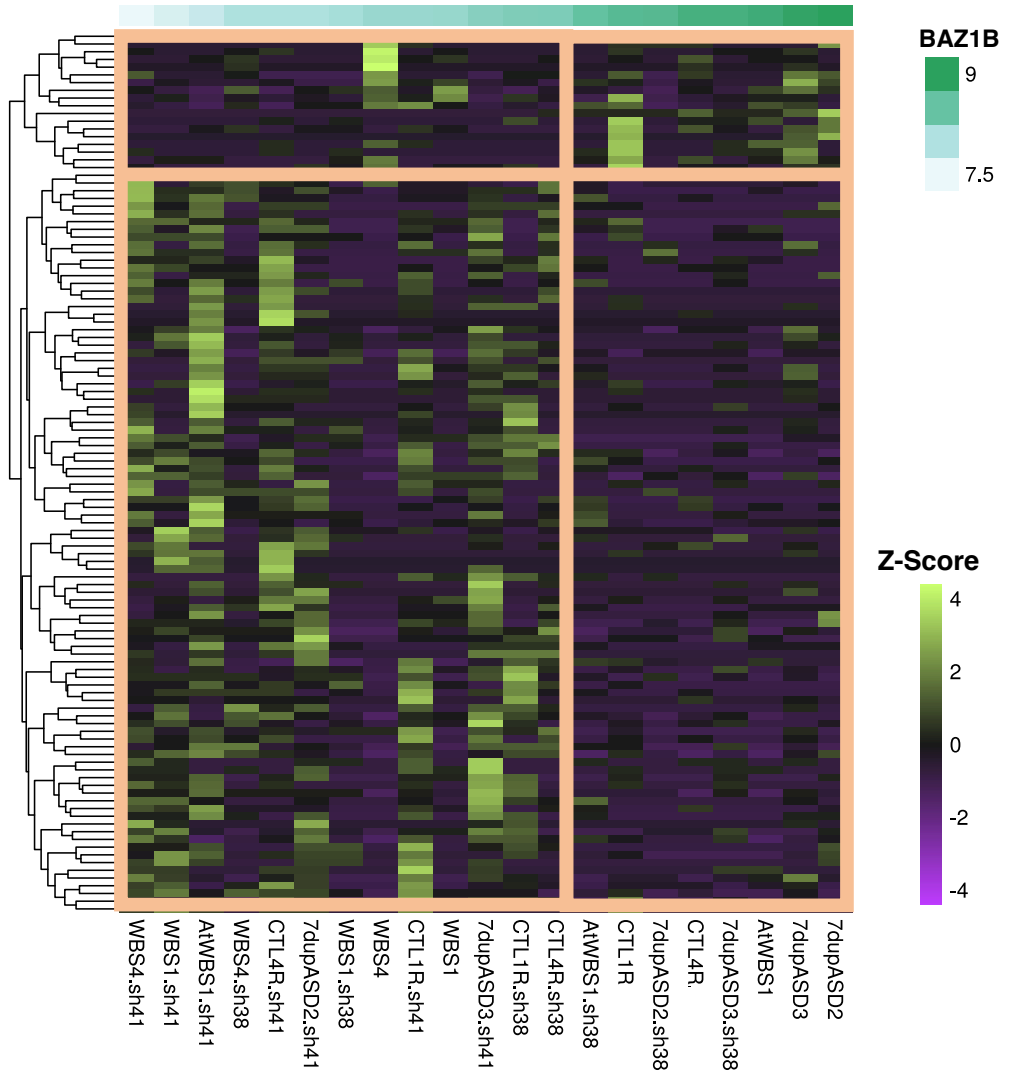


Figure 28. H3K27me3 ChIP-seq signal profile of the chromosomal regions in which H3K27me3 directly or inversely follows BAZ1B levels

The 114 represented regions were all identified as putative active enhancers in the control samples and are ordered based on BAZ1B levels. p value < 0.005.

To better appreciate the impact of *BAZ1B* KD on the distribution of the H3K27ac, H3K4me1 and H3K27me3 marks, we ordered the samples by sh (non-infected, sh38- and sh41-interfered lines). The sh41-mediated KD, in the right part of the heatmap, was responsible for a stronger redistribution of all the three marks, when compared to the non-infected samples on the left (**Figure 29, 30 and 31**), in the same way as in the RNA-seq analysis (**Figure 25**).

Both the H3K4me1 (**Figure 29**) and the H3K27ac mark (**Figure 30**) followed *BAZ1B* levels in a direct way in about half of the identified regions (upper part of the heatmap) and in an inverse way in the remaining half (lower part). On the contrary, the H3K27me3 mark mostly followed BAZ1B levels in an inverse way (**Figure 31**), as already observed in the previous heatmap on the same mark.

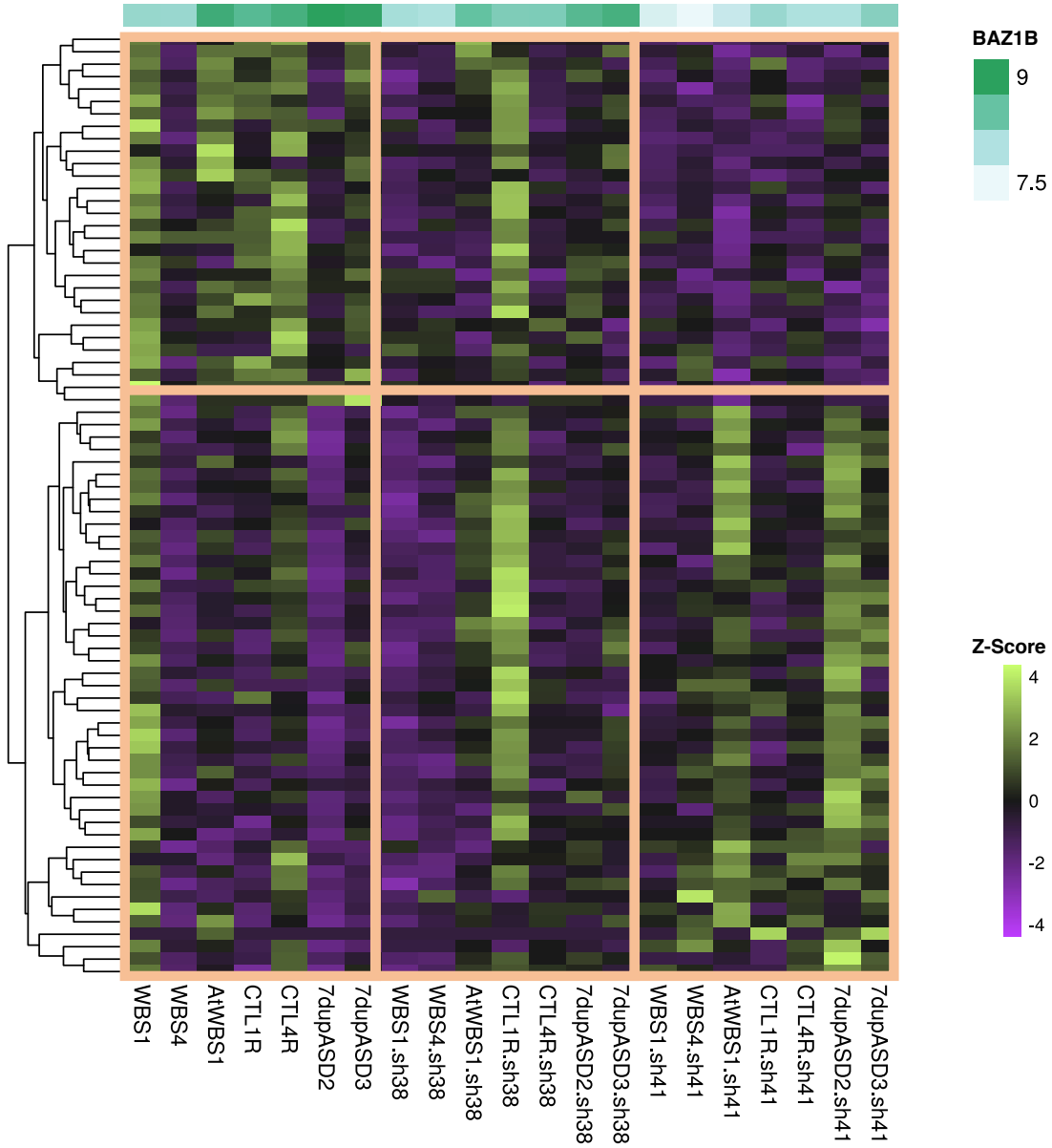


Figure 29. H3K4me1 ChIP-seq signal profile of the chromosomal regions in which H3K4me1 directly or inversely follows BAZ1B levels

The 76 represented regions were all identified as putative active enhancers in the control samples and are ordered based on the type of NCSC sample (non-infected, sh38- or sh41-interfered). p value < 0.005.

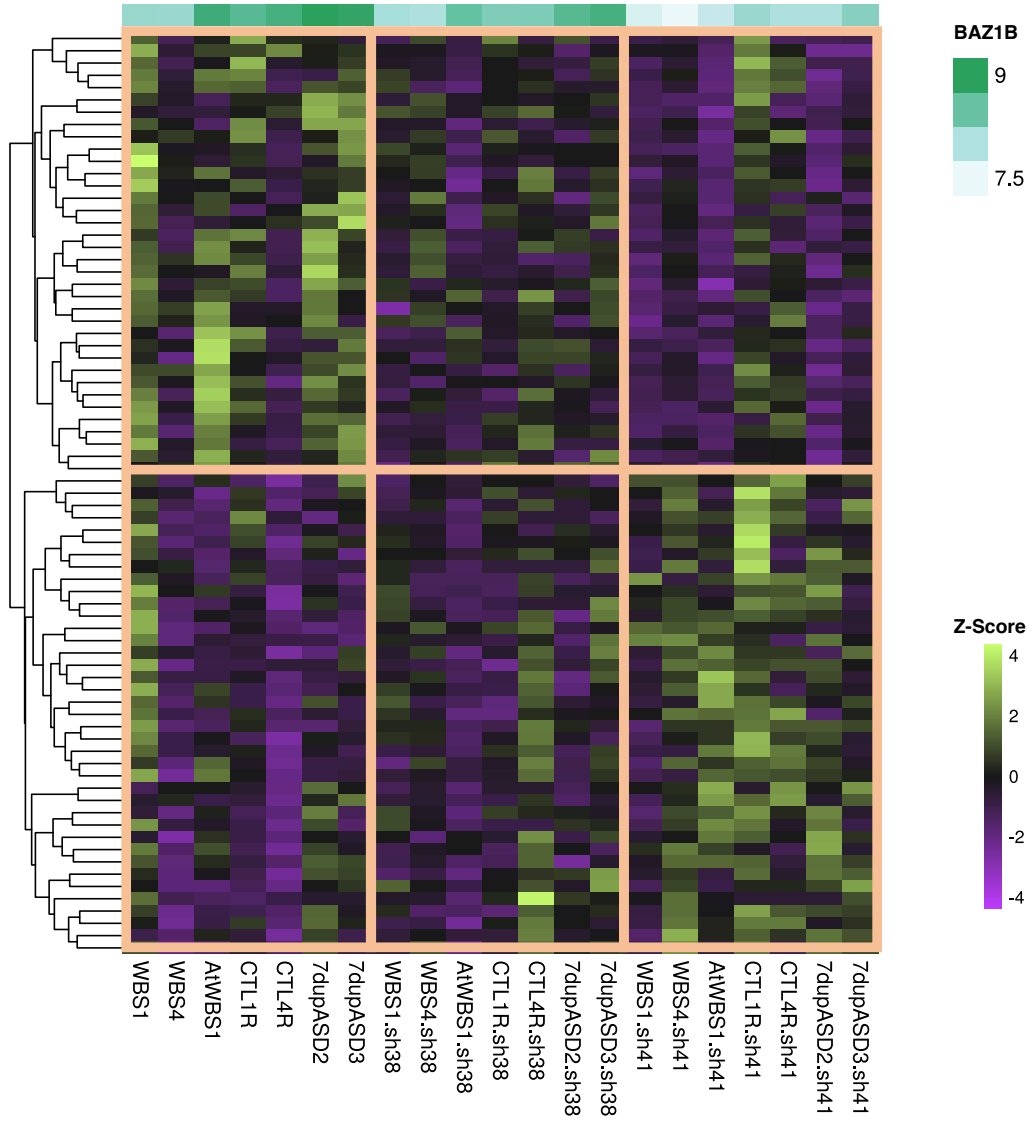


Figure 30. H3K27ac ChIP-seq signal profile of the chromosomal regions in which H3K27ac directly or inversely follows BAZ1B levels

The 75 represented regions were all identified as putative active enhancers in the control samples and are ordered based on the type of NCSC sample (non-infected, sh38- or sh41-interfered). p value < 0.005.

