UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

IGBMC-CNRS UMR 7104-INSERM U 964

THÈSE présentée par :

Baptiste BIDON

soutenue le : **03 Novembre 2017**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Biologie/Aspects moléculaires et cellulaires de la biologie

Mediator and NER factors in transcription initiation

Médiateur et facteurs NER lors de l'initiation de la transcription

THÈSE dirigée par :

Dr EGLY Jean-Marc	DR INSERM, IGBMC, Illkirch
Dr COIN Frédéric	DR INSERM, IGBMC, Illkirch

RAPPORTEURS :

Dr DURR Alexandra	PUPH, Hôpital Pitié Salpêtrière, Paris
Dr BULAVIN Dmitry	DR INSERM, IRCAN, Nice

AUTRES MEMBRES DU JURY :

Dr WEIXLBAUMER Albert	CR INSERM, IGBMC, Illkirch
Dr BORG Jean-Paul	PUPH, Institut Paoli-Calmettes, Marseille

Acknowledgments

I am grateful to Mrs Alexandra Durr, Mr Dmitry Bulavin, Mr Albert Weixlbaumer and Mr Jean-Paul Borg for examining and reviewing my thesis.

I am thankful to Mr Jean-Marc Egly and Mr Frédéric Coin to have given me the opportunity to work on such exciting projects.

I am largely obliged to Mr Jean-Marc Egly and Mr Nicolas Le May for all the scientific work and experiments (functioning or not) that have been achieved and for discussions we had. Thank you…

I do have a special thought to 4-0-2-5 members where I spend the last four years. Not only science is about sharing and this PhD would not have been possible without Lise-Marie, Cathy, Philippe, Emmanuel and Carlos.

I am also thankful to all members of Egly/Coin's lab for technical support, conversation, time or just the beer we shared. Grazie mille…

Thank you to my parents to have raised me as a free spirit.

Summary

Tables and Figures

Abbreviations

6-4 PP	6-4 pyrimidine pyrimidone photoproducts
AP-1	Activator Protein-1
AR	Androgen Receptor
ATAC	Ada Two A-Containing
ATP	Adenosine TriPhosphate
ATRA	All-Trans Retinoic Acid
bp	base pairs
ATAC	Ada Two A-Containing
BER	Base Excision Repair
BRE	TFIIB Recognition Elements
CAK	CDK-Activating Kinase
CBP	CREB-Binding Protein
CCNC/H	Cyclin C/H
CDK	Cyclin-Dependent Kinase
CETN2	Centrin2
COFS	cerebello-oculo-facio-skeletal syndrome
CPD	cyclobutane-pyrimidine dimers
CPDD	CTCF-Pair-Defined Domain
CPE	Core Promoter Elements
CPSF	Cleavage and Polyadenylation Specificity Factor
CS	Cockaine syndrome
CSA/B	Cockaine Syndrome complementation group A/B protein
CstF	Cleavage stimulatory Factor
CTCF	CCCTC-binding protein
CTD	C-Terminal Domain
DBD	DNA Binding Domain
DCE	Downstream Core Element
DDB	Damage specific DNA Binding protein
DNA	Deoxy-riboNucleic Acid
DnmT	DNA methyltranferase
DP1-3	Dimerization Partner 1-3
DSIF	DBR Sensitivity-Inducing Factor
DPE	Downstream Promoter Element
E2F	Andenovirus E2 gene Factor
EGR-1	Early Growth Response-1 protein
ELL	Eleven-nineteen Lysine-rich Leukaemia protein
ER	oEstrogen Receptor
ES	Embryonic Stem
GCN5	General Control of Nutrition protein 5
GG-NER	Global-Genome NER
GNAT	GCN5-related N-Acetyl Transferase
GR	Glucocorticoid Receptor
GTF	General Transcription Factors
HAT	Histone Acetyl Transferase
HDAC	Histone De-ACetylases
HMGN1	High Mobility Group Nucleosome-biding domain-containing protein 1
HMT	Histone Methyl Transferase

HNF-4	Hepatocyte Nuclear Factor-4
HPTM	Histones Post-Translational Modification
HR	Homologous Repair
HR23B	Homologue Rad 23 B
HRE	Hormone Response Element
HSP	Heat Shock Proteins
IEG	Immediate Early response Genes
Inr	Initiator
KAT2A	Lysine (K) AcetylTransferase 2A
LBD	Ligand Binding Factor
LMR	Low Methylated Regions
LRG	Late Response Genes
mRNA	messenger RNA
MAT1	Ménage-À-Trois 1 protein
MMR	Mis Match Repair
MR	Mineralocorticoid Receptor
MTE	Motif Ten Element
MYST	MOF - Ybf2/Sas3 - Sas2 - Tip60
NCOR	Nuclear receptor Co-Repressor
NELF	Nasal Embryonic LHRH Factor
NER	Nucleotide Excision Repair
NHEJ	Non-Homologous End Joining
PCAF	p300/CBP-Associated Factor
PCNA	Proliferating Cell Nuclear Antigen
PIC	Pre-Initiation Complex
PR	Progesterone Receptor
Pol	Polymerase
PPAR	Peroxisome Proliferator-Activated-Receptor
PRMT	aRginine Methyl Transferases
P-TEFβ	Positive Transcription Elongation Factor β
PTM	Post-Translational Modification
RNA	RiboNucleic Acid
RAR	Retinoic Acid Receptor
RARE	Retinoic Acid Response Element
Rpb	RNA Polymerase B
RPA	Replication Factor A
RXR	Retinoid X Receptor
SAGA	Spt-Ada-Gcn5 Acetyltransferase
Sen1	Senatoxin1
Ser	Serine
Sirt	Sirtuin
SMRT	Silencing Mediator of Retinoic acid and Thyroid hormone receptor
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
SCC	Stem Cell Coactivator
T3R	Thyroid hormone Receptor
TAD	Trans-Activation Domain
TAF	TBP Associated Factors
TBP	TATA Binding Protein
TC-NER	Transcription-Coupled NER
TFIIA-H/S	general Transcription Factor of Pol II A-H/S
Thr	Threonine

TDG	Thymine DNA Glycosylase
TPA	TetradecanoylPhorbol Acetate
TR2	Testicular Receptor
TRE	TPA Responsive Element
TSS	Transcription Start Site
TTD	Trichothiodystrophy
Tyr	Tyrosine
UV	Ultra-Violet
VAD	Vitamin A Deficiency syndrome
VDR	Vitamin D Receptor
XCPE1/2	X Core Promoter Element 1/2
XP	Xeroderma Pigmentosum
XPA-G	Xeroderma Pigmentosum group A-G protein

INTRODUCTION

Transcription is a fundamental process of living cell. It allows the genetic information present in **D**eoxy-ribo**N**ucleic **A**cid (DNA) to be duplicated on a **R**ibo**N**ucleic **A**cid (RNA) molecule. A recent study, performed In human cells, has shown that nearly three quarters of the genome can potentially be transcribed(Djebali et al., 2012).