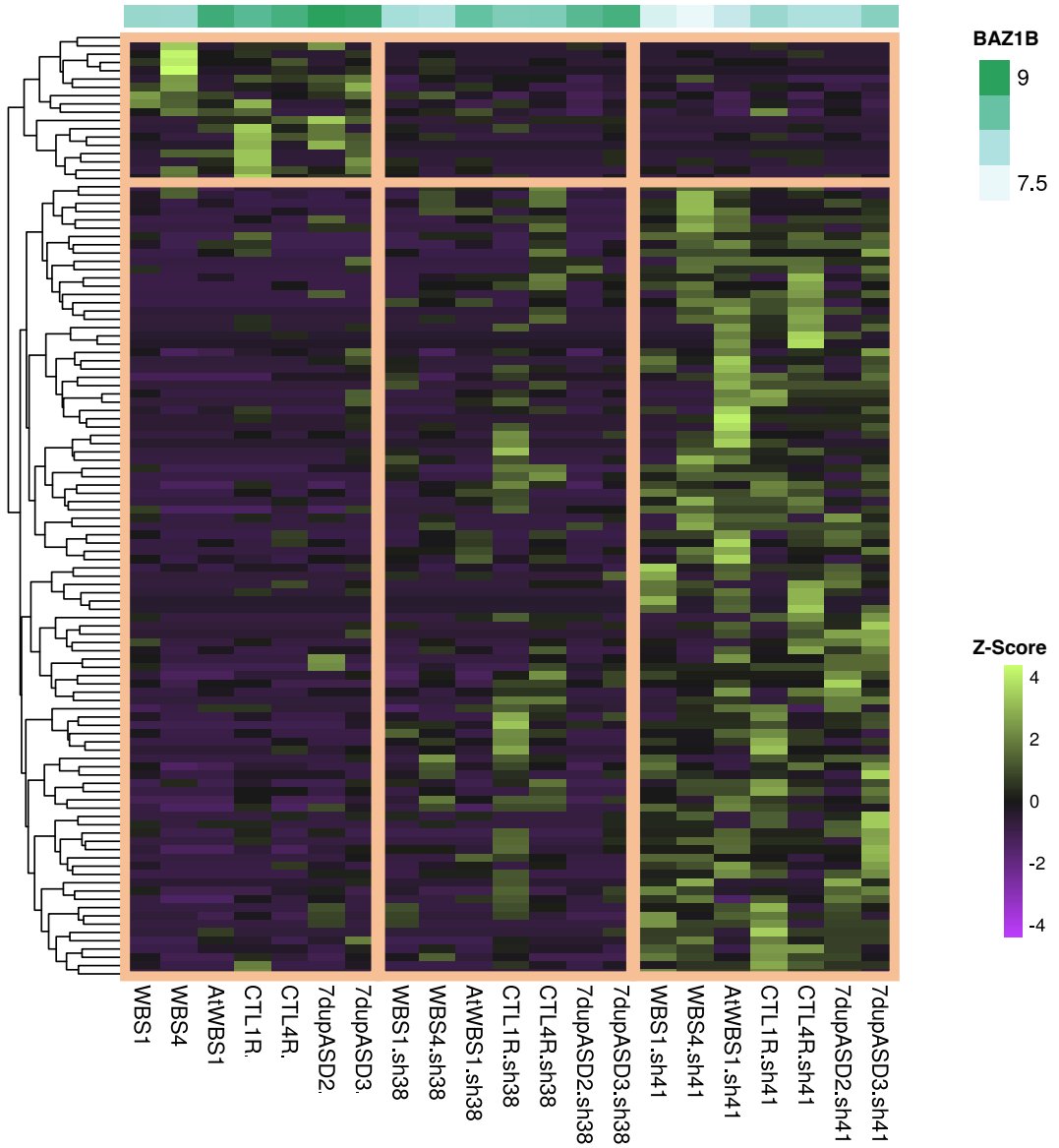


Figure 31. H3K27me3 ChIP-seq signal profile of the chromosomal regions in which H3K27me3 directly or inversely follows BAZI1B levels (alternative sample distribution)

The 114 represented regions were all identified as putative active enhancers in the control samples and are ordered based on the type of NCSC sample (non-infected, sh38- or sh41-interfered). p value < 0.005.

The dysregulated regions identified for each mark did not show any overlap, suggesting that BAZ1B might affect the deposition of all the three marks in regions that are distinct.

Thus, to have some insights into their biological function, we annotate the regions identified in the differential analysis for each histone mark, using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010). This online tool

(http://bejerano.stanford.edu/great/public/html/) is able to predict the association of cis-regulatory regions with one or more genes and, thus, with specific biological functions.

We distinguished between regions in which the mark followed *BAZ1B* levels in a direct or indirect fashion, we isolated genes associated with each specific subgroup of regions and we performed GO enrichment analysis with FDR < 0.05. The genes predicted to be associated with regions in which the H3K4me1 mark directly or inversely followed *BAZ1B* levels, were enriched for relevant GO categories, including cell motility and migration, muscle cell differentiation and several morphogenesis-related processes, in line with NC-dependent formation of craniofacial structures (**Figure 32**).

Genes associated with H3K27 differentially acetylated regions were only enriched for odontogenesis in line with patient-dental defects, while no enrichments were identified for the H3K27me3 mark.

Figure 32. Top most-specific enrichments for GO biological processes among the genes that were predicted to be associated with differential H3K4me1 regions

Only enrichments with FDR < 0.05 are shown. Values in brackets represent the significance (FDR) of each enrichment.

4.3.4.1. Integration of results from RNA-seq and ChIP-seq analysis

In order to integrate results obtained in the last two chapters, we looked at the overlap between all the genes associated with *BAZ1B*-level sensitive putative active enhancer regions and the *BAZ1B*-level sensitive genes identified in the RNA-seq analysis with FDR < 0.1 (see **Figure 27**). Surprisingly, only 15 of them (*i.e.*, *ALPK2*, *ARFGAP1*, *CNTNAP1*, *DISP1*, *GPR124*, *HMGB1*, *LARP4B*, *NKAIN4*, *NXN*, *POU3F1*, *RASL11B*, *RBMS3*, *RHOU*, *SULF2* and *TBX2*) were included in both datasets, first suggesting a poor coherence between the results obtained with the two analyses.

Notably, the majority of these genes are directly on inversely correlated with *BAZ1B* levels in a coherent way in both RNA-seq and ChIP-seq analyses. For example, both *DISP1*, *POU3F* and *RHOU* mRNA levels and H3K27ac levels displayed a direct correlation with *BAZ1B* expression.

Gene	**Correlation with BAZ1B levels in RNA-seq**	**Correlation with BAZ1B levels in ChIP-seq**
ALPK2	Inverse	Inverse (H3K27me3)
ARFGAP1	Direct	Inverse (H3K27me3)
CNTNAP1	Inverse	Inverse (H3K27me3)
DISP1	Direct	Direct (H3K27ac)
GPR124	Inverse	Inverse (H3K27me3)
HMGB1	Direct	Inverse (H3K27me3)
LARP4B	Direct	Inverse (H3K27ac)
NKAIN4	Direct	Inverse (H3K27me3)
NXN	Inverse	Inverse (H3K4me1)
POU3F1	Direct	Direct (H3K27ac)
RASL11B	Inverse	Inverse (H3K27me3)
RBMS3	Inverse	Inverse (H3K4me1)
RHOU	Direct	Direct (H3K27ac)
SULF2	Inverse	Inverse (H3K4me1)
TBX2	Inverse	Inverse (H3K4me1)

Table 6. List of genes identified as commonly dysregulated in RNA-seq and ChIP-seq analyses

I indicated whether each gene expression level and the type of histone mark found differentially enriched in the associated putative enhancer regions, were directly on inversely correlating with *BAZ1B* gene expression levels.

These genes have different biological functions, including cell migration (*RHO* and *GPR124*) and cell signaling (*ARFGAP1*, *RHOU*, *SULF2* and *DISP1*). Interestingly, *RHOU* was shown to play a crucial role in the generation of craniofacial cartilages through the positive regulation of cranial NC migration (Fort et al., 2011).

Moreover, preliminary analysis aimed at detecting possible interactions between proteins encoded by all the genes included in the RNA-seq and ChIP-seq datasets (https://string-db.org), identified a significant number of interactions (**Figure 33**). The interactions were based on experimental data either performed on the very same proteins or on putative homologs known to interact in other species. Interestingly, many of the described interactions involved a protein included in the RNA-seq dataset and another one identified in the ChIP-seq analysis (indicated by the red lines). This initial result suggests a first link between the *BAZ1B* dosage-mediated redistribution of histone marks along putative active enhancer regions and the transcriptional dysregulation observed with the RNA-seq analysis.

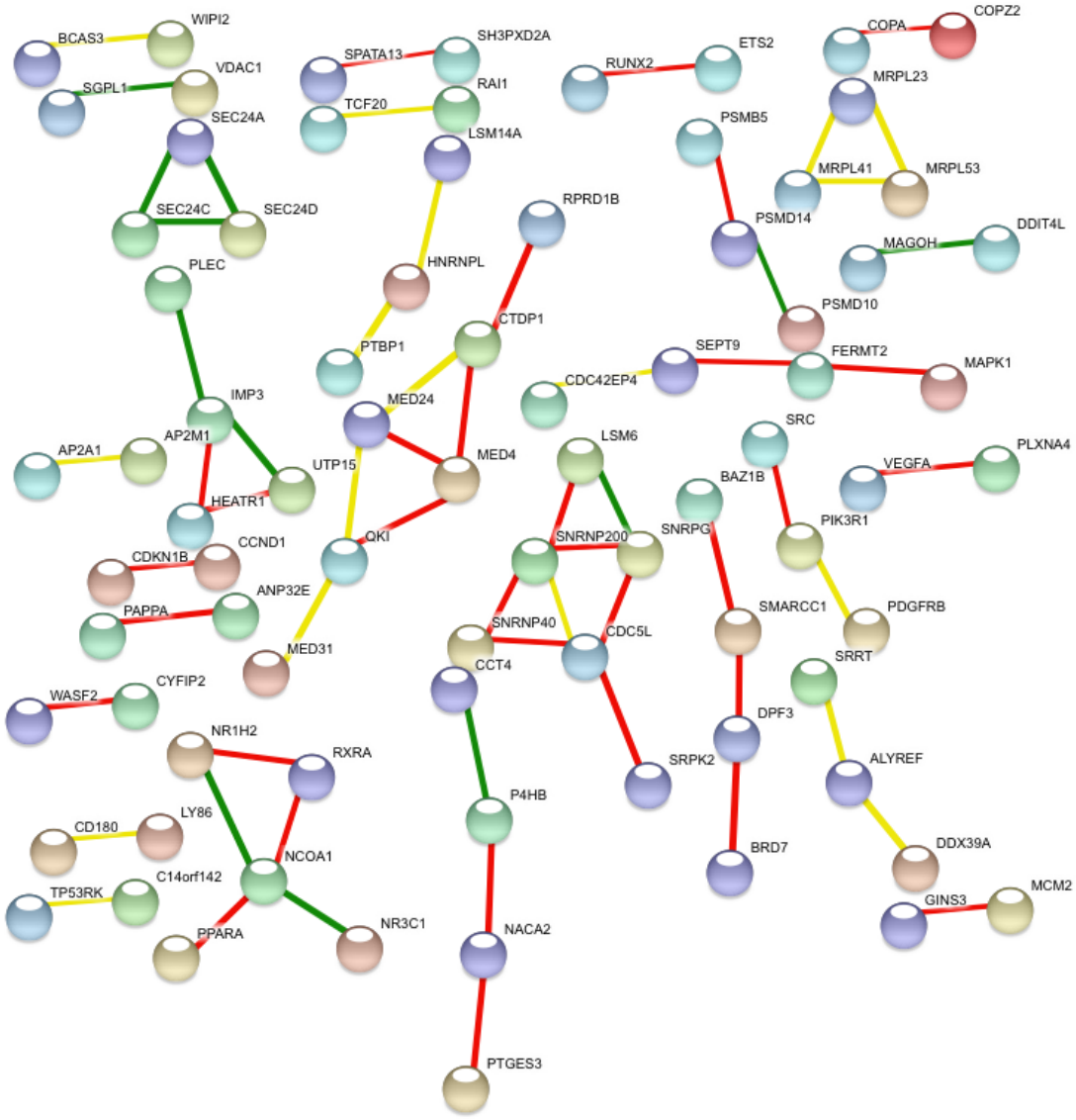


Figure 33. STRING analysis on *BAZ1B*-level sensitive genes identified in RNA-seq and ChIP-seq

The green line represents an interaction between two proteins encoded by genes from the RNA-seq dataset, the yellow line represents an interaction between two proteins encoded by genes from the ChIP-seq dataset and the red line indicates an interaction between a protein from one dataset and a protein from the other dataset. The minimum required interaction score used was 0.700 (high confidence) and only interactions supported by experiments were considered.

4.3.5. Investigation of BAZ1B direct interactors

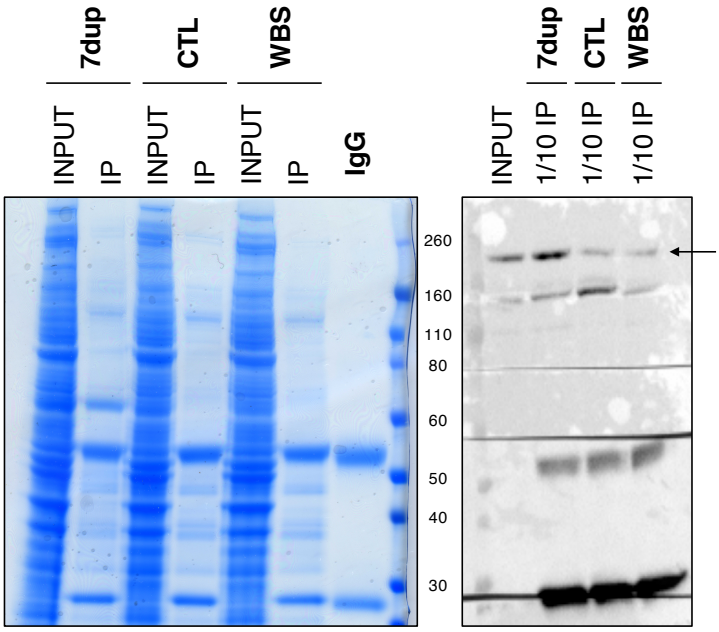
RNA-seq analysis alone is not able to distinguish between direct and indirect effects of BAZ1B in the dysregulation of NCSC transcriptional programs described in the previous chapters. Preliminary results suggest some interactions between proteins whose expression resulted dependent on BAZ1B levels (see **Figure 32**), but is still not sufficient for the identification of BAZ1B direct partners.

Thus, in order to investigate BAZ1B direct interactors and to evaluate possible BAZ1B dosage-dependent differences in the formation of specific complexes, I performed BAZ1B immunoprecipitation (IP) followed by Mass Spectrometry (MS) analysis.

In total, I tested two different antibodies known to work in IP, but with none of them I could obtain good enrichments for the bait (*i.e.*, BAZ1B).