For protein coding genes, transcription is first of two processes that allow protein synthesis. In such case, the molecule is called messenger RNA (mRNA) and is subsequently translated into a functional protein by the ribosomal machinery.

In living cells, a dedicated enzyme called RNA polymerase (Pol) is accomplishing the process of transcription. While there is only one RNA Polymerase in Archea and Bacteria, distinct ones have been isolated and characterised in Eucharyota(Pikaard et al., 2008; Roeder and Rutter, 1970). Pol I is mostly dedicated to the synthesis of ribosomal RNA while Pol III is responsible for the synthesis of transfer RNA and other non-coding RNAs like snRNA, snoRNA or microRNA. Pol IV and V had only been identified in Plantae Kingdom and are dedicated to the synthesis of small interfering RNA. Finally, Pol II is the polymerase eliciting mRNA synthesis. In this manuscript, I will describe the mechanisms of transcription by Pol II and the results we obtained when we investigated it.

I - RNA polymerase II mediated transcription

1) The RNA polymerase II

RNA polymerase II was first isolated by Roeder and colleagues(Roeder and Rutter, 1970). RNA polymerase II also called **R**NA **P**olymerase **B** (Rpb) is a large protein complex composed of twelve subunits named from Rpb1 to Rpb 12 and collectively weighting more than 500 kDa (Table 1). The five subunits Rpb 5, 6, 8, 10 and 12 are in common with the RNA polymerase I and III(Woychik and Young, 1990). Some of them are also conserved with their yeast counterparts(McKune et al., 1995) and to a certain extent with their bacterial counterparts.

Pol II consists of a ten subunits core sub-complex associated with the stalk sub-complex(Armache et al., 2003). The whole complex is first assembled in the cellular cytoplasm with the help of several Heat Shock Proteins (HSP) before being

transferred in the nucleus through the nuclear import machinery(Boulon et al., 2010). This assembly starts with the formation of several sub-complexes including Rpb1-Rbp8-HSP90, before the full complex to be assembled. During transcription, the two largest subunits Rpb1 and Rpb 2 are forming a positively charged “cleft” where negatively charged DNA can bind. Rpb1 also forms a mobile “clamp”, which is then close around the DNA-RNA duplex, building a tunnel like structure. Rpb2 forms the “wall” that delineates the extremity of the tunnel(Armache et al., 2003).

Table 1: RNA Polymerase II

Sub-complexes	Sub-assemblies	Subunits	Features	Size (kDa)
Core	Rpb1 sub-assembly	Rpb1	Phosphorylation sites	191.6
		Rpb5	Common in Pol I, Pol II, Pol III	25.1
		Rpb6	Common in Pol I, Pol II, Pol III	17.9
		Rpb8	Common in Pol I, Pol II, Pol III	16;5
	Rpb2 sub-assembly	Rpb2	NTP binding site	138.8
		Rpb9		14.3
	Rpb3 sub-assembly	Rpb3	Promoter recognition	35.3
		Rpb10	Common in Pol I, Pol II, Pol III	8.3
		Rpb11		13.6
		Rpb12	Common in Pol I, Pol II, Pol III	7.7
Stalk		Rpb4		25.4
		Rpb7	Unique to Pol II	19.1

The **C**-**T**erminal **D**omain (CTD) of Rpb1 subunit forms a special domain of the polymerase not necessary for its catalytic activity *in vitro* but allowing the specific control of its function *in vivo*(Serizawa et al., 1993). It is composed of a hepta-peptide repeated 26 times in yeast and 52 times in human with almost regular sequence of YSPTSPS. Regulation occurs through phosphorylation and dephosphorylation of the tyrosine, threonine and serine residues Tyr1-Ser2-Thr4-Ser5-Ser7, numbered relative to their position in the hepta-peptide. The **P**re-**I**nitiation **C**omplex (PIC), including the general transcription factors, has been shown to preferentially associate with the dephosphorylated form of Pol II(Lu et al., 1991; Robinson et al., 2012). Ser5 and Ser7 are phosphorylated by the **C**yclin-**D**ependent **K**inase 7 (CDK7), part of the general transcription factor TFIIH, leading to the initiation of transcription and to the escape of the polymerase from the promoter(Wong et al., 2014; Zhou et al., 2000).

Ser 5 is also implicated in the recruitment of the capping machinery at the 5' extremity of the transcribed gene(Schroeder et al., 2000). Both marks are removed shortly after transcription initiation. Ser2 phosphorylation allows the promoter–proximal pause release(Jonkers and Lis, 2015). This mark is retained on the CTD during elongation and is required for splicing and cleavage machinery to be co-recruited(Gu et al., 2013). In addition to Ser2, Tyr1 and Thr4 phosphorylation are needed to allow the arrival of the poly-adenylation machinery(Harlen et al., 2016; Mayer et al., 2012).

2) Promoter

1. Organisation

Promoter is a region of the gene where the transcription machinery assembled. This region includes the **T**ranscription **S**tart **S**ite (TSS) where the transcription actually initiates.

Two major types of promoter organisation have been described in the past decade, focused and dispersed(Carninci et al., 2006). The first one is characterised by the presence of one TSS, positioned at a specific nucleotide or eventually some contiguous nucleotides. On the other hand, the second promoter type corresponds to several weakly actionable TSS that are spread on a 50 to 100 bp region. Aside from those two main types, a few peculiar promoters have been described to contain a principal TSS surrounded by several weak ones.

If the focused type seems the more frequent in simpler organisms, it represents only a third of human genes. Analysing them has led to the discovery of major **C**ore **P**romoter **E**lements (CPE) like TATA box, BRE, XPCE1 or DPE. Focused promoters are mostly present at developmental and highly regulated genes. Dispersed promoters have been less studied even though they represent at least two third of human promoters. They are typically found in CpG regions and are controlling housekeeping or constitutive genes. The manuscript will concentrate on focused promoters.

2. Structure

The core promoter is the DNA sequence where the PIC assembles and the transcription initiation occurs. It is organized around the TSS referred to as +1 nucleotide and is generally considered to span 35 to 40 nucleotides upstream and downstream of it. Core promoters contain variable number of elements with various nucleotide sequences; the first identified being the TATA box element by Chambon and colleagues. All sequences are given in accordance with the IUPAC nomenclature and resumed in Table 2.

Table 2: Core promoter elements

Elements	Position	Consensus seq.	GTF
Inr	-2 to +5	YYANWYY	TFIID (TAF1/TAF2)
sInr	-2 to +5	CCATYTT	
TCT	-1 to +6	YCTYTYY	
XCPE1	-8 to +2	DSGYGGRASM	Unknow
XCPE2	-9 to +2	VCYCRTTRCMY	
TATA box	-30 to -24	TATAWAAR	TBP
BREu	-37 to -31	SSRCGCC	TFIIA/TFIIB
BREd	-23 to -17	RTDKKK	
DPE	+28 to +33	DSWYVY	TFIID
MTE	+18 to +27	CSARCSSAAC	TFIID
DCE I	+6 to +11	CTTC	TFIID
II	+16 to +21	CTGT	
III	+30 to +34	AGC	

2.a. Initiator (Inr)

In 1980, Chambon described Inr, as a sequence enriched around the TSS of Pol II transcribed genes(Corden et al., 1980). Smale latterly defined this sequence encompassing the TSS, as able to drive transcription initiation *in vivo* and *in* vitro, without the need of a TATA box. Indeed, the presence of the TATA box or other CPE seems only needed to potentiate the initiation. Pyrimidine rich Inr corresponds to YYANWYY, where the A is defined as the +1 nucleotide(Javahery et al., 1994; Smale and Baltimore, 1989), and the whole sequence is present at nearly half of human genes(Gershenzon and Ioshikhes, 2005). It is recognized by the general transcription

factor TFIID(Kaufmann and Smale, 1994), especially TAF1 and TAF2(Chalkley and Verrijzer, 1999).