In the first attempt, I immunoprecipitated BAZ1B in all the three genetic conditions using an antibody produced by Santa Cruz (**Figure 34a**) and I identified around 700 proteins as potential interactors. The most enriched ones were involved in splicing and RNA binding, suggesting novel functions for BAZ1B (**Figure 34b**). SMARCA5, known to be associated with BAZ1B in the WICH complex (chapter 2.4.3.1.), was also identified in all the conditions. However, I could only detect few peptides attributable to BAZ1B, possibly suggesting a low antibody efficiency.

a



b

		Peptide count			
Identified proteins	**MW**	**CTL**	**WBS**	**7dup**	**IgG**
116 kDA U5 small nuclear ribonucleoprotein component	109 kDa	60	47	50	0
Splicing factor 3B subunit 1	146 kDa	47	13	44	0
Splicing factor 3B subunit 3	136 kDa	53	17	37	0
Heterogeneous nuclear ribonucleoprotein A3	40 kDa	22	24	20	0
RNA-binding protein 25	100 kDa	25	23	31	0
Splicing factor 3B subunit 2	98 kDa	35	13	25	0
Pre-mRNA-splicing factor SYF1	100 kDA	33	27	25	0
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 (SMARCA5)	122 kDa	8	3	9	0
Tyrosine-protein kinase BAZ1B	**171 kDa**	**1**	**4**	**6**	**0**

Identified proteins: 709

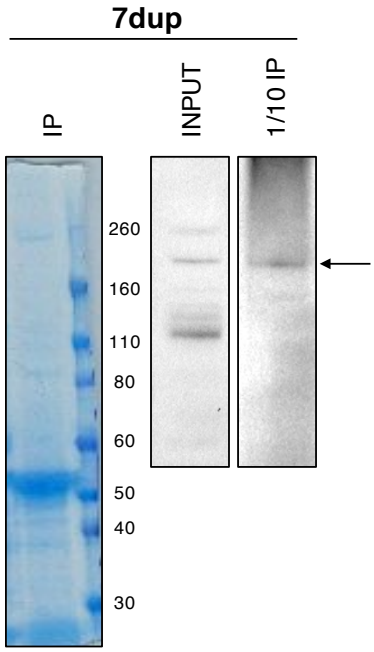
Figure 34. First attempt of BAZ1B immunoprecipitation using an antibody produced by Santa Cruz

a) 9/10 of the IP were stained with Comassie blue and subjected to MS analysis, while 1/10 of the IP was analyzed by Western Blot as internal control. **b**) Most enriched interactors according to Mass Spectrometry analysis of the immunoprecipitates showed in panel a.

Given the small BAZ1B enrichment, I tested the efficiency of a different antibody (produced by Abcam) by performing BAZ1B IP in a NCSC line from a 7dup patient

(**Figure 35a**). In this second attempt, I obtained a significantly higher enrichment for the bait: 25 peptides versus the 6 identified on the very same line in the previous immunoprecipitation (**Figure 35b**). In addition, I observed a good enrichment for proteins involved in the maintenance of chromosome structures and SMARCA5 was still detected, together with another member of the SWI/SNF family, SMARCA4. Splicing proteins were still present, but with a lower enrichment.

a



b

Identified proteins	MW	7dup
Structural maintenance of chromosomes protein 1A	143 kDA	25
Structural maintenance of chromosomes protein 3	142 kDA	25
Structural maintenance of chromosomes protein 4	144 kDA	12
116 kDA U5 small nuclear ribonucleoprotein component	109 kDa	10
Splicing factor 3B subunit 1	146 kDa	8
Splicing factor 3B subunit 3	136 kDa	10
Transcription activator BRG1 (SMARCA4)	189 kDa	20
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 (SMARCA5)	122 kDa	3
Tyrosine-protein kinase BAZ1B	**171 kDa**	**25**

Identified proteins: 587

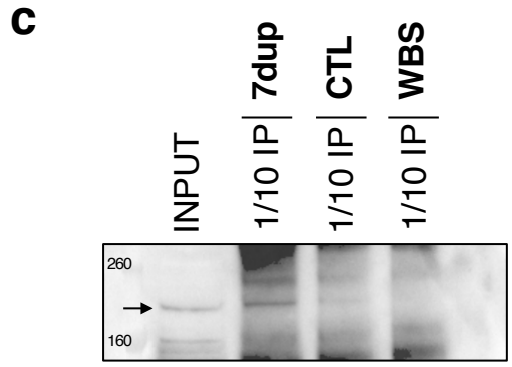


Figure 35. Second attempt of BAZ1B immunoprecipitation using an antibody produced by Abcam

a) Coomassie blue staining and Western Blot analysis of BAZ1B IP in a 7dup NCSC line. **b**) Relevant interactors according to Mass Spectrometry analysis of the immunoprecipitates showed in panel a. **c**) Western Blot analysis of BAZ1B IP in NSCSs from the three genetic conditions.

Next, using the very same experimental conditions, I performed BAZ1B IP in all the three genetic conditions (including the 7dup sample tested as a pilot). As before, 1/10 of the IP from each condition was assayed in western blot as internal controls (**Figure 35c**). Unexpectedly, BAZ1B was not detected in any of the samples according to MS analysis. Immunoprecipitation and analysis were repeated a second time on fresh protein extracts but, again, I could not obtain any enrichment for the bait.

The difficulties in BAZ1B enrichment can be attributed to:

1. the antibody efficiency. Any of the antibodies tested allowed good enrichments of the bait, even though both have already been tested either in other cell types or in previous published works from other groups (Almuzzaini et al., 2016).
2. the inefficient optimization of all the steps required for the identification of the bait and its interactors. I tested different protocols and IP/wash buffers during the different steps of immunoprecipitation; while sample processing was always performed in the same way by the internal Mass Spectrometry Unit.

3. very low BAZ1B expression levels in patient derived-NCSCs.

Overexpression of BAZ1B in our cell lines can't be considered a possible option since our main aim is to investigate BAZ1B dosage-dependent alterations. Thus, we resorted to a targeted genome editing strategy that takes advantage of the CRISPR/Cas9 system to add a specific tag to the endogenous BAZ1B gene sequence. I am currently investigating the best tag option to identify BAZ1B interactors, working with high efficiency both in IP and ChIP experiments. Preliminary maps of two possible donor plasmids are reported in **Figure 36**.

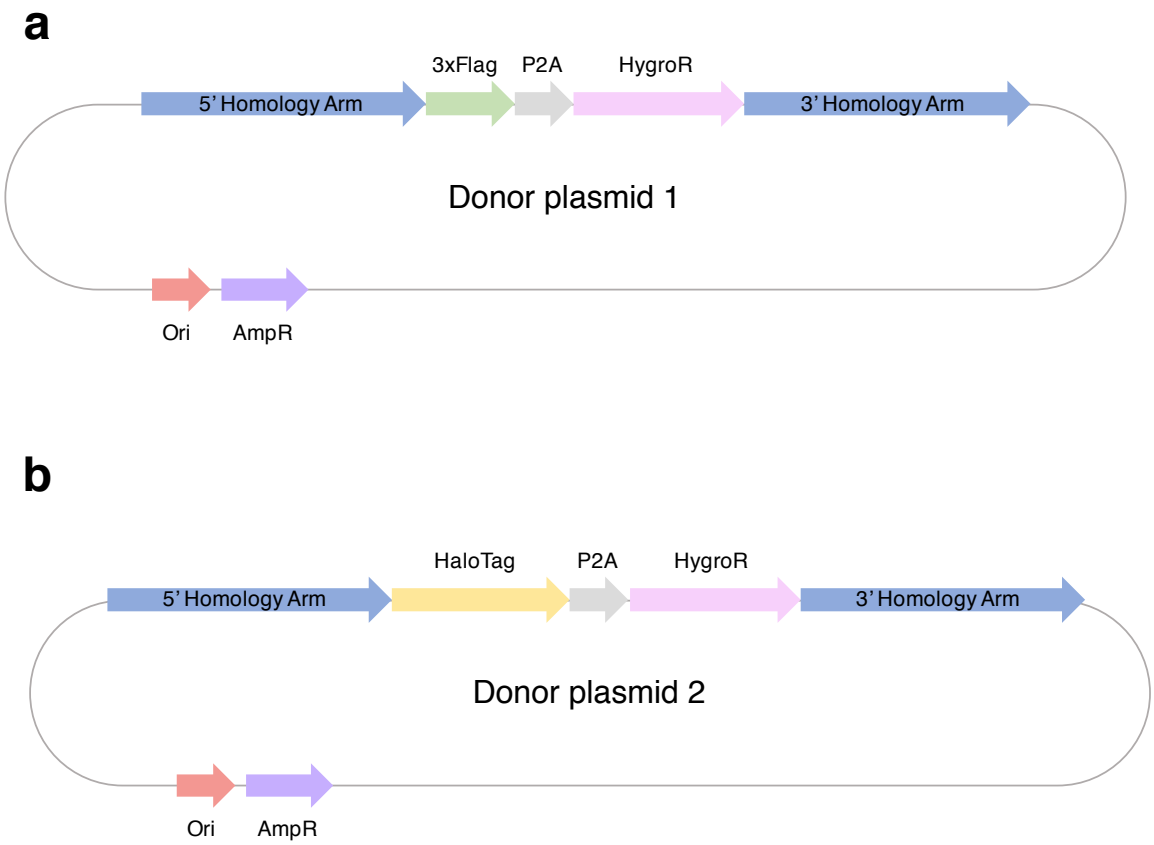


Figure 36. Schematic representation of two alternative donor plasmids to be used in CRISPR/Cas9-mediated BAZ1B tagging

a) The donor plasmid will contain either **a**) three flag tag or **b**) a Halotag followed by a self-cleaving peptide (P2A) and a Hygromycin resistance (HygroR). The Tag-P2A-HygroR cassette will be flanked by BAZ1B specific Homology Arms to promote homologous recombination and then subcloned into a bacterial backbone.

5. DISCUSSION

In this thesis, I have investigated the molecular basis of NC-related alterations typical of both WBS and 7dup patients, taking advantage of NCSCs directly derived from patient-specific iPSCs (Adamo et al., 2015).

The association of NC/NC-derived cells defects with brain alterations is a constant of most neurodevelopmental disorders, given the fact that both tissues derive from ectoderm and the genetic causes are indeed rooted at very early stages of development.

This explains why it is extremely crucial to focus on progenitor cells to have better insights into the molecular mechanisms responsible for this type of diseases. The recent possibilities of reproducing *in vitro* poorly accessible cell types/tissues and to specifically explore the patient genetic information, represent extremely advanced tools in the study of neurodevelopmental disorders.

WBS and 7dup syndrome constitute a very relevant example of the coexistence of NC-related craniofacial malformations and central nervous system (CNS)-related cognitive and social alterations, as extensively discussed in chapter 2.1.

5.1. *Bona fide* multipotent NCSCs can be derived from patient specific-iPSC

Taking advantage of WBS and 7dup patient-derived iPSCs and of a highly efficient differentiation protocol we generated at least three NCSC lines per each genetic condition, including healthy controls. To our knowledge, this is the largest cohort of patient-specific NCSCs described so far.

We did not observe any differences in the efficiency of differentiation of both patient and control iPSCs and we were able to obtain a high percentage of *bona fide* NCSCs in all the

conditions, indicating that iPSC-to-NCSC differentiation ability is not impaired in WBS and 7dup patients. Moreover, NCSCs from patient and control individuals are indistinguishable in their morphology and they all displayed self-renewal ability, even after a high number of passages in culture (>40). Multipotency is also maintained both in WBS/7dup and control NCSCs, as shown by the successful generation of three distinct NC-derivatives (*i.e.,* MSCs, osteocytes and SMCs). In a sense, this is expected, since both types of patients are able to develop bone and facial structures, as well as smooth muscle in vessels or neurons in the peripheral nervous system.

Given these evidences, I hypothesized that 7q11.23 gene dosage imbalances do not impair the cell-specific differentiation potential and the differentiation process itself, but most probably affect NC and NC-derived cell function and the regulation of their generation and distribution in the developing embryo. NCSCs are indeed known to migrate in specific areas of the embryo and to give rise to their derivatives upon an extremely tight regulation of numerous signaling pathways (chapter 2.2.1.). It is reasonable to think that WBSCR genes mediated-alterations of these pathways might lead either to i) an over/underproduction of NC or NC-derived cells and/or ii) to migration defects and, consequently, to an uncontrolled deposition in the target regions of the developing embryo. Indeed, SVAS in WBS patients is caused by an overgrowth of smooth muscle cells in arteries, thus supporting the overproduction hypothesis (Pober, 2010). Moreover, some symmetrically opposite features of WBS and 7dup patients, such as the philtrum length and the lip size, could be explained by a gene dosage-dependent defective cell deposition.

5.2. iPSC derived-NCSC transcriptome is dysregulated in pathways involved in NC development

The presence of several important transcriptional regulators within the 7q11.23 region, suggests transcription to be significantly impaired in WBS and 7dup patients. We were able to prove this and, more interestingly, we showed that dysregulation at the RNA level was already visible at very early stages of development, *i.e.,* the pluripotent state, and pointed to a series of biological processes known to determine the patient clinical phenotype (Adamo et al., 2015).

This impairment is then maintained upon differentiation into NSCSs, MSCs and neural progenitors (NPCs) ((Adamo et al., 2015) and chapter 4.2) but also amplified in a lineage-specific manner, so that disease-relevant pathways ascribable to a particular differentiated lineage are progressively and preferentially more affected. For example, while at the iPSC stage, the dysregulated genes encompass a wide range of processes relevant for the diseases, in NPCs DEGs are specifically enriched for categories involving neuronal function and development (Adamo et al., 2015). In the same way, microarray profiling of NCSC transcriptomes shows altered expression of genes specifically implicated in NC development and NC-derived tissues including craniofacial structures, peripheral neurons and smooth muscle. Notably, it was shown that the enhancement of the small GTPase RhoA signaling, that resulted dysregulated in our NCSC dataset, could rescue the poorly organized networks of actin bundles in smooth muscle cells from patients with SVAS (Adamo et al., 2015; Ge et al., 2012).

5.3. BAZ1B KD affects patient-specific NCSC migration

A plethora of evidences from different models indicate BAZ1B involvement in the acquisition of NC-dependent alterations.

In particular, results from scratch assays indicate a clear impairment of migration upon BAZ1B KD in NCSCs from at least 4 different individuals (*i.e.*, 1 7dup and 3 controls). NCSC migration in the WBS sample appears already affected in the 'scr' line, when compared to the 'scr' lines from other individuals, while the KD in the same line doesn't have relevant consequences. Thus, I hypothesized that a lowering of BAZ1B levels in cells in which its expression is already low and where probably cell motility is already altered, might not further affect migration, but replicates are needed.

NC cells represent a typical model of 'contact inhibition of locomotion', as shown by both *in vitro* and *in vivo* experiment (Carmona-Fontaine et al., 2008). This means that when two NC cells enter in contact, they tend to collapse their protrusions and to establish connections, leading both to a change in direction and to a reduction of their velocity. Thus, BAZ1B dosage alterations might impair this specific process, given our observations on NCSC migration. In line with this, the analysis of NCSC random migration suggests a reduction of cell velocity and, consequently, of accumulated distance upon KD; while results on cell directionality are still difficult to interpret because of a high variability within cells (data not shown). The automated tracking that I am currently setting up, will allow the measurement of a significantly higher number of cells thus, possibly, reducing variability.

Notably, BAZ1B-mediated positive regulation of migration, has been previously described in NC cells from *Xenopus laevis* (Barnett et al., 2012) and in human lung cancer cells (Meng et al., 2016). However, the first evidence was only based on the observed dysregulation of genes involved in NC cell migration upon BAZ1B KD and no migration-specific assays were performed. In the second case, the accelerated migration and aggressiveness due to BAZ1B overexpression were also confirmed with pertinent experiments (*i.e.*, scratch and invasion assays) but these observations were anyway limited to cancerous cells.

5.4. BAZ1B KD alters NCSC transcriptome in a specific subset of biological processes

RNA interference-mediated reduction of BAZ1B levels in NCSCs, aimed at dissecting its specific contribution to the disease, had a limited but still relevant impact on NCSC transcriptomes, in line with previous reports describing BAZ1B as a key transcriptional regulator (Poot et al., 2004).

Notably, the strength of our analysis resides in considering the two shRNA against BAZ1B not only as a single entity (pairwise comparative analysis, chapter 4.3.3.1) but also as two distinct ones (multifactorial comparative analysis, chapter 4.3.3.2), given the fact that their efficiencies in knocking down BAZ1B expression levels were proven to be different. When gene dosage is altered in regions that are extremely sensitive to dosage variations, such as the 7q11.23 region, we also expect to have specific consequences that are dependent on the entity of dysregulation. This is why we took in consideration the different KD efficiencies of the two shRNA and we assigned numerical variables to each to each shRNA group. Thanks to this approach, we could identify a subset of common genes, whose expression is affected upon both shRNA-mediated KD, and a group of genes only dysregulated upon a more substantial alteration of BAZ1B levels.

In the last more sophisticated analysis, we specifically looked at genes whose expression was dependent on BAZ1B level variations and we could identify a higher number of genes dysregulated in both shBAZ1B groups. Moreover, this analysis confirmed the stronger transcriptional impact of sh41, and, at the same time, was able to detect all those genes whose expression was affected, even in a milder fashion, also upon the less efficient sh38-mediated KD. Given the stringency of this last approach, we considered the '279 BAZ1B-level sensitive genes' dataset as the most complete and reliable and we focused on this for the subsequent analysis.

Among these genes, we found *OLFM1*, a key regulator of embryonic nervous system development, to directly follow BAZ1B levels. Interestingly, its overexpression induces an excess of NC cells emigration from the neural tube (Barembaum et al., 2000), thus supporting my hypothesis of an uncontrolled NC cell production as a consequence of gene dosage imbalances.

An interesting gene that, according to our analysis, inversely follows BAZ1B levels, is *NR2F2*, better known as Chicken Ovalbumin Upstream Promoter Transcription Factor II (COUP-TFII). It is involved in the development of heart, blood vessels and muscles and its inactivation in mesenchymal progenitors promotes differentiation towards osteoblast and myoblast, while impairing the formation of adipocytes and chondrocytes (Xie et al., 2011). Thus, an aberrant deposition of bone and cartilage in facial areas or of smooth muscle cells in vessels, as a consequence of BAZ1B-mediated *NR2F2* dysregulation, could be possible explanations for craniofacial and vascular defects in WBS and 7dup patients. In line with this, NR2F2 and its partner NR2F1 were shown to modulate craniofacial development through the binding of NC-specific enhancer regions (Rada-Iglesias et al., 2012, 2013).

Moreover, NR2F2 promotes cancer cell migration and metastasis in several types of tumors (Qin et al., 2010), thus its involvement in the regulation of NC cell migration is plausible.

The repetition, in most cases, of the same small subset of genes (such as *POSTN*, *NR2F2*, *ERBB4*) throughout the different most enriched GO categories, supports the dysregulation of a limited group of biological processes upon *BAZ1B* KD, mainly encompassing vascular and skeletal development and cell migration. For example, the presence of several dysregulated genes involved in cardiovascular development suggests that ELN might not be the only causative gene for SVAS and aortic dilation in WBS and 7dup patients, respectively. A deep characterization of SMCs derived from *BAZ1B* KD NCSCs will possibly give some useful insights into cardiovascular defects, that, to date,

are the most common cause of death in WBS patients (Pober, 2010). Likewise, many genes within our dataset (*e.g.*, *POSTN*, *NR2F2*, *ERBB4*, *THSD7A* and *RAB3C*) play a crucial role in cell migration and for some of them there are evidences of a direct involvement in NC cell motility. This result suggests that the migration phenotype observed in our NCSCs might be caused by a dysregulation of pathways that control NC cell motility and adhesion, as a consequence of BAZ1B dosage imbalances.

5.5. BAZ1B-dosage imbalances impact on NCSC enhancer state

Our hypothesis, claiming an impact of BAZ1B dosage imbalances on NCSC putative cis-regulatory regions is partially supported by our preliminary results, specifically focused on putative active enhancer regions. A more accurate investigation of BAZ1B chromatin occupancy, through a BAZ1B-specific ChIP-seq, will shed more lights on its preferential binding sites.

Differential analysis on the distribution of H3K27ac, H3K4me1 and H3K27me3 marks in the selected group of regions, show a limited role for BAZB1 in the dynamics of these marks, since only few regions were identified as differentially marked based on BAZ1B levels. In particular, the most affected mark is H3K27me3, given the higher number of BAZ1B-level sensitive regions identified when compared to the other two marks. Moreover, only in this context, BAZ1B acts in a univocal manner, since almost all the regions differentially marked inversely follow its levels. Indeed, higher BAZ1B levels correspond to a lower presence of the repressive mark, in line with a BAZ1B positive role in transcriptional regulation (Poot et al., 2005).

On the contrary, the H3K4me1 and H3K27ac marks follow BAZ1B levels either directly or inversely in a similar number of regions, suggesting that i) other epigenetic remodelers

could contribute, together with BAZ1B, to the deposition/erasure of these marks and/or ii) a second, not yet described, negative role of BAZ1B in transcriptional regulation.
Moreover, BAZ1B involvement in migration is further supported by ChIP-seq analysis, specifically on H3K4 differentially monomethylated regions.

I hypothesized different explanations for the limited overlap between the genes associated with differential putative enhancer regions and those identified in the RNA-seq dataset.

1. Results from the two analyses display two distinct levels of transcriptional dysregulation, respectively involving or not involving enhancer remodeling. However, the identification of several interactions between proteins encoded by genes from the RNA-seq and the ChIP-seq dataset, show a preliminary connection between the apparently distinct outcomes of the two approaches.
2. BAZ1B-mediated remodeling of these regions is reflected in a milder alteration of enhancer-mediated gene expression, not detected in the RNA-seq analysis. However, mRNA expression levels of genes predicted to be associated with differential enhancers do not show a clear sensitivity to BAZ1B dosage (data not shown).
3. BAZ1B-mediated remodeling of these regions does not result in an impairment of enhancer-dependent gene expression, because other factors or gene promoters themselves are able to compensate.
4. GREAT prediction approach, exclusively based on the association of a peak with the closest gene body, is suboptimal and potentially inefficient.

5.6. Preliminary analysis suggests novel interactors for BAZ1B

The suboptimal identification of the bait doesn't allow us to draw final conclusions on BAZ1B interactors. However, the presence of a known BAZ1B partner (*i.e.*, SMARCA5) using two different antibodies, suggests not to completely exclude the results that we obtained. For example, the recognition of novel potential interactors, mostly involved in splicing-related processes (*e.g.*, different subunit of splicing factor 3B) or in the structural maintenance of chromosomes (*e.g.*, SMC1A, SMC3 and SMC4), might indicate unknown levels of BAZ1B-dependent regulation. Within splicing-related factors, we identified Eukaryotic Initiation Factor 4A3 (EIF4A3), encoding for a protein involved in alternative splicing and translation initiation, whose partial loss of function causes a craniofacial disorder known as Richieri-Costa-Pereira syndrome (RCPS) (Favaro et al., 2014). Interestingly, it was recently shown that NC cells derived from RCPS patient-iPSCs have a reduced migration capacity when compared to control cells (Miller et al., 2017).

In addition, mutations in some subunits of the cohesion complex (such as SMC1A and SMC3) are responsible for a variant of Cornelia de Lange syndrome (CdLS), characterized by a predominant mental disability and milder craniofacial alterations (Deardorff et al., 2007). More recently, it was shown that these mutations impair transcription initiation and elongation and this might explain the transcriptional dysregulation observed in CdLS patients (Mannini et al., 2015).

Thus, we can hypothesize that the observed interactions between BAZ1B and EIF4A3 or cohesin subunits might be involved in the alterations in transcription or in the migration defects caused by BAZ1B KD, but further confirmation studies are needed.

5.7. Atypical patients are critical in the study of CNV-related disorders

In the future, we plan to further increase our patient cohort including, in particular, other atypical WBS patients with different types of smaller deletions, since they were already shown to represent extremely valuable tools, both in our and in previous works.

For example, the atypical WBS patient already included in our analysis, in which BAZ1B and few other genes are not deleted and that presents milder craniofacial alterations, further supports our hypothesis on BAZ1B role in NC regulation. In line with this, we observed that the transcriptional profile of NCSCs from the atypical patient associates closely to the profile of at least two control NCSCs and to none of the typical WBS NCSCs. This indicates that a relevant portion of the observed transcriptional dysregulation might be attributed to the few WBSCR genes that are still present in the atypical deletion, including BAZ1B.

In addition, a second WBS patient bearing a smaller deletion that spares both *BAZ1B* and *GTF2I*, also displayed a less severe craniofacial phenotype, but still presented some typical WBS facial features (Fusco et al., 2013). Moreover, a third peculiar atypical WBS patient in which only GTF2I and other few genes are deleted, did not showed typical behavioural problems but presented the typical WBS facies, including iris stellata, medial eyebrow flare, ptosis of the eyelids, fullness of the periorbital region, anteverted nares, prominent and long philtrum and full lips (Bert de Vries, unpublished).

These last observations indicate that *BAZ1B* is not acting alone in the regulation of facial development, but *GTF2I* and other WBSCR genes might be involved. In line with this, mice lacking *Gtf2i* are characterized by craniofacial defects (Enkhmandakh et al., 2009).

Thus, our hypothesis indicating a combined action of different WBSCR genes in the definition of NC-related defects, rather than a univocal attribution of a single gene to a specific clinical trait, is now supported by two main observations:

- BAZ1B possible involvement in patient cardiovascular defects together with ELN, as indicated by RNA-seq results;
- BAZ1B- and GTF2I-mediated regulation of the development of distinct craniofacial structures.

5.8. Final remarks

In conclusion, our observations highlight a multilevel *modus operandi* for BAZ1B in the regulation of WBS and 7dup patient-specific NCSCs, involving enhancer remodeling, transcriptional regulation and cell-type specific functions.

We showed that BAZ1B dosage imbalances are responsible for a transcriptional alteration involving migration-related genes and that this process is indeed impaired upon BAZ1B KD in our NCSC cohort.

In addition, BAZ1B dosage imbalances affect the distribution of histone marks (in particular H3K27me3) in a limited subset of putative active enhancer regions, but the entity and the consequences of the remodeling still need to be better elucidated and, thus, require further and more reliable analyses.

The identification of BAZ1B direct interactors and a better characterization of NCSCs derived from patients with atypical deletions (sparing or not BAZ1B), together with the results already presented in this thesis, will hopefully contribute to the reconstruction of the molecular dysregulation responsible for NC-related defects in WBS and 7dup patient-derived NCSCS.

6. REFERENCES

Van der Aa, N., Rooms, L., Vandeweyer, G., van den Ende, J., Reyniers, E., Fichera, M., Romano, C., Delle Chiaie, B., Mortier, G., Menten, B., et al. (2009). Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome. Eur. J. Med. Genet. *52*, 94–100.

Achilleos, A., and Trainor, P.A. (2012). Neural crest stem cells: discovery, properties and potential for therapy. Cell Res. *22*, 288–304.

Adamo, A., Atashpaz, S., Germain, P.-L., Zanella, M., D'Agostino, G., Albertin, V., Chenoweth, J., Micale, L., Fusco, C., Unger, C., et al. (2015). 7q11.23 dosage-dependent dysregulation in human pluripotent stem cells affects transcriptional programs in disease-relevant lineages. Nat. Genet. *47*, 132–141.

Almuzzaini, B., Sarshad, A.A., Rahmanto, A.S., Hansson, M.L., Von Euler, A., Sangfelt, O., Visa, N., Farrants, A.-K.O., and Percipalle, P. (2016). In -actin knockouts, epigenetic reprogramming and rDNA transcription inactivation lead to growth and proliferation defects. FASEB J. *30*, 2860–2873.

Antonell, A., Campo, M. Del, Magano, L.F., Kaufmann, L., Martínez De La Iglesia, J., Gallastegui, F., Flores, R., Schweigmann, U., Fauth, C., Kotzot, D., et al. Partial 7q11.23 deletions further implicate GTF2I and GTF2IRD1 as the main genes responsible for the WilliamseBeuren syndrome neurocognitive profile.