Recently, a strict Inr (sInr), with a more restrictive sequence has been found in TATA less promoters where it seems to bypass TATA box by cooperating with the Sp1 sequence(Yarden et al., 2009). CCATYTT sequence was defined as sInr and is present in 1,5% of human genes.

Finally, an alternative pyrimidine rich initiator element called TCT has been described to encompass the TSS of many genes and notably members of the ribosome gene family. Its sequence is YCTYTYY, where C appears to be the TSS(Parry et al., 2010).

2.b. X gene core promoter element (XCPE1/2)

The **X** **C**ore **P**romoter **E**lement 1 (XCPE1) has been characterised as able to induce transcription initiation of hepatitis virus X gene and lately found in 1% of human promoter, especially TATA less ones. It spans from -8 to +2 with the sequence DSGYGGRASM, where A is the TSS(Tokusumi et al., 2007).

The XCPE2 core promoter element directs transcription initiation of the second promoter of hepatitis virus X gene and have the sequence VCYCRTTRCMY, where M is the TSS(Anish et al., 2009). It is found at multiple TATA less human promoters.

2.c. TATA box

The TATA box has also been described by Chambon and named after its sequence composition. It was first thought to be a general transcription CPE. Nowadays, it is assumed that it is only present in 10% to 20% of human genes. Its consensus sequence, TATAWAAR, is present upstream to the TSS at nucleotide -28. Inr is associated with around 60% of all TATA box (Gershenzon and Ioshikhes, 2005). The TATA box is well conserved from Archea to *human*.

2.d. TFIIB Recognition Elements (BRE)

The TFII**B** **R**ecognition **E**lements (BRE) are two sequences located directly upstream (BREu)(Lagrange et al., 1998) and downstream (BREd)(Deng and Roberts, 2005) of the TATA box, but may also be present in TATA less promoters. Both sequences are well conserved from Archea to *human* and their sequences are

SSRCGCC and RTDKKK respectively. They are recognised by the general transcription factors TFIIA and TFIIB.

2.e. Downstream Promoter Element (DPE)

As its name suggests, the **D**ownstream **P**romoter **E**lement (DPE) is located downstream the +1 nucleotide between +28 and +33, with the consensus sequence DSWYVY. Its presence is apparently independent of either TATA box or Inr(Burke and Kadonaga, 1997; Gershenzon and Ioshikhes, 2005). The DPE element is notably associated with developmental regulatory genes(Zehavi et al., 2014). It is recognised and bound by the general transcription factor TFIID.

2.f. Motif Ten Element (MTE)

The **M**otif **T**en **E**lement (MTE), named after computational analysis(Ohler et al., 2002), is located downstream of the TSS and directly upstream of the DPE. It spans from +18 to +27 with the sequence CSARCSSAAC. MTE requires association with Inr sequence. Even though it can act independently of TATA box or DPE, their presences strongly reinforce its transcription initiation activity(Lim et al., 2004). MTE is also recognised by TFIID.

More recently, MTE and DPE combination have been found to form a bridge motif and to constitute an independent core promoter element(Theisen et al., 2010).

2.g. Downstream Core Element (DCE)

First observed in β-globin gene, the **D**ownstream **C**ore **E**lement (DCE) is a downstream element alternative to DPE, MTE or Bridge motif(Lewis et al., 2000). It is composed of three sequences, spanning from +6 to +11 (necessary motif CTTC), from +16 to +21 (necessary motif CTGT) and from +30 to +34 (necessary motif AGC). It is recognised by TFIID(Lee et al., 2005).

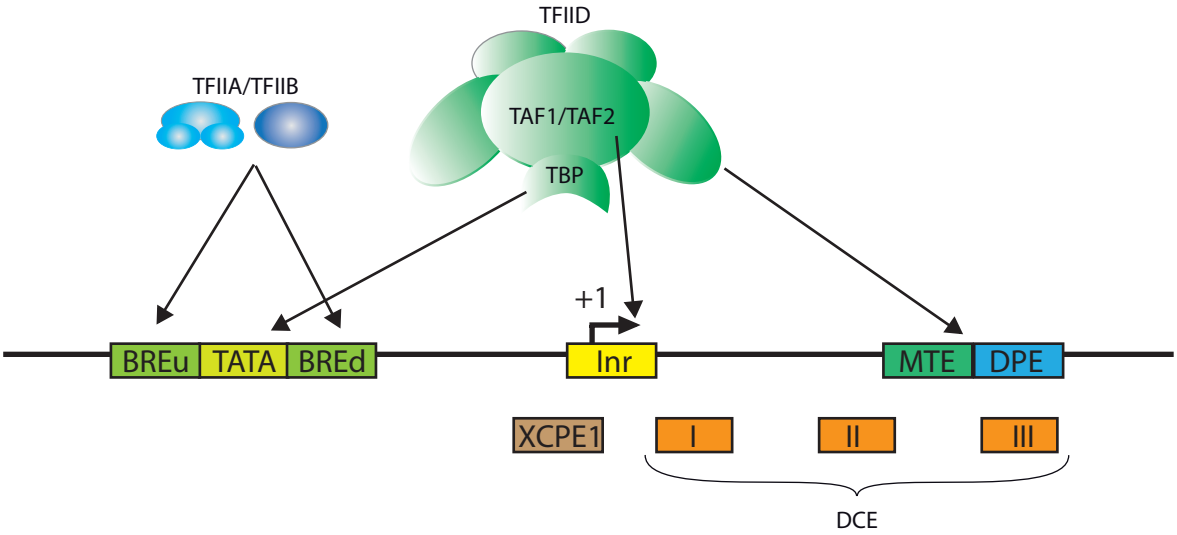


Figure 1: Core promoter elements and respective binding GTF

The diagram shows a promoter with most common CPE and the proteins from the transcription machinery that contact them.

3) Pre-initiation complex

Even if RNA polymerase II alone is able to transcribe DNA templates *in vitro*, it requires the assembly of a large **P**re-**I**nitiation **C**omplex (PIC) for transcription to start at the right nucleotide in a well-coordinated manner *in vivo*(Weil et al., 1979).

1. General transcription factors

General transcription factors (GTF) have first been purified from human cells in 1980 (Matsui et al., 1980) and further characterized(Reinberg and Roeder, 1987; Reinberg et al., 1987; Samuels et al., 1982). They are named TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH for basal/general **T**ranscription **F**actors associated with RNA Polymerase **II**; the last letter corresponds to purification fractions. When assembled with Pol II on the core promoter, they form the Pre-Initiation Complex. They allow the correct positioning of Pol II on the promoter and the transcription to start at the right nucleotide.

As well as core promoters largely vary in their sequence compositions; nowadays it is assumed that PIC composition may vary, especially for TFIID. Some subunits of general transcription factors seem to be specific to certain set of genes, cell types or even tissues(Akhtar and Veenstra, 2011). Here, we will describe their widely observable composition (Table 3).

1.a. TFIIA

The general transcription factor TFIIA is a three subunit complex composed of TFIIAα, TFIIAβ and TFIIAγ. It has been shown to interact with the BRE elements to stabilize TFIID(Buratowski et al., 1989; Lee et al., 1992).

1.b. TFIIB

The general transcription factor TFIIB is a single protein. It stabilizes the TFIID-DNA-TFIIA complex and helps the subsequent recruitment of the RNA polymerase II(Ha et al., 1993; Maldonado et al., 1990) together with TFIIF. TFIIB participates in determining TSS position(Li et al., 1994) and recognizes BRE elements. It regulates itself by auto-acetylation(Choi et al., 2003) and participates in promoter escape(Westover et al., 2004).

1.c. TFIID

The general transcription factor TFIID is the largest complex among all GTF. It is composed of the **T**ATA **B**inding **P**rotein (TBP) and 14 **T**BP **A**ssociated **F**actors (TAF). As its name indicates, TBP binds the consensus sequence of the TATA box(Corden et al., 1980; Sawadogo and Roeder, 1985). For either TATA less or TATA containing promoters, the whole TFIID complex binds the core promoter to initiate the PIC formation. TAF1/TAF2 dimer binds Inr element(Chalkley and Verrijzer, 1999). TAF1 is also able to bind DCE(Lee et al., 2005) and TAF6/TAF9 is able to bind both DPE(Burke and Kadonaga, 1996) and MTE(Shao et al., 2005). Some of the subunits are also able to interact with certain nuclear receptors(Lavigne et al., 1999) or to modify histone proteins(Mizzen et al., 1996; Pham and Sauer, 2000), in order to regulate transcription. Furthermore, TFIID interacts with TFIIF(Dikstein et al., 1996).

1.d. TFIIE

The general transcription factor TFIIE is composed of two copies of each subunit TFIIEα and TFIIEβ, forming a hetero-tetramer. It directly interacts with Pol II and DNA promoter as well as TFIID, TFIIF and TFIIH(Maxon et al., 1994; Okuda et al., 2004; Tanaka et al., 2009). It is mostly implicated in the control of TFIIH enzymatic activities and required for Pol II transcription, driving both

initiation(Ohkuma and Roeder, 1994) and transition to elongation(Watanabe et al., 2003).

1.e. TFIIF

The general transcription factor TFIIF is a complex of two proteins, RAP74 and RAP30. It is responsible, especially the small subunit, for the incorporation of Pol II into the PIC and its stable binding on promoter(Flores et al., 1991; Killeen et al., 1992). TFIIF is also essential in promoter escape and proper elongation(Yan et al., 1999; Zhang and Burton, 2004).

1.f. TFIIH

TFIIH is a dual protein complex, as it acts both in transcription and in the **N**ucleotide **E**xcision **R**epair (NER) pathway of DNA damage response. It is composed of two sub-complexes, the core and the **C**DK-**A**ctivating **K**inase (CAK). The first one is a six subunit ensemble containing Xeroderma Pigmentosum group B protein (XPB), p34/44/52/62 and the lastly discovered p8/TTDA(Giglia-Mari et al., 2004). The CAK is a three subunits complex composed by **M**énage-**À**-**T**rois 1 protein (MAT1), CDK7 and its associated cyclin H (CCNH). Those two sub-complexes are stabilized and held together by **X**eroderma **P**igmentosum group **D** protein (XPD)(Coin et al., 1998; Sandrock and Egly, 2001). In addition, **X**eroderma **P**igmentosum group **G** protein (XPG) that is known to participate with TFIIH in NER, have recently been found to directly interact with the core in order to stabilize the whole complex and to control its nuclear receptor phosphorylation activity(Ito et al., 2007).

The XPB protein possesses an ATP-dependent DNA helicase activity. A more recent study enlightened a XPB ATP-dependent translocase activity(Fishburn et al., 2015). TFIIH alters transcription initiation by making the PIC unstable(Plaschka et al., 2016). This block seems to be released by the translocase activity of XPB that would allow the promoter opening(Alekseev et al., 2017).

CDK7 works in collaboration with its partner CCNH and their interaction are stabilised by MAT1, the third member of the CAK(Devault et al., 1995). CDK7 is then able to phosphorylate the CTD Ser5 and Ser7 residues of the Pol II subunit Rpb1(Wong et al., 2014; Zhou et al., 2000). It also controls indirectly the phosphorylation of Ser2 residues by CDK9(Larochelle et al., 2012). Finally, CDK7

also phosphorylates TFIID, TFIIE and TFIIF general transcription factors(Ohkuma and Roeder, 1994).

CDK7 is also implicated in the cell cycle progression. XPD controls CDK7 activity (Chen et al., 2003). XPD notably control the cellular localisation of CDK7 to regulate its mitotic kinase activity and the chromosomal segregation(Li et al., 2010).

Table 3: General transcription factors

General Transcription Factors	Protein composition	Functions
TFIIA	TFIIAα, TFIIAβ and TFIIAγ	Stabilize TFIID
TFIIB	TFIIB	Stabilize TFIID-DNA-TFIIA, recruitment of Pol II/TFIIF
TFIID	TBP and TAF1-14	Bind to core promoter, initiate PIC formation, HPTM
TFIIE	TFIIEα and TFIIEβ (x2)	Control TFIIH enzymatic activity, elongation
TFIIF	RAP74 and RAP30	Recruitment of Pol II, promoter escape and elongation
TFIIH	XPB, XPD, P34, p44, p52, p62, p8/TTDA, MAT1, CCNH and CDK7 (XPG)	DNA opening, phosphorylation of Pol II CTD

Side to its roles in transcription, the TFIIH complex is a major actor of the NER DNA repair pathway, as described in the second part of the introduction.

2. PIC assembly and transcription process

There are several critical steps for transcription to be productive. It starts by the binding of PolI II and general transcription factors on the promoter, which assemble to form the PIC and to initiate the transcription. Then, Pol II performs the elongation along the gene body. Finally, termination occurs and a newly synthesised mRNA is released (Figure 2). While the Mediator complex is largely implicated in the formation of the PIC by helping the recruitment of the transcription machinery and by bridging transcription factors with the transcription machinery, its roles will be detailed in the third part of the introduction.

2.a. PIC formation

Transcription starts with the assemblage of the PIC around the TSS. It can happen in two different manners: the stepwise assembly or the holoenzyme assembly.