Antonell, A., Vilardell, M., and Pérez Jurado, L.A. (2010). Transcriptome profile in Williams-Beuren syndrome lymphoblast cells reveals gene pathways implicated in glucose intolerance and visuospatial construction deficits. Hum. Genet. *128*, 27–37.

Arya, M., Srinivasan, M., and Rajasekharan, R. (2017). Human alpha beta hydrolase domain containing protein 11 and its yeast homolog are lipid hydrolases. Biochem. Biophys. Res. Commun. *487*, 875–880.

Ashe, A., Morgan, D.K., Whitelaw, N.C., Bruxner, T.J., Vickaryous, N.K., Cox, L.L., Butterfield, N.C., Wicking, C., Blewitt, M.E., Wilkins, S.J., et al. (2008). A genome-wide

screen for modifiers of transgene variegation identifies genes with critical roles in development. Genome Biol. *9*, R182.

Axelrod, F.B., Liebes, L., Gold-Von Simson, G., Mendoza, S., Mull, J., Leyne, M., Norcliffe-Kaufmann, L., Kaufmann, H., and Slaugenhaupt, S.A. (2011). Kinetin improves IKBKAP mRNA splicing in patients with familial dysautonomia. Pediatr. Res. *70*, 480–483.

Aydin, Ö.Z., Marteijn, J.A., Ribeiro-Silva, C., Rodríguez López, A., Wijgers, N., Smeenk, G., van Attikum, H., Poot, R.A., Vermeulen, W., and Lans, H. (2014). Human ISWI complexes are targeted by SMARCA5 ATPase and SLIDE domains to help resolve lesion-stalled transcription. Nucleic Acids Res. *42*, 8473–8485.

Bajpai, R., Chen, D.A., Rada-Iglesias, A., Zhang, J., Xiong, Y., Helms, J., Chang, C.-P., Zhao, Y., Swigut, T., and Wysocka, J. (2010). CHD7 cooperates with PBAF to control multipotent neural crest formation. Nature *463*, 958–962.

Barembaum, M., Moreno, T.A., LaBonne, C., Sechrist, J., and Bronner-Fraser, M. (2000). Noelin-1 is a secreted glycoprotein involved in generation of the neural crest. Nat. Cell Biol. *2*, 219–225.

Barnett, C., and Krebs, J.E. (2011). WSTF does it all: a multifunctional protein in transcription, repair, and replication. Biochem. Cell Biol. *89*, 12–23.

Barnett, C., Yazgan, O., Kuo, H.-C., Malakar, S., Thomas, T., Fitzgerald, A., Harbour, W., Henry, J.J., and Krebs, J.E. (2012). Williams Syndrome Transcription Factor is critical for neural crest cell function in Xenopus laevis. Mech. Dev. *129*, 324–338.

Basch, M.L., Bronner-Fraser, M., and García-Castro, M.I. (2006). Specification of the neural crest occurs during gastrulation and requires Pax7. Nature *441*, 218–222.

Bayés, M., Magano, L.F., Rivera, N., Flores, R., and Pérez Jurado, L.A. (2003). Mutational mechanisms of Williams-Beuren syndrome deletions. Am. J. Hum. Genet. *73*, 131–151.

Beaudet, A.L. (2013). The utility of chromosomal microarray analysis in developmental and behavioral pediatrics. Child Dev. *84*, 121–132.

Begbie, J. (2013). Chapter 13 – Induction and Patterning of Neural Crest and Ectodermal Placodes and their Derivatives. In Patterning and Cell Type Specification in the

Developing CNS and PNS, pp. 239–258.

Bellin, M., Marchetto, M.C., Gage, F.H., and Mummery, C.L. (2012). Induced pluripotent stem cells: the new patient? Nat. Rev. Mol. Cell Biol.

Benish, B.M. (1975). The neurocristopathies: a unifying concept of disease arising in neural crest development. Hum. Pathol. *6*, 128.

Bennett, M.K., García-Arrarás, J.E., Elferink, L.A., Peterson, K., Fleming, A.M., Hazuka, C.D., and Scheller, R.H. (1993). The syntaxin family of vesicular transport receptors. Cell *74*, 863–873.

Beuren, A.J., Apitz, J., and Harmjanz, D. (1962). Supravalvular Aortic Stenosis in Association with Mental Retardation and a Certain Facial Appearance. Circulation *XXVI*, 1235–1240.

Bhatt, S., Diaz, R., and Trainor, P.A. (2013). Signals and switches in Mammalian neural crest cell differentiation. Cold Spring Harb. Perspect. Biol. *5*, a008326–a008326.

Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D., and Amit, I. (2013). High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nat. Protoc. *8*, 539–554.

Bozhenok, L., Wade, P.A., and Varga-Weisz, P. (2002). WSTF-ISWI chromatin remodeling complex targets heterochromatic replication foci. EMBO J. *21*, 2231–2241.

Brandl, C., Florian, C., Driemel, O., Weber, B.H.F., and Morsczeck, C. (2009). Identification of neural crest-derived stem cell-like cells from the corneal limbus of juvenile mice. Exp. Eye Res. *89*, 209–217.

Bronner, M.E., and LeDouarin, N.M. (2012). Development and evolution of the neural crest: An overview. Dev. Biol. *366*, 2–9.

Buecker, C., and Wysocka, J. (2012). Enhancers as information integration hubs in development: lessons from genomics. Trends Genet. *28*, 276–284.

Burden, S., and Yarden, Y. (1997). Neuregulins and their receptors: a versatile signaling module in organogenesis and oncogenesis. Neuron *18*, 847–855.

Burn, J. (1986). Syndrome of the month Williams syndrome. J. Med. Genet. *23*, 389–395.

Cagle, A.P., Waguespack, S.G., Buckingham, B.A., Shankar, R.R., and Dimeglio, L.A. (2004). Severe infantile hypercalcemia associated with Williams syndrome successfully treated with intravenously administered pamidronate. Pediatrics *114*, 1091–1095.

Cairo, S., Merla, G., Urbinati, F., Ballabio, A., and Reymond, A. (2001). WBSCR14, a gene mapping to the Williams--Beuren syndrome deleted region, is a new member of the Mlx transcription factor network. Hum. Mol. Genet. *10*, 617–627.

Capossela, S., Muzio, L., Bertolo, A., Bianchi, V., Dati, G., Chaabane, L., Godi, C., Politi, L.S., Biffo, S., D'Adamo, P., et al. (2012). Growth defects and impaired cognitivebehavioral abilities in mice with knockout for Eif4h, a gene located in the mouse homolog of the williams-beuren syndrome critical region. Am. J. Pathol.

Carmona-Fontaine, C., Matthews, H.K., Kuriyama, S., Moreno, M., Dunn, G.A., Parsons, M., Stern, C.D., and Mayor, R. (2008). Contact inhibition of locomotion in vivo controls neural crest directional migration. Nature *456*, 957–961.

Cavellán, E., Asp, P., Percipalle, P., and Farrants, A.-K.O. (2006). The WSTF-SNF2h chromatin remodeling complex interacts with several nuclear proteins in transcription. J. Biol. Chem. *281*, 16264–16271.

Chai, Y., Jiang, X., Ito, Y., Bringas, P., Han, J., Rowitch, D.H., Soriano, P., McMahon, A.P., and Sucov, H.M. (2000). Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. Development *127*, 1671–1679.

Chailangkarn, T., Trujillo, C.A., Freitas, B.C., Hrvoj-Mihic, B., Herai, R.H., Yu, D.X., Brown, T.T., Marchetto, M.C., Bardy, C., McHenry, L., et al. (2016). A human neurodevelopmental model for Williams syndrome. Nature.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat. Biotechnol. *27*, 275–280.

Chandler, R.L., and Magnuson, T. (2016). The SWI/SNF BAF-A complex is essential for neural crest development. Dev. Biol. *411*, 15–24.

Chen, P., Chen, J., Shao, C., Li, C., Zhang, Y., Lu, W., Fu, Y., Gu, P., and Fan, X. (2014). Treatment with retinoic acid and lens epithelial cell-conditioned medium in vitro directed the differentiation of pluripotent stem cells towards corneal endothelial cell-like cells. Exp. Ther. Med. *9*, 351–360.

Cheung, C., Bernardo, A.S., Pedersen, R.A., and Sinha, S. (2014). Directed differentiation of embryonic origin–specific vascular smooth muscle subtypes from human pluripotent stem cells. Nat. Protoc. *9*, 929–938.

Cheung, M., Chaboissier, M.-C., Mynett, A., Hirst, E., Schedl, A., and Briscoe, J. (2005). The transcriptional control of trunk neural crest induction, survival, and delamination. Dev. Cell *8*, 179–192.

Corley, S.M., Canales, C.P., Carmona-Mora, P., Mendoza-Reinosa, V., Beverdam, A., Hardeman, E.C., Wilkins, M.R., and Palmer, S.J. (2016). RNA-Seq analysis of Gtf2ird1 knockout epidermal tissue provides potential insights into molecular mechanisms underpinning Williams-Beuren syndrome. BMC Genomics *17*, 450.

Crackower, M.A., Kolas, N.K., Noguchi, J., Sarao, R., Kikuchi, K., Kaneko, H., Kobayashi, E., Kawai, Y., Kozieradzki, I., Landers, R., et al. (2003). Essential role of Fkbp6 in male fertility and homologous chromosome pairing in meiosis. Science *300*, 1291–1295.

Culler, F.L., Jones, K.L., and Deftos, L.J. (1985). Imparied calcitonin secretion in patients with Williams syndrome. J. Pediatr. *107*, 720–723.

Culver-Cochran, A.E., and Chadwick, B.P. (2013). Loss of WSTF results in spontaneous fluctuations of heterochromatin formation and resolution, combined with substantial changes to gene expression. BMC Genomics *14*, 740.

Curran, M.E., Atkinson, D.L., Ewart, A.K., Morris, C.A., Leppert, M.F., and Keating, M.T. (1993). The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis. Cell *73*, 159–168.

Cus, R., Maurus, D., and Kühl, M. (2006). Cloning and developmental expression of WSTF during Xenopus laevis embryogenesis.

Davies, M., Udwin, O., and Howlin, P. (1998). Adults with Williams syndrome. Preliminary

study of social, emotional and behavioural difficulties. Br. J. Psychiatry *172*, 273–276.

Deardorff, M.A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S., Pie, J., Gil-Rodríguez, C., Arnedo, M., Loeys, B., Kline, A.D., et al. (2007). Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. Am. J. Hum. Genet. *80*, 485–494.

Deruelle, C., Schön, D., Rondan, C., and Mancini, J. (2005). Global and local music perception in children with Williams syndrome. Neuroreport *16*, 631–634.

Dirscherl, S.S., Henry, J.J., and Krebs, J.E. (2005). Neural and eye-specific defects associated with loss of the Imitation Switch (ISWI) chromatin remodeler in Xenopus laevis. Mech. Dev. *122*, 1157–1170.

Doerks, T., Copley, R., and Bork, P. (2001). DDT -- a novel domain in different transcription and chromosome remodeling factors. Trends Biochem. Sci. *26*, 145–146.

Elçioglu, N., Mackie-Ogilvie, C., Daker, M., and Berry, A.C. (1998). FISH analysis in patients with clinical diagnosis of Williams syndrome. Acta Paediatr. *87*, 48–53.

Enkhmandakh, B., Makeyev, A. V, Erdenechimeg, L., Ruddle, F.H., Chimge, N.-O., Tussie-Luna, M.I., Roy, A.L., and Bayarsaihan, D. (2009). Essential functions of the Williams-Beuren syndrome-associated TFII-I genes in embryonic development. Proc. Natl. Acad. Sci. U. S. A. *106*, 181–186.

Etchevers, H.C., Amiel, J., and Lyonnet, S. (2006). Molecular bases of human neurocristopathies. Adv. Exp. Med. Biol. *589*, 213–234.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Ewart, A.K., Morris, C.A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., Stock, A.D., Leppert, M., and Keating, M.T. (1993). Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. Nat. Genet. *5*, 11–16.

Fattah, F.J., Hara, K., Fattah, K.R., Yang, C., Wu, N., Warrington, R., Chen, D.J., Zhou, P., Boothman, D.A., and Yu, H. (2014). The transcription factor TFII-I promotes DNA translesion synthesis and genomic stability. PLoS Genet. *10*, e1004419.

Faury, G., Pezet, M., Knutsen, R.H., Boyle, W.A., Heximer, S.P., McLean, S.E., Minkes, R.K., Blumer, K.J., Kovacs, A., Kelly, D.P., et al. (2003). Developmental adaptation of the mouse cardiovascular system to elastin haploinsufficiency. J. Clin. Invest. *112*, 1419–1428.

Favaro, F.P., Alvizi, L., Zechi-Ceide, R.M., Bertola, D., Felix, T.M., de Souza, J., Raskin, S., Twigg, S.R.F., Weiner, A.M.J., Armas, P., et al. (2014). A Noncoding Expansion in EIF4A3 Causes Richieri-Costa-Pereira Syndrome, a Craniofacial Disorder Associated with Limb Defects. Am. J. Hum. Genet. *94*, 120–128.

Fort, P., Guémar, L., Vignal, E., Morin, N., Notarnicola, C., de Santa Barbara, P., and Faure, S. (2011). Activity of the RhoU/Wrch1 GTPase is critical for cranial neural crest cell migration. Dev. Biol. *350*, 451–463.

Francke, U., Hsieh, C.L., Foellmer, B.E., Lomax, K.J., Malech, H.L., and Leto, T.L. (1990). Genes for two autosomal recessive forms of chronic granulomatous disease assigned to 1q25 (NCF2) and 7q11.23 (NCF1). Am. J. Hum. Genet. *47*, 483–492.

Friedman, W.F., and Mills, L.F. (1969). The relationship between vitamin D and the craniofacial and dental anomalies of the supravalvular aortic stenosis syndrome. Pediatrics *43*, 12–18.

Friedman, W.F., and Roberts, W.C. (1966). Vitamin D and the supravalvar aortic stenosis syndrome. The transplacental effects of vitamin D on the aorta of the rabbit. Circulation *34*, 77–86.