The stepwise assembly (aka sequential assembly) starts by the recognition of the promoter by TFIID. TBP first recognizes the TATA box, if it is present. In any case, other TFIID subunits recognise Inr and DPE to interact with the promoter. This interaction is then stabilized and strengthened by the arrival of TFIIA and TFIIB. Subsequently, Pol II joins the promoter in association with TFIIF. TFIIE and TFIIH are the last to be recruited to the PIC(He et al., 2013).

The Pol II holoenzyme consists of a large already assembled complex containing the Pol II, the Mediator complex and most of the general transcription factors except TFIID/TFIIA. In the holoenzyme pathway, it assembles without DNA and comes in almost one single step onto the promoter(Koleske and Young, 1995).

2.b. Initiation

From now on, TFIIH is essential for the opening of DNA. This ATP dependent step allows the transition from a closed conformation to an open conformation PIC. As mentioned above, various scenarios are possible for TFIIH. The DNA opening creates an 11-15 bp bubble and the non-coding strand is then inserted in the active site of Pol II. The first nucleotide synthesis converts the open conformation complex into an initially transcribing complex(Hantsche and Cramer, 2017). This complex enters several abortive initiations that only produces short oligonucleotides of less than 10 bp. If the newly synthetized RNA reaches 10 bp, then Pol II is phosphorylated by TFIIH and dissociates from the initiation complex, in order to start elongation. This step is called promoter clearance. After the RNA reaches a length of 25 bp, the capping machinery is recruited and this step allows productive transcription to start.

2.c. Proximal pausing/Elongation

Pol II encounters a pausing event shortly after starting productive transcription due to DSIF and NELF. Indeed, they bind Pol II and repress its elongation activity. The phosphorylation of NELF by P-TEFβ-associated CDK9 leads to its dissociation.

Pol II is phosphorylated on Ser2. DSIF is also phosphorylated and turn into an activator to promote productive elongation(Bernecky et al., 2017; Yamada et al., 2006).

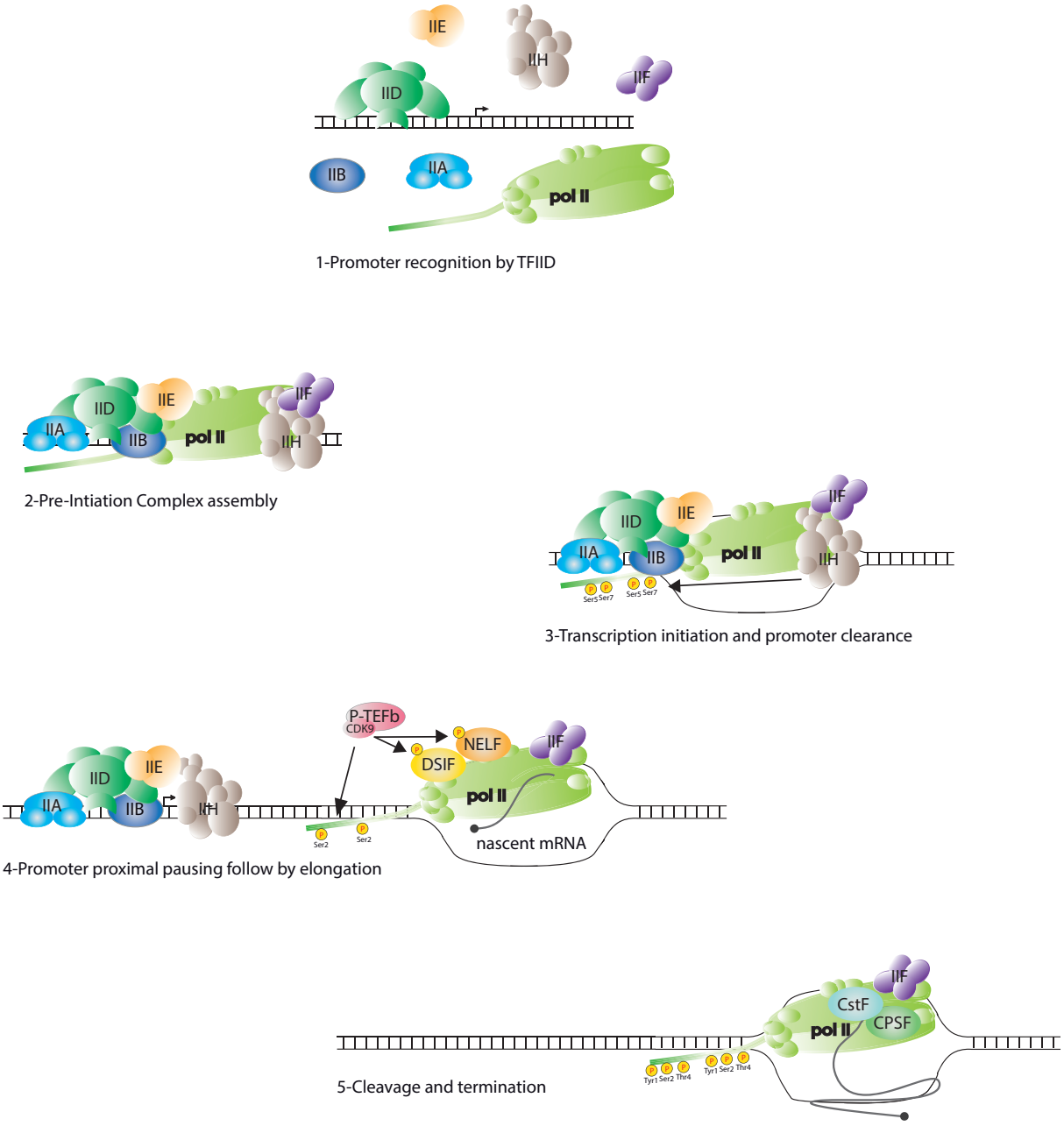


Figure 2: Transcription mechanism by Pol II

The diagram shows the promoter recognition and the recruitment of the different components of the transcription machinery as well as the cycle of transcription by Pol II.

Pol II continues to slide down to the 3' end of the gene, associated with the elongation factors TFIIS, ELL, Elongin and P-TEFβ. These factors sustain the DNA/RNA hybrid, that is necessary for Pol II processivity during elongation(Wind and Reines, 2000).

2.d. Termination/poly-adenylation

The termination occurs when the elongation complex reaches the 3' end of the gene. There are two main ways to end transcription, depending of the gene sequence(Kuehner et al., 2011).

The first pathway involves the **C**leavage and **P**olyadenylation **S**pecificity **F**actor (CPSF), which interacts with the Pol II CTD and recognizes a specific poly A signal at the 3' end of the gene. It induces the elongation to stop. Then, the **C**leavage **st**imulatory **F**actor (CstF) induces the cleavage of the RNA. Pol II is then released and the newly synthesised RNA is polyadenylated. The remaining part of the RNA, downstream of the cleavage site, is therefore degraded by XRN2. Most of the *human* genes follow this path.

But for other genes, the termination is achieved through the dissociation of the RNA/DNA complex by **Sen**ataxin 1 protein (Sen1). In such case, the resultant RNA is not polyadenylated.

4) Specific transcription factors

A large number of actors are required for transcription to occur in the right full place of the genome, in a well-controlled and coordinated manner. Several proteins and protein complexes are involved to initiate transcription including specific transcription factors, coactivators and the Mediator complex. The last one will extensively be explained in the third part of the introduction.

Specific transcription factors have been selected through evolution by the cells to answer various stimuli and stresses. Their role is to stimulate (activator) or to inhibit (repressor) transcription of specific genes to allow cell division, cell growth and cell death but also cell differentiation, stem cell maintenance or cellular response to stresses.

Specific transcription factors operate through enhancer and silencer regions. These genomic regions are enriched in specific DNA sequences called **R**esponse **E**lements (RE), where transcription factors can bind. Transcription factors are all characterised by a **D**NA **B**inding **D**omain (DBD) to directly contact DNA on specific RE and a **T**rans-**A**ctivation **D**omain (TAD) to transfer signal to the transcription machinery. The sum of activator and repressor signals finally determines the level of gene activity, ranging from no transcription to high transcribing activity.

Several types of transcription factor have been identified, depending on their structure, their mode of activation and the biological processes in which they are implicated.

1. E2F transcription factor family

The first E2F protein has been discovered as a cellular factor required for adenovirus E2 gene induction(Kovesdi et al., 1986), and lately defined as E2F1. Height different E2F transcription factors have been identified in this family and assigned as activators (E2F1-3) or repressors (E2F4-8). They bind DNA as heterodimer with a **D**imerization **P**artner (DP1-3). This partner is also able to modulate E2F activity(Bandara et al., 1994; Girling et al., 1993; Ingram et al., 2011; Ormondroyd et al., 1995). E2F transcription factors are involved in cell cycle progression and apoptosis, mainly through transcription regulation(DeGregori and Johnson, 2006; DeGregori et al., 1997).

E2F1 is implicated in G1-S phase progression(Yao et al., 2008), where it controls an all-or-nothing process. Division has to occur or cell would enter apoptosis(Wu et al., 2017). E2F1 mainly acts as transcription activator during G1/S phases(Araki et al., 2003; Burke et al., 2014). Indeed, it activates genes responsible for DNA synthesis. Moreover, acetylation of histone H3 and H4 accumulates during G1 phase in an E2F1 dependant manner, in order to open chromatin and allow gene activation as well as S phase entry(Taubert et al., 2004).

Then, it is made inactive during S/G2 phases(Schulze et al., 1995; Zhang et al., 2000). Notably, E2F1 is phosphorylated by TFIIH via an interaction with its p62 subunit. This interaction leads to E2F1 degradation during S phase(Vandel and Kouzarides, 1999).

E2F1 also have some roles in the NER pathway that will be presented in the second part of the introduction.

2. Nuclear receptors

Nuclear receptors are ligand-dependent transcription factors, limited to *Metazoans*. They are characterised by a particular **L**igand **B**inding **D**omain (LBD). They act either as monomers or dimers but need to bind their specific ligand to be activated and thus modulate transcription(Escriva et al., 2004).

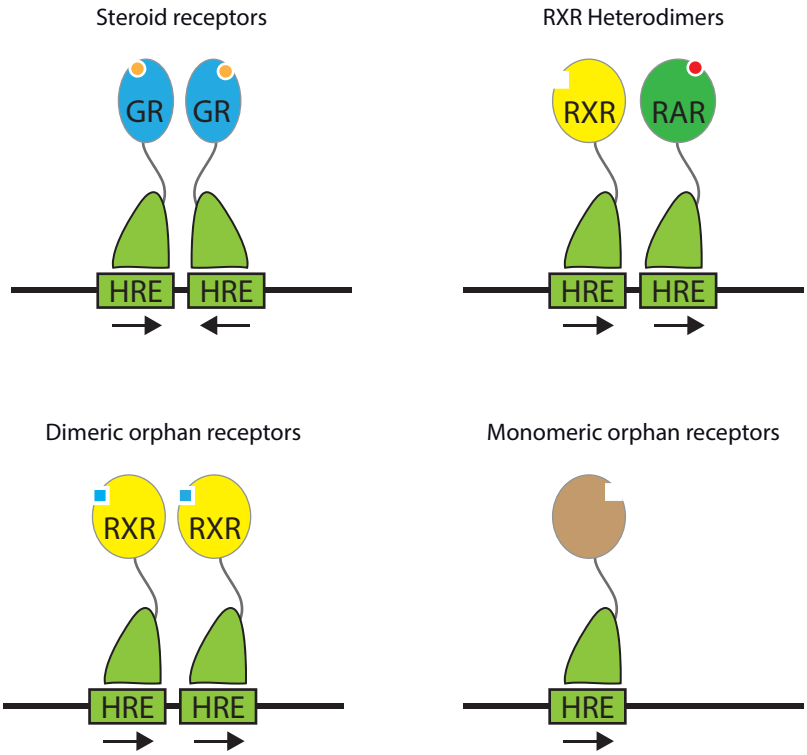


Figure 3: Nuclear receptor classes (adapted from Mangelsdorf et al., 1995)

2.a. Mechanistic classes

There are four main classes of nuclear receptors(Mangelsdorf et al., 1995). The first one includes steroid receptors that bind as homodimers to two repeated but inverted **H**ormone **R**esponse **E**lements (HRE), separated by several nucleotides. They are activated when associated to their steroid ligand. It includes **E**strogen **R**eceptor (ER), **P**rogesterone **R**eceptor (PR), **A**ndrogen **R**eceptor (AR), **G**lucocorticoid **R**eceptor (GR) and **M**ineralocorticoid **R**eceptor (MR). The second class corresponds to receptors that bind to DNA as a heterodimer complex with **R**etinoid **X** **R**eceptor (RXR), to direct repeat HRE. Among many others, it includes

Retinoic **A**cid **R**eceptor (RAR), **T**hyroid hormone **R**eceptor (T3R), **V**itamin **D** **R**eceptor (VDR) or **P**eroxisome **P**roliferator-**A**ctivated-**R**eceptor (PPAR). The third class corresponds to homodimeric nuclear receptors that bind direct repeat HRE. RXR, **T**esticular **R**eceptor (TR2) or **H**epatocyte **N**uclear **F**actor-4 (HNF-4) belongs to this class. Finally, the fourth and last class includes receptors that act either as monomers or dimers but when only one DNA binding domain contacts one half-site HRE. Most of the orphan receptors fall in the last two classes.

2.b. Retinoic acid receptor

Vitamin A is essential for several biological processes like embryogenesis, organ development, homeostasis, immunity and reproduction. Its biological significance is enlightened by fundamental studies but also by clinical observations. Indeed, deficiency in vitamin A can lead to growth retardation, intellectual disability and other symptoms of the **V**itamin **A** **D**eficiency (VAD) syndrome(Clagett-Dame and DeLuca, 2002; Zile, 2001).