Fujiwara, T., Sanada, M., Kofuji, T., and Akagawa, K. (2016). Unusual social behavior in HPC-1/syntaxin1A knockout mice is caused by disruption of the oxytocinergic neural system. J. Neurochem. *138*, 117–123.

Fuller, D.M., Zhu, M., Ou-Yang, C.-W., Sullivan, S.A., and Zhang, W. (2011). A tale of two TRAPs: LAT and LAB in the regulation of lymphocyte development, activation, and autoimmunity. Immunol. Res. *49*, 97–108.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai

virus, an RNA virus that does not integrate into the host genome. Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci. *85*, 348–362.

Fusco, C., Micale, L., Augello, B., Pellico, M.T., Menghini, D., Alfieri, P., Digilio, M.C., Mandriani, B., Carella, M., Palumbo, O., et al. (2013). Smaller and larger deletions of the Williams Beuren syndrome region implicate genes involved in mild facial phenotype, epilepsy and autistic traits. Eur. J. Hum. Genet. Adv. Online Publ.

Gagliardi, C., Frigerio, E., Burt, D.M., Cazzaniga, I., Perrett, D.I., and Borgatti, R. (2003). Facial expression recognition in Williams syndrome. Neuropsychologia *41*, 733–738.

Garcia, R.E., Friedman, W.F., Kaback, M.M., and Rowe, R.D. (1964). Idiopathic hypercalcemia and supravalvular aortic stenosis. Documentation of a new syndrome. N. Engl. J. Med. *271*, 117–120.

Garnett, A.T., Square, T.A., and Medeiros, D.M. (2012). BMP, Wnt and FGF signals are integrated through evolutionarily conserved enhancers to achieve robust expression of Pax3 and Zic genes at the zebrafish neural plate border. Development *139*, 4220–4231.

Gassmann, M., Casagranda, F., Orioli, D., Simon, H., Lai, C., Klein, R., and Lemke, G. (1995). Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. Nature *378*, 390–394.

Ge, X., Ren, Y., Bartulos, O., Lee, M.Y., Yue, Z., Kim, K.-Y., Li, W., Amos, P.J., Bozkulak, E.C., Iyer, A., et al. (2012). Modeling supravalvular aortic stenosis syndrome with human induced pluripotent stem cells. Circulation *126*, 1695–1704.

van der Geest, J.N., Lagers-van Haselen, G.C., van Hagen, J.M., Govaerts, L.C.P., de Coo, I.F.M., de Zeeuw, C.I., and Frens, M.A. (2004). Saccade dysmetria in Williams-Beuren syndrome. Neuropsychologia *42*, 569–576.

Geng, F.-S., Abbas, L., Baxendale, S., Holdsworth, C.J., Swanson, A.G., Slanchev, K., Hammerschmidt, M., Topczewski, J., and Whitfield, T.T. (2013). Semicircular canal morphogenesis in the zebrafish inner ear requires the function of gpr126 (lauscher), an adhesion class G protein-coupled receptor gene. Development *140*, 4362–4374.

Germain, P.-L., Vitriolo, A., Adamo, A., Laise, P., Das, V., and Testa, G. (2016).

RNAontheBENCH: computational and empirical resources for benchmarking RNAseq quantification and differential expression methods. Nucleic Acids Res. *44*, 5054–5067.

Gillan, L., Matei, D., Fishman, D.A., Gerbin, C.S., Karlan, B.Y., and Chang, D.D. (2002). Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility. Cancer Res. *62*, 5358–5364.

Golding, J.P., Trainor, P., Krumlauf, R., and Gassmann, M. (2000). Defects in pathfinding by cranial neural crest cells in mice lacking the neuregulin receptor ErbB4. Nat. Cell Biol. *2*, 103–109.

González, F., Boué, S., and Izpisúa Belmonte, J.C. (2011). Methods for making induced pluripotent stem cells: reprogramming à la carte. Nat. Rev. Genet. *12*, 231–242.

Gorovoy, M., Niu, J., Bernard, O., Profirovic, J., Minshall, R., Neamu, R., and Voyno-Yasenetskaya, T. (2005). LIM kinase 1 coordinates microtubule stability and actin polymerization in human endothelial cells. J. Biol. Chem. *280*, 26533–26542.

Guenat, D., Merla, G., Deconinck, E., Borg, C., and Rohrlich, P. (2017). DNA damage response defect in Williams-Beuren syndrome. Int. J. Mol. Med.

Günzel, D., and Yu, A.S.L. (2013). Claudins and the modulation of tight junction permeability. Physiol. Rev. *93*, 525–569.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. *10*, 622–640.

Haag, S., Kretschmer, J., and Bohnsack, M.T. (2015). WBSCR22/Merm1 is required for late nuclear pre-ribosomal RNA processing and mediates N7-methylation of G1639 in human 18S rRNA. RNA *21*, 180–187.

van Hagen, J.M., Eussen, H.J.F.M.M., van Schooten, R., van Der Geest, J.N., Lagers-van Haselen, G.C., Wouters, C.H., De Zeeuw, C.I., and Gille, J.J.P. (2007). Comparing Two Diagnostic Laboratory Tests for Williams Syndrome: Fluorescent *In Situ* Hybridization versus Multiplex Ligation-Dependent Probe Amplification. Genet. Test. *11*, 321–327.

Hahn, Y., and Lee, B. (2005). Identification of nine human-specific frameshift mutations by comparative analysis of the human and the chimpanzee genome sequences.

Bioinformatics *21 Suppl 1*, i186-94.

Hall, B.K. (2000). The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. Evol. Dev. *2*, 3–5.

Heilmann, A., Schinke, T., Bindl, R., Wehner, T., Rapp, A., Haffner-Luntzer, M., Nemitz, C., Liedert, A., Amling, M., and Ignatius, A. (2013). The Wnt serpentine receptor Frizzled-9 regulates new bone formation in fracture healing. PLoS One *8*, e84232.

His, W. (1868). Untersuchungen uber die erste Anlage des Wirbeltierleibes. Leipzig: Die erste Entwickelung des Hühnchens im Ei (FCW Vogel).

Ho, H.-K. V, Jang, J.J., Kaji, S., Spektor, G., Fong, A., Yang, P., Hu, B.S., Schatzman, R., Quertermous, T., and Cooke, J.P. (2004). Developmental endothelial locus-1 (Del-1), a novel angiogenic protein: its role in ischemia. Circulation *109*, 1314–1319.

Hockenhull, E.L., Carette, M.J., Metcalfe, K., Donnai, D., Read, A.P., and Tassabehji, M. (1999). A complete physical contig and partial transcript map of the Williams syndrome critical region. Genomics *58*, 138–145.

Hoogenraad, C.C., Eussen, B.H., Langeveld, A., van Haperen, R., Winterberg, S., Wouters, C.H., Grosveld, F., De Zeeuw, C.I., and Galjart, N. (1998). The murine CYLN2 gene: genomic organization, chromosome localization, and comparison to the human gene that is located within the 7q11.23 Williams syndrome critical region. Genomics *53*, 348–358.

Hoogenraad, C.C., Koekkoek, B., Akhmanova, A., Krugers, H., Dortland, B., Miedema, M., van Alphen, A., Kistler, W.M., Jaegle, M., Koutsourakis, M., et al. (2002). Targeted mutation of Cyln2 in the Williams syndrome critical region links CLIP-115 haploinsufficiency to neurodevelopmental abnormalities in mice. Nat. Genet. *32*, 116–127.

Howard, M.L., Palmer, S.J., Taylor, K.M., Arthurson, G.J., Spitzer, M.W., Du, X., Pang, T.Y.C., Renoir, T., Hardeman, E.C., and Hannan, A.J. (2012). Mutation of Gtf2ird1 from the Williams-Beuren syndrome critical region results in facial dysplasia, motor dysfunction, and altered vocalisations. Neurobiol. Dis.

Hsu, H., Lacey, D.L., Dunstan, C.R., Solovyev, I., Colombero, A., Timms, E., Tan, H.L.,

Elliott, G., Kelley, M.J., Sarosi, I., et al. (1999). Tumor necrosis factor receptor family member RANK mediates osteoclast differentiation and activation induced by osteoprotegerin ligand. Proc. Natl. Acad. Sci. U. S. A. *96*, 3540–3545.

Ichiyanagi, T., Ichiyanagi, K., Ogawa, A., Kuramochi-Miyagawa, S., Nakano, T., Chuma, S., Sasaki, H., and Udono, H. (2014). HSP90α plays an important role in piRNA biogenesis and retrotransposon repression in mouse. Nucleic Acids Res. *42*, 11903–11911.

Iizuka, K., Bruick, R.K., Liang, G., Horton, J.D., and Uyeda, K. (2004). Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis. Proc. Natl. Acad. Sci. U. S. A. *101*, 7281–7286.

Ito, T., Levenstein, M.E., Fyodorov, D. V, Kutach, A.K., Kobayashi, R., and Kadonaga, J.T. (1999). ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. Genes Dev. *13*, 1529–1539.

Jain, R.K. (2003). Molecular regulation of vessel maturation. Nat. Med. *9*, 685–693.

Janebodin, K., Horst, O. V, Ieronimakis, N., Balasundaram, G., Reesukumal, K., Pratumvinit, B., and Reyes, M. (2011). Isolation and characterization of neural crest-derived stem cells from dental pulp of neonatal mice. PLoS One *6*, e27526.

Jiang, X., Rowitch, D.H., Soriano, P., McMahon, A.P., and Sucov, H.M. (2000). Fate of the mammalian cardiac neural crest. Development *127*, 1607–1616.

Jiang, X., Gwye, Y., McKeown, S.J., Bronner-Fraser, M., Lutzko, C., and Lawlor, E.R. (2009). Isolation and Characterization of Neural Crest Stem Cells Derived From In Vitro–Differentiated Human Embryonic Stem Cells. Stem Cells Dev. *18*, 1059–1071.

Jiao, J., Xiong, W., Wang, L., Yang, J., Qiu, P., Hirai, H., Shao, L., Milewicz, D., Chen, Y.E., and Yang, B. (2016). Differentiation defect in neural crest-derived smooth muscle cells in patients with aortopathy associated with bicuspid aortic valves. EBioMedicine *10*, 282–290.

Jones, M.H., Hamana, N., Nezu, J., and Shimane, M. (2000). A Novel Family of Bromodomain Genes. Genomics *63*, 40–45.

Jordan-Sciutto, K.L., Dragich, J.M., Rhodes, J.L., and Bowser, R. (1999). Fetal Alz-50 clone 1, a novel zinc finger protein, binds a specific DNA sequence and acts as a transcriptional regulator. J. Biol. Chem. *274*, 35262–35268.

Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P., and Udwin, O. (1997). Language and Williams syndrome: how intact is "intact"? Child Dev. *68*, 246–262.

Katzmann, D.J., Babst, M., and Emr, S.D. (2001). Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell *106*, 145–155.

Kawano, T., Zhu, M., Troiano, N., Horowitz, M., Bian, J., Gundberg, C., Kolodziejczak, K., and Insogna, K. (2013). LIM kinase 1 deficient mice have reduced bone mass. Bone.

Knight, R.D., Nair, S., Nelson, S.S., Afshar, A., Javidan, Y., Geisler, R., Rauch, G.-J., and Schilling, T.F. (2003). lockjaw encodes a zebrafish tfap2a required for early neural crest development. Development *130*, 5755–5768.

Kruger, G.M., Mosher, J.T., Bixby, S., Joseph, N., Iwashita, T., and Morrison, S.J. (2002). Neural crest stem cells persist in the adult gut but undergo changes in self-renewal, neuronal subtype potential, and factor responsiveness. Neuron *35*, 657–669.

Lai, F.P.-L., Lau, S.-T., Wong, J.K.-L., Gui, H., Wang, R.X., Zhou, T., Lai, W.H., Tse, H.-F., Tam, P.K.-H., Garcia-Barcelo, M.-M., et al. (2017). Correction of Hirschsprung-associated Mutations in Human Induced Pluripotent Stem Cells, via CRISPR/Cas9, Restores Neural Crest Cell Function. Gastroenterology.

Lalli, M.A., Jang, J., Park, J.-H.C., Wang, Y., Guzman, E., Zhou, H., Audouard, M., Bridges, D., Tovar, K.R., Papuc, S.M., et al. (2016). Haploinsufficiency of BAZ1B contributes to Williams syndrome through transcriptional dysregulation of neurodevelopmental pathways. Hum. Mol. Genet. *25*, 1294–1306.

Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A., et al. (2009). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. Nature *461*, 402–406.

Leyfer, O.T., Woodruff-Borden, J., Klein-Tasman, B.P., Fricke, J.S., and Mervis, C.B. (2006). Prevalence of psychiatric disorders in 4 to 16-year-olds with Williams syndrome. Am. J. Med. Genet. B. Neuropsychiatr. Genet. *141B*, 615–622.

Li, D.Y., Brooke, B., Davis, E.C., Mecham, R.P., Sorensen, L.K., Boak, B.B., Eichwald, E., and Keating, M.T. (1998a). Elastin is an essential determinant of arterial morphogenesis. Nature *393*, 276–280.

Li, D.Y., Faury, G., Taylor, D.G., Davis, E.C., Boyle, W.A., Mecham, R.P., Stenzel, P., Boak, B., and Keating, M.T. (1998b). Novel arterial pathology in mice and humans hemizygous for elastin. J. Clin. Invest. *102*, 1783–1787.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91–95.

Li, H.H., Roy, M., Kuscuoglu, U., Spencer, C.M., Halm, B., Harrison, K.C., Bayle, J.H., Splendore, A., Ding, F., Meltzer, L.A., et al. (2009). Induced chromosome deletions cause hypersociability and other features of Williams-Beuren syndrome in mice. EMBO Mol. Med. *1*, 50–65.

Li, J., Sarosi, I., Yan, X.Q., Morony, S., Capparelli, C., Tan, H.L., McCabe, S., Elliott, R., Scully, S., Van, G., et al. (2000). RANK is the intrinsic hematopoietic cell surface receptor that controls osteoclastogenesis and regulation of bone mass and calcium metabolism. Proc. Natl. Acad. Sci. U. S. A. *97*, 1566–1571.