Vitamin A corresponds to a group of organic compound including retinol, retinal and **A**ll-**T**rans **R**etinoic **A**cid (ATRA), the last one being the most biologically active form. Its physiological level is tightly regulated(Shannon et al., 2017). The retinoic acid receptor is a ligand-dependent transcriptional regulator. It includes three homologous receptors called RARα, RARβ and RARγ. They form a heterodimer with RXR and activate transcription when they are fixed by their ligand.

The heterodimer RAR/RXR binds to the **R**etinoic **A**cid **R**esponse **E**lement (RARE) that corresponds to two successive repeats of RGKTCA separated by variable number of nucleotides. When no ligand is associated with RAR/RXR dimer, co-repressors like **N**uclear receptor **Co**-**R**epressor (NCoR aka NCOR1) or **S**ilencing **M**ediator of **R**etinoic acid and **T**hyroid hormone receptor (SMRT aka NCOR2)(Perissi et al., 2010) are bound to it and serve as a platform for the recruitment of **H**istone **D**e-**AC**etylases (HDAC). The latter maintains chromatin in heterochromatin state, blocking transcription.

Upon ligand binding, RAR encounters conformational changes. This process leads to co-repressor release and coactivator arrival(Perissi et al., 2004). It is

accompanied by the recruitment of several **H**istone **A**cetyl **T**ransferase (HAT) and **H**istone **M**ethyl **T**ransferase (HMT), which open chromatin and enable transcription.

Retinoic acid seems able to partially modify the genomic location of certain RAR receptors through a dynamic process and to induce the transcription of previously unbound but very specific genes(Mahony et al., 2011).

3. Immediate early genes

3.a. Primary response

The transcription of certain genes can be activated very quickly and transiently by several stimuli without the need of new protein synthesis. These genes are divided into two classes: the **I**mmediate **E**arly response **G**enes (IEG) and the delay primary response genes(Tullai et al., 2007). Some of them code for transcription factors that subsequently induce the secondary response(Winkles, 1998).

IEG mRNA appear in the cell few minutes after stimulation. Their transcription is activated by various stimuli like growth factors, mitogens, phorbol esters, immunological and neurological signals or stresses (**U**ltra-**V**iolet (UV), toxins)(Fowler et al., 2011; Greenberg and Ziff, 1984; Herschman, 1991; Morgan and Curran, 1991).

The two most characterised IEG are JUN and FOS(Healy et al., 2013). They are playing roles in cell proliferation, differentiation and survival. They can be regulated by **P**ost-**T**ranslational **M**odifications (PTM), which influence their ability to form a dimer, thus to bind DNA and activate transcription. FOS and JUN form an active heterodimer complex called **A**ctivator **P**rotein-1 (AP-1), via a leucine zipper motif. When associated, they form a bipartite DBD. The complex binds DNA at specific responsive element called **T**etradecanoyl**P**horbol **A**cetate (TPA) **R**esponsive **E**lement (TRE).

Early **G**rowth **R**esponse-1 protein (EGR1) is also a transcription factor-coding IEG, involved in cell growth and differentiation. It binds DNA on specific consensus sequences and interacts with the transcription machinery(Liu et al., 2001; Zhang et al., 2003).

3.b. Secondary response

The secondary response corresponds to the late induction of a set of genes by the products of the primary response. They are called **L**ate **R**esponse **G**enes (LRGs), as well as secondary response genes.

5) Chromatin regulation of transcription

If DNA is bound by a large variety of protein in different dynamic processes, including transcription, it is rarely naked or freely accessible inside the nucleus. Indeed, DNA can be chemically modified to alter its accessibility. Moreover, it is incorporated into chromatin where it is physically compacted.

1. DNA methylation

DNA methylation is well conserved through evolution. It corresponds to the addition of a methyl group to the 5 position of a cytosine residue (5mC) in the context of a CpG dinucleotide. In mammals, DNA methylation is implicated in transcription regulation but also in X-inactivation, genomic imprinting and the silencing of transposable elements. The genomic pattern of methylation is established during embryonic life by DNA methyltranferase Dnmt3 and is latterly maintained by Dnmt1 during mitosis.

DNA methylation of promoter regions is associated with repression. The presence of methyl groups inside the major groove of DNA is supposed to engendered hindrance that perturbed DNA/TF contacts and block gene activation(Watt and Molloy, 1988). In the opposite, the binding of TF to DNA can also influence its methylation status and generate **L**ow **M**ethylated **R**egions (LMR)(Stadler et al., 2011).

DNA methylation also participates to gene repression by attracting proteins of the **M**ethyl CpG **B**inding **D**omain (MBD) family that are able to repress transcription(Bird and Wolffe, 1999).

To actively regulate transcription, DNA methylation is dynamically altered in cell. Notably, TET deoxygenase is responsible for 5mC oxidation to 5-hydroxymethylcytosine (5hmC), then to 5-formylcytosine (5fC) and finally to 5-

carboxycytosine (5caC). Both 5fC and 5caC can be removed by DNA glycosylase to be restored to regular cytosine by BER (Figure 4)(Weber et al., 2016).

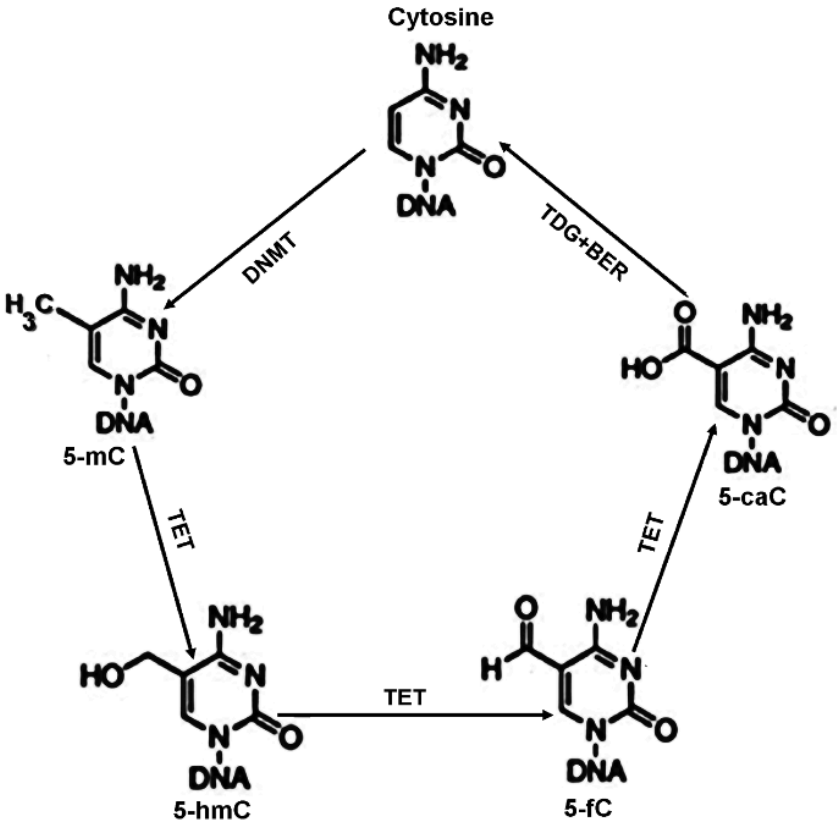


Figure 4: Cytosine methylation and demethylation (from Kang and Hyun, 2017)

2. Chromatin organisation

To protect it from physical damages but also as a transcription regulation mechanism, DNA is incorporated into a nucleoprotein complex called chromatin. The DNA molecule wrap around several histone proteins, in order to form a nucleosome. The chromatin is composed of a succession of nucleosomes and inter-nucleosome linker sequences. It appears as beads regularly positioned along a string. This three-dimensional conformation, called euchromatin, constitutes the first level of a necessary condensation of the DNA molecule into the nucleus. This is also a way to regulate its accessibility and thus transcription.