Liu, J.A., and Cheung, M. (2016). Neural crest stem cells and their potential therapeutic applications. Dev. Biol. *419*, 199–216.

Lu, X., Meng, X., Morris, C.A., and Keating, M.T. (1998). A Novel Human Gene,WSTF,Is Deleted in Williams Syndrome. Genomics *54*, 241–249.

Lundqvist, J., Kirkegaard, T., Laenkholm, A.-V., Duun-Henriksen, A.K., Bak, M., Feldman, D., and Lykkesfeldt, A.E. (2017). Williams syndrome transcription factor (WSTF) acts as an activator of estrogen receptor signaling in breast cancer cells and the effect can be abrogated by 1α,25-dihydroxyvitamin D 3. J. Steroid Biochem. Mol. Biol.

M Hao, M. (2016). Development of Neural Activity in the Enteric Nervous System: Similarities and Differences to Other Parts of the Nervous System. Adv. Exp. Med. Biol. *891*, 43–51.

Majesky, M.W. (2007). Developmental Basis of Vascular Smooth Muscle Diversity. Arterioscler. Thromb. Vasc. Biol. *27*, 1248–1258.

Makeyev, A. V, Erdenechimeg, L., Mungunsukh, O., Roth, J.J., Enkhmandakh, B., Ruddle, F.H., and Bayarsaihan, D. (2004). GTF2IRD2 is located in the Williams-Beuren syndrome critical region 7q11.23 and encodes a protein with two TFII-I-like helix-loop-helix repeats. Proc. Natl. Acad. Sci. U. S. A. *101*, 11052–11057.

Mancilla, A., and Mayor, R. (1996). Neural crest formation in Xenopus laevis: mechanisms of Xslug induction. Dev. Biol. *177*, 580–589.

Mannini, L., C Lamaze, F., Cucco, F., Amato, C., Quarantotti, V., Rizzo, I.M., Krantz, I.D., Bilodeau, S., and Musio, A. (2015). Mutant cohesin affects RNA polymerase II regulation in Cornelia de Lange syndrome. Sci. Rep. *5*, 16803.

Marintchev, A., Edmonds, K.A., Marintcheva, B., Hendrickson, E., Oberer, M., Suzuki, C., Herdy, B., Sonenberg, N., and Wagner, G. (2009). Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. Cell *136*, 447–460.

Martens, M.A., Wilson, S.J., and Reutens, D.C. (2008). Research Review: Williams syndrome: A critical review of the cognitive, behavioral, and neuroanatomical phenotype. J. Child Psychol. Psychiatry Allied Discip.

Mass, E., and Belostoky, L. (1993). Craniofacial Morphology of Children with Williams Syndrome. Cleft Palate-Craniofacial J. *30*, 343–349.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Menendez, L., Yatskievych, T.A., Antin, P.B., and Dalton, S. (2011). Wnt signaling and a Smad pathway blockade direct the differentiation of human pluripotent stem cells to multipotent neural crest cells. Proc. Natl. Acad. Sci. U. S. A. *108*, 19240–19245.

Menendez, L., Kulik, M.J., Page, A.T., Park, S.S., Lauderdale, J.D., Cunningham, M.L., and Dalton, S. (2013). Directed differentiation of human pluripotent cells to neural crest stem cells. Nat. Protoc. *8*, 203–212.

Meng, J., Zhang, X.-T., Liu, X.-L., Fan, L., Li, C., Sun, Y., Liang, X.-H., Wang, J.-B., Mei, Q.-B., Zhang, F., et al. (2016). WSTF promotes proliferation and invasion of lung cancer cells by inducing EMT via PI3K/Akt and IL-6/STAT3 signaling pathways. Cell. Signal. *28*, 1673–1682.

Meng, X., Lu, X., Morris, C.A., and Keating, M.T. (1998). A novel human gene FKBP6 is deleted in Williams syndrome. Genomics *52*, 130–137.

Meng, Y., Zhang, Y., Tregoubov, V., Falls, D.L., and Jia, Z. (2003). Regulation of spine morphology and synaptic function by LIMK and the actin cytoskeleton. Rev. Neurosci. *14*, 233–240.

Merla, G., Ucla, C., Guipponi, M., and Reymond, A. (2002). Identification of additional transcripts in the Williams-Beuren syndrome critical region. Hum. Genet. *110*, 429–438.

Mervis, C.B., Dida, J., Lam, E., Crawford-Zelli, N.A., Young, E.J., Henderson, D.R., Onay, T., Morris, C.A., Woodruff-Borden, J., Yeomans, J., et al. (2012). Duplication of GTF2I results in separation anxiety in mice and humans. Am. J. Hum. Genet.

Mervis, C.B., Klein-Tasman, B.P., Huffman, M.J., Velleman, S.L., Pitts, C.H., Henderson, D.R., Woodruff-Borden, J., Morris, C.A., and Osborne, L.R. (2015). Children with 7q11.23 duplication syndrome: Psychological characteristics. Am. J. Med. Genet. Part A.

Micale, L., Fusco, C., Augello, B., Napolitano, L.M.R., Dermitzakis, E.T., Meroni, G., Merla, G., and Reymond, A. (2008). Williams–Beuren syndrome TRIM50 encodes an E3 ubiquitin ligase. Eur. J. Hum. Genet. *16*, 1038–1049.

Miller, E.E., Kobayashi, G.S., Musso, C.M., Allen, M., Ishiy, F.A.A., de Caires, L.C., Goulart, E., Griesi-Oliveira, K., Zechi-Ceide, R.M., Richieri-Costa, A., et al. (2017). EIF4A3 deficient human iPSCs and mouse models demonstrate neural crest defects that underlie Richieri-Costa-Pereira syndrome. Hum. Mol. Genet. *26*, 2177–2191.

Mitchell, K.J. (2014). The genetic architecture of neurodevelopmental disorders. Genet.

Neurodev. Disord. bioRxiv (P. 009449).

Mizuseki, K., Sakamoto, T., Watanabe, K., Muguruma, K., Ikeya, M., Nishiyama, A., Arakawa, A., Suemori, H., Nakatsuji, N., Kawasaki, H., et al. (2003). Generation of neural crest-derived peripheral neurons and floor plate cells from mouse and primate embryonic stem cells. Proc. Natl. Acad. Sci. *100*, 5828–5833.

Mogha, A., Benesh, A.E., Patra, C., Engel, F.B., Schöneberg, T., Liebscher, I., and Monk, K.R. (2013). Gpr126 functions in Schwann cells to control differentiation and myelination via G-protein activation. J. Neurosci. *33*, 17976–17985.

Monsoro-Burq, A.-H., Wang, E., and Harland, R. (2005). Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during Xenopus neural crest induction. Dev. Cell *8*, 167–178.

Morris, C.A., Demsey, S.A., Leonard, C.O., Dilts, C., and Blackburn, B.L. (1988). Natural history of Williams syndrome: physical characteristics. J. Pediatr. *113*, 318–326.

Morris, C.A., Mervis, C.B., Hobart, H.H., Gregg, R.G., Bertrand, J., Ensing, G.J., Sommer, A., Moore, C.A., Hopkin, R.J., Spallone, P.A., et al. (2003). GTF2I hemizygosity implicated in mental retardation in Williams syndrome: Genotype-phenotype analysis of five families with deletions in the Williams syndrome region. Am. J. Med. Genet.

Morris, C.A., Mervis, C.B., Paciorkowski, A.P., Abdul-Rahman, O., Dugan, S.L., Rope, A.F., Bader, P., Hendon, L.G., Velleman, S.L., Klein-Tasman, B.P., et al. (2015). 7q11.23 Duplication syndrome: Physical characteristics and natural history. Am. J. Med. Genet. Part A.

Muñoz, W.A., and Trainor, P.A. (2015). Neural crest cell evolution: how and when did a neural crest cell become a neural crest cell. Curr. Top. Dev. Biol. *111*, 3–26.

Nagano, F., Kawabe, H., Nakanishi, H., Shinohara, M., Deguchi-Tawarada, M., Takeuchi, M., Sasaki, T., and Takai, Y. (2002). Rabconnectin-3, a novel protein that binds both GDP/GTP exchange protein and GTPase-activating protein for Rab3 small G protein family. J. Biol. Chem. *277*, 9629–9632.

Nishi, M., Aoyama, F., Kisa, F., Zhu, H., Sun, M., Lin, P., Ohta, H., Van, B., Yamamoto,

S., Kakizawa, S., et al. (2012). TRIM50 protein regulates vesicular trafficking for acid secretion in gastric parietal cells. J. Biol. Chem. *287*, 33523–33532.

Oka, K., Honda, M.J., Tsuruga, E., Hatakeyama, Y., Isokawa, K., and Sawa, Y. (2012). Roles of collagen and periostin expression by cranial neural crest cells during soft palate development. J. Histochem. Cytochem. *60*, 57–68.

Osborne, L.R., Soder, S., Shi, X.M., Pober, B., Costa, T., Scherer, S.W., and Tsui, L.C. (1997). Hemizygous deletion of the syntaxin 1A gene in individuals with Williams syndrome. Am. J. Hum. Genet. *61*, 449–452.

Pagon, R.A., Graham, J.M., Zonana, J., and Yong, S.L. (1981). Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. J. Pediatr. *99*, 223–227.

Palmer, S.J., Tay, E.S.E., Santucci, N., Cuc Bach, T.T., Hook, J., Lemckert, F.A., Jamieson, R. V., Gunnning, P.W., and Hardeman, E.C. (2007). Expression of Gtf2ird1, the Williams syndrome-associated gene, during mouse development. Gene Expr. Patterns *7*, 396–404.

Parisi, M.A. (1993). Hirschsprung Disease Overview.

Parrott, A., James, J., Goldenberg, P., Hinton, R.B., Miller, E., Shikany, A., Aylsworth, A.S., Kaiser-Rogers, K., Ferns, S.J., Lalani, S.R., et al. (2015). Aortopathy in the 7q11.23 microduplication syndrome. Am. J. Med. Genet. Part A.

Patthey, C., Edlund, T., and Gunhaga, L. (2008). Wnt-regulated temporal control of BMP exposure directs the choice between neural plate border and epidermal fate. Development *136*.

Peoples, R., Perez-Jurado, L., Wang, Y.K., Kaplan, P., and Francke, U. (1996). The gene for replication factor C subunit 2 (RFC2) is within the 7q11.23 Williams syndrome deletion. Am. J. Hum. Genet. *58*, 1370–1373.

Pérez Jurado, L.A., Wang, Y.K., Francke, U., and Cruces, J. (1999). TBL2, a novel transducin family member in the WBS deletion: characterization of the complete sequence, genomic structure, transcriptional variants and the mouse ortholog. Cytogenet.

Cell Genet. *86*, 277–284.

Perrier, A.L., Tabar, V., Barberi, T., Rubio, M.E., Bruses, J., Topf, N., Harrison, N.L., and Studer, L. (2004). Derivation of midbrain dopamine neurons from human embryonic stem cells. Proc. Natl. Acad. Sci. *101*, 12543–12548.

Pober, B.R. (2010). Williams–Beuren Syndrome. *3*.

Pober, B.R., Johnson, M., and Urban, Z. (2008). Mechanisms and treatment of cardiovascular disease in Williams-Beuren syndrome. J. Clin. Invest. *118*, 1606–1615.

Poot, R.A., Bozhenok, L., van den Berg, D.L.C., Steffensen, S., Ferreira, F., Grimaldi, M., Gilbert, N., Ferreira, J., and Varga-Weisz, P.D. (2004). The Williams syndrome transcription factor interacts with PCNA to target chromatin remodelling by ISWI to replication foci. Nat. Cell Biol. *6*, 1236–1244.

Poot, R.A., Bozhenok, L., Berg, D.L.C. van den, Hawkes, N., and Varga-Weisz, P.D. (2005). Chromatin Remodelling by WSTF-ISWI at the Replication Site: Opening a Window of Opportunity for Epigenetic Inheritance? Cell Cycle *4*, 543–546.

Prescott, S.L., Srinivasan, R., Marchetto, M.C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F.H., Swigut, T., and Wysocka, J. (2015). Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. Cell *163*, 68–83.

Purves, D., Augustine, G.J., Fitzpatrick, D., Hall, W.C., LaMantia, A.-S., McNamara, J.O., and White, L.E. (2008). Neuroscience (Sinauer).

Qin, J., Chen, X., Xie, X., Tsai, M.-J., and Tsai, S.Y. (2010). COUP-TFII regulates tumor growth and metastasis by modulating tumor angiogenesis. Proc. Natl. Acad. Sci. U. S. A. *107*, 3687–3692.

Quong, M.W., Massari, M.E., Zwart, R., and Murre, C. (1993). A new transcriptional-activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. Mol. Cell. Biol. *13*, 792–800.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest. Cell Stem Cell *11*, 633–648.

Rada-Iglesias, A., Prescott, S.L., and Wysocka, J. (2013). Human genetic variation within neural crest enhancers: molecular and phenotypic implications. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *368*, 20120360.

Rando, T.A., and Chang, H.Y. (2012). Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. Cell *148*, 46–57.

Ranheim, E.A., Kwan, H.C.K., Reya, T., Wang, Y.-K., Weissman, I.L., and Francke, U. (2005). Frizzled 9 knock-out mice have abnormal B-cell development. Blood *105*, 2487–2494.

Rios, H., Koushik, S. V, Wang, H., Wang, J., Zhou, H.-M., Lindsley, A., Rogers, R., Chen, Z., Maeda, M., Kruzynska-Frejtag, A., et al. (2005). periostin null mice exhibit dwarfism, incisor enamel defects, and an early-onset periodontal disease-like phenotype. Mol. Cell. Biol. *25*, 11131–11144.

Robinson, W.P., Waslynka, J., Bernasconi, F., Wang, M., Clark, S., Kotzot, D., and Schinzel, A. (1996). Delineation of 7q11.2 deletions associated with Williams-Beuren syndrome and mapping of a repetitive sequence to within and to either side of the common deletion. Genomics *34*, 17–23.

Sadaghiani, B., and Thiébaud, C.H. (1987). Neural crest development in the Xenopus laevis embryo, studied by interspecific transplantation and scanning electron microscopy. Dev. Biol. *124*, 91–110.

Sadeghifar, F., Böhm, S., Vintermist, A., and Östlund Farrants, A.-K. (2015). The B-WICH chromatin-remodelling complex regulates RNA polymerase III transcription by promoting Max-dependent c-Myc binding. Nucleic Acids Res. *43*, 4477–4490.

Sadler, L.S., Robinson, L.K., Verdaasdonk, K.R., and Gingell, R. (1993). The Williams syndrome: Evidence for possible autosomal dominant inheritance. Am. J. Med. Genet. *47*, 468–470.

Sakurai, T., Dorr, N.P., Takahashi, N., McInnes, L.A., Elder, G.A., and Buxbaum, J.D. (2011). Haploinsufficiency of Gtf2i, a gene deleted in Williams Syndrome, leads to increases in social interactions. Autism Res. *4*, 28–39.

Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., Chu, S.H., Moreau, M.P., Gupta, A.R., Thomson, S.A., et al. (2011). Multiple Recurrent De Novo CNVs, Including Duplications of the 7q11.23 Williams Syndrome Region, Are Strongly Associated with Autism. Neuron.

Sareila, O., Jaakkola, N., Olofsson, P., Kelkka, T., and Holmdahl, R. (2013). Identification of a region in p47phox/NCF1 crucial for phagocytic NADPH oxidase (NOX2) activation. J. Leukoc. Biol. *93*, 427–435.

Savina, N. V, Smal, M.P., Kuzhir, T.D., Egorova, T.M., Khurs, O.M., Polityko, A.D., and Goncharova, R.I. (2011). Chromosomal instability at the 7q11.23 region impacts on DNA-damage response in lymphocytes from Williams-Beuren syndrome patients. Mutat. Res. *724*, 46–51.

Schlesinger, B.E., By, F.R.C.P., Butler, N.R., and Black, D.C.H.J.A. (1956). Severe type of infantile hypercalcaemia. Br Med J *1*, 127–134.

Schosserer, M., Minois, N., Angerer, T.B., Amring, M., Dellago, H., Harreither, E., Calle-Perez, A., Pircher, A., Gerstl, M.P., Pfeifenberger, S., et al. (2015). Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. Nat. Commun. *6*, 6158.

Schubert, C. (2009). The genomic basis of the Williams - Beuren syndrome. Cell. Mol. Life Sci.

Schubert, C., and Laccone, F. (2006). Williams-Beuren syndrome: determination of deletion size using quantitative real-time PCR. Int. J. Mol. Med. *18*, 799–806.

Seki, D., Takeshita, N., Oyanagi, T., Sasaki, S., Takano, I., Hasegawa, M., and Takano-Yamamoto, T. (2015). Differentiation of Odontoblast-Like Cells From Mouse Induced Pluripotent Stem Cells by Pax9 and Bmp4 Transfection. Stem Cells Transl. Med. *4*, 993–997.

Selleck, M.A., and Bronner-Fraser, M. (1995). Origins of the avian neural crest: the role of neural plate-epidermal interactions. Development *121*, 525–538.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. *68*, 850–858.

Shi, H., Gong, Y., Qiang, L., Li, X., Zhang, S., Gao, J., Li, K., Ji, X., Tian, L., Gu, X., et al. (2016). Derivation of Schwann cell precursors from neural crest cells resident in bone marrow for cell therapy to improve peripheral nerve regeneration. Biomaterials *89*, 25–37.

Slaugenhaupt, S.A., Blumenfeld, A., Gill, S.P., Leyne, M., Mull, J., Cuajungco, M.P., Liebert, C.B., Chadwick, B., Idelson, M., Reznik, L., et al. (2001). Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. Am. J. Hum. Genet. *68*, 598–605.

Somerville, M.J., Mervis, C.B., Young, E.J., Seo, E.-J., del Campo, M., Bamforth, S., Peregrine, E., Loo, W., Lilley, M., Pérez-Jurado, L.A., et al. (2005). Severe Expressive-Language Delay Related to Duplication of the Williams–Beuren Locus. N. Engl. J. Med.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*, 543–549.

Sommer, C.A., Sommer, A.G., Longmire, T.A., Christodoulou, C., Thomas, D.D., Gostissa, M., Alt, F.W., Murphy, G.J., Kotton, D.N., and Mostoslavsky, G. (2010). Excision of reprogramming transgenes improves the differentiation potential of iPS cells generated with a single excisable vector. Stem Cells *28*, 64–74.

Stemple, D.L., and Anderson, D.J. (1992). Isolation of a stem cell for neurons and glia from the mammalian neural crest. Cell *71*, 973–985.

Strømme, P., Bjørnstad, P.G., and Ramstad, K. (2002). Prevalence Estimation of Williams Syndrome. J. Child Neurol. *17*, 269–271.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and

Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell.

Tamura, Y., Sano, M., Nakamura, H., Ito, K., Sato, Y., Shinmura, K., Ieda, M., Fujita, J., Kurosawa, H., Ogawa, S., et al. (2016). Neural crest-derived resident cardiac cells contribute to the restoration of adrenergic function of transplanted heart in rodent. Cardiovasc. Res. *109*, 350–357.

Tassabehji, M. (2003). Williams-Beuren syndrome: a challenge for genotype-phenotype correlations. Hum. Mol. Genet.

Tassabehji, M., Metcalfe, K., Fergusson, W.D., Carette, M.J.A., Dore, J.K., Donnai, D., Read, A.P., Pröschel, C., Gutowski, N.J., Mao, X., et al. (1996). LIM–kinase deleted in Williams syndrome. Nat. Genet. *13*, 272–273.

Tassabehji, M., Hammond, P., Karmiloff-Smith, A., Thompson, P., Thorgeirsson, S.S., Durkin, M.E., Popescu, N.C., Hutton, T., Metcalfe, K., Rucka, A., et al. (2005). GTF2IRD1 in craniofacial development of humans and mice. Science *310*, 1184–1187.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Toma, J.G., Akhavan, M., Fernandes, K.J.L., Barnabé-Heider, F., Sadikot, A., Kaplan, D.R., and Miller, F.D. (2001). Isolation of multipotent adult stem cells from the dermis of mammalian skin. Nat. Cell Biol. *3*, 778–784.

Tomita, Y., Matsumura, K., Wakamatsu, Y., Matsuzaki, Y., Shibuya, I., Kawaguchi, H., Ieda, M., Kanakubo, S., Shimazaki, T., Ogawa, S., et al. (2005). Cardiac neural crest cells contribute to the dormant multipotent stem cell in the mammalian heart. J. Cell Biol. *170*, 1135–1146.

Torniero, C., Dalla Bernardina, B., Novara, F., Cerini, R., Bonaglia, C., Pramparo, T., Ciccone, R., Guerrini, R., and Zuffardi, O. (2008). Dysmorphic features, simplified gyral pattern and 7q11.23 duplication reciprocal to the Williams-Beuren deletion. Eur. J. Hum. Genet. *16*, 880–887.

Tsukumo, Y., Tsukahara, S., Furuno, A., Iemura, S., Natsume, T., and Tomida, A. (2016). TBL2 Associates With ATF4 mRNA Via Its WD40 Domain and Regulates Its Translation During ER Stress. J. Cell. Biochem. *117*, 500–509.

Turner, D.J., Miretti, M., Rajan, D., Fiegler, H., Carter, N.P., Blayney, M.L., Beck, S., and Hurles, M.E. (2008). Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. Nat. Genet. *40*, 90–95.

Tuttle, A.M., Hoffman, T.L., and Schilling, T.F. (2014). Rabconnectin-3a regulates vesicle endocytosis and canonical Wnt signaling in zebrafish neural crest migration. PLoS Biol. *12*, e1001852.

Uehara, T., Kage-Nakadai, E., Yoshina, S., Imae, R., and Mitani, S. (2015). The Tumor Suppressor BCL7B Functions in the Wnt Signaling Pathway. PLoS Genet. *11*, e1004921.

Urban, A.E., and Purmann, C. (2015). Using iPSCs and genomics to catch CNVs in the act. Nat. Genet. *47*, 100–101.

Vega-Lopez, G.A., Cerrizuela, S., and Aybar, M.J. (2017). Trunk neural crest cells: formation, migration and beyond. Int. J. Dev. Biol. *61*, 5–15.

Vignali, M., Hassan, A.H., Neely, K.E., and Workman, J.L. (2000). ATP-dependent chromatin-remodeling complexes. Mol. Cell. Biol. *20*, 1899–1910.

Vintermist, A., Böhm, S., Sadeghifar, F., Louvet, E., Mansén, A., Percipalle, P., and Ostlund Farrants, A.-K. (2011). The chromatin remodelling complex B-WICH changes the chromatin structure and recruits histone acetyl-transferases to active rRNA genes. PLoS One *6*, e19184.

Waldo, K., Miyagawa-Tomita, S., Kumiski, D., and Kirby, M.L. (1998). Cardiac Neural Crest Cells Provide New Insight into Septation of the Cardiac Outflow Tract: Aortic Sac to Ventricular Septal Closure. Dev. Biol. *196*, 129–144.

Wang, C.-H., Su, P.-T., Du, X.-Y., Kuo, M.-W., Lin, C.-Y., Yang, C.-C., Chan, H.-S., Chang, S.-J., Kuo, C., Seo, K., et al. (2010). Thrombospondin type I domain containing 7A (THSD7A) mediates endothelial cell migration and tube formation. J. Cell. Physiol. *222*, 685–694.

Wang, L., Ming Wang, L., Chen, W., and Chen, X. (2016). Bicuspid Aortic Valve: A Review of its Genetics and Clinical Significance. J. Heart Valve Dis. *25*, 568–573.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., et al. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell *7*, 618–630.

Wen, Z. (2017). Modeling neurodevelopmental and psychiatric diseases with human iPSCs. J. Neurosci. Res.

Wessel, A., Gravenhorst, V., Buchhorn, R., Gosch, A., Partsch, C.-J., and Pankau, R. (2004). Risk of sudden death in the Williams-Beuren syndrome. Am. J. Med. Genet. A *127A*, 234–237.

Williams, J.C.P., Barratt-Boyes, B.G., and Lowe, J.B. (1961). Supravalvular Aortic Stenosis. Circulation *24*, 1311–1318.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810–813.

Xiao, A., Li, H., Shechter, D., Ahn, S.H., Fabrizio, L.A., Erdjument-Bromage, H., Ishibe-Murakami, S., Wang, B., Tempst, P., Hofmann, K., et al. (2009). WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. Nature *457*, 57–62.

Xie, X., Qin, J., Lin, S.-H., Tsai, S.Y., and Tsai, M.-J. (2011). Nuclear receptor chicken ovalbumin upstream promoter-transcription factor II (COUP-TFII) modulates mesenchymal cell commitment and differentiation. Proc. Natl. Acad. Sci. U. S. A. *108*, 14843–14848.

Yoshioka, N., Gros, E., Li, H.-R., Kumar, S., Deacon, D.C., Maron, C., Muotri, A.R., Chi, N.C., Fu, X.-D., Yu, B.D., et al. (2013). Efficient generation of human iPSCs by a synthetic self-replicative RNA. Cell Stem Cell *13*, 246–254.

Young, E.J., Lipina, T., Tam, E., Mandel, A., Clapcote, S.J., Bechard, A.R., Chambers, J., Mount, H.T.J., Fletcher, P.J., Roder, J.C., et al. (2008). Reduced fear and aggression and altered serotonin metabolism in Gtf2ird1-targeted mice. Genes. Brain. Behav. *7*, 224–234.

Zarate, Y.A., Lepard, T., Sellars, E., Kaylor, J.A., Alfaro, M.P., Sailey, C., Schaefer, G.B., and Collins, R.T. (2014). Cardiovascular and genitourinary anomalies in patients with duplications within the Williams syndrome critical region: Phenotypic expansion and review of the literature. Am. J. Med. Genet. Part A.

Zeltner, N., Fattahi, F., Dubois, N.C., Saurat, N., Lafaille, F., Shang, L., Zimmer, B., Tchieu, J., Soliman, M.A., Lee, G., et al. (2016). Capturing the biology of disease severity in a PSC-based model of familial dysautonomia. Nat. Med. *22*, 1421–1427.

Zhu, M., Koonpaew, S., Liu, Y., Shen, S., Denning, T., Dzhagalov, I., Rhee, I., and Zhang, W. (2006). Negative regulation of T cell activation and autoimmunity by the transmembrane adaptor protein LAB. Immunity *25*, 757–768.

7. ACKNOWLEDGMENTS

My first thanks go to my supervisor, Giuseppe Testa, who believed in me five years ago, despite of my shyness, guided me and helped me growing both professionally and personally.

I also want to thank my external advisor Oliver Brüstle and my internal advisor Bruno Amati for their good advice and suggestions, and all the facilities and collaborators in the Campus (the cell culture, the sequencing, the imaging and the mass spectrometry unit and Giorgio Scita's group) for their technical support.

Special thanks go to all my past and present colleagues: Pietro, Marija, Maddalena, Italia, Michele, Veronica, Pierre-Luc, Maria Elena, Mattia, Natascia, Silvia, Francesco, Federico, Patricio, Sebastiano, Celeste, Danila, Reinald, Nicolò and Alejandro.

In particular, I want to thank Giulia for her support throughout these years and till the very last days; Alessandro for the fruitful and, sometimes, pretty lively discussions and for his precious help with bioinformatic analysis; Sina for his useful teachings and 'nice' nicknames; Elena for being there anytime I had a doubt and Giuseppe for being the best bench mate I could ever ask for.

I would like to thank Matteo for his unconditional support and for tolerating all the weekends I had to spend in the lab.

Of course, I thank my family, mum, dad, Luca and Chiara, for always supporting my professional and personal choices and for helping me in reaching all my achievements.

Finally, I am extremely grateful to all WBS and 7dup patients and their families, especially the ones I could personally meet, for truly believing in our research.