2.a. Nucleosome

The nucleosome also known as nucleosome core particle is composed of 146 bp of DNA wrapped around a core histone octamer in 1.67 turns, for a total of 206

kDa (Figure 5). This octamer is composed of two copies of four histone proteins named H2A, H2B, H3 and H4(Luger et al., 1997). Histones are small proteins weighting from 11 to 22 kDa and well-conserved trough evolution. They are enriched in positive amino acids, like lysine and arginine, to interact with negatively charged DNA. The assembling of the octamer starts by the formation of H3-H4 dimers. Next, two H3-H4 dimers are assembled in a tetramer. Finally, two H2A-H2B dimers are added to the tetramer to form a complete octamer(Arents and Moudrianakis, 1995).This “beads on a string” fibre has a diameter of 11 nm and corresponds to the first level of DNA condensation.

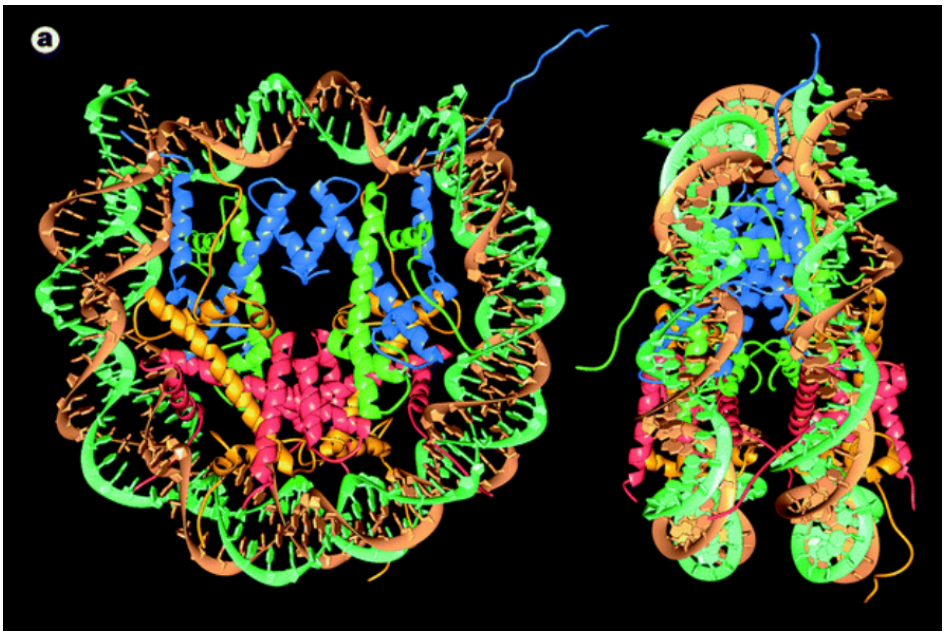


Figure 5: Nucleosome core particle (from Luger et al., 1997)

A ninth “linker” histone named H1 can bind to linker sequences and form a higher ordered and more stable structure with a diameter of 30 nm(Robinson et al., 2006). Each nucleosome then encompasses 160 bp(Syed et al., 2010). Two hypotheses have been proposed for the structure of the 30 nm fibres. The solenoid model predicts that adjacent nucleosome follow a super-helicoidal path with 6 to 8 nucleosome per turn(Widom and Klug, 1985). The zigzag model postulates two nucleosomes with straight linker and successively interleaved(Chen and Li, 2010). Higher structure compactions are less clear but also affect transcription, as it alters access to DNA. This more compacted chromatin is called heterochromatin and does not allow transcription. In the cellular context, chromatin seems heterogeneous

depending on genomic regions or cell cycles phases. It implies dynamic processes to pack and unpack the chromatin and control the access to genetic information.

2.b. Histone variants

Histone proteins are divided into two classes. The canonical one includes the most abundant form of histones. They are produced during the S phase and are the major support for genomic DNA. They are produced along with the newly synthesised DNA and assembled with it. But beside the canonical forms, several histone variants have emerged through evolution(Talbert et al., 2012). They are produced all along the cell cycle and are dedicated to specific functions like transcription initiation and termination, cell division or maintenance of genome integrity. Variants exist for the core histone H2A, H2B, H3 and for the linker histone H1. They differ by few amino acids or additional domains but keep the common histone structure to allow an easy exchange with their canonical counterpart(Venkatesh and Workman, 2015).

Amongst many, the histone variant **H3.3** is notably found at enhancer, promoter and in the body of actively transcribed genes, along with elongating Pol II. H3.3 turnover seems to compensate Pol II associated nucleosome displacement(Wirbelauer et al., 2005). In addition, H3.3 deposition at active promoters has been correlated to replication-independent E2F gene activation(Daury et al., 2006).

H2A.Z variant is highly conserved across evolution. It has strong sequence differences with canonical H2A but their similar three-dimensional structures allow them to be exchanged(Talbert and Henikoff, 2010). Subtle structural differences make H2A.Z containing nucleosomes less stable. Indeed, H2A.Z/H2B dimer has an affected interface with H3/H4 tetramer but also with canonical H2A/H2B dimer(Suto et al., 2000).

H2A.Z/H3.3 containing nucleosomes are highly unstable and accumulate at previously considered “nucleosome free region”, around promoters, enhancers and insulators. They maintain an open chromatin state and facilitate transcription factor accessibility(Jin et al., 2009).

2.c. Histone post-translational modifications

Histones Post-Translational Modifications (HPTM) also impacts transcription. Indeed, unfolded N-terminal tail of histones protrudes outside the nucleosome and is subjected to various enzymatic activities like phosphorylation, methylation, acetylation, ubiquitination, phosphorylation, citrullination, isomerisation and few others more recently discovered (Figure 6)(Kouzarides, 2007). Some modifications have also been identified inside the octamer core(Yu et al., 2012) or on the small protruding C-terminal part of certain histones. Modifications are generally present at specific spatial and/or temporal localisation and associated with either active transcription or repression.

Figure 6: Histones post-translational modifications (inspired from Kouzarides, 2007)

Histone modifications can modify nucleosome electric charges and impact histone/histone or histone/DNA contacts. They can also serve as a recruitment platform for other proteins that will further act on chromatin or transcription.

A specific nomenclature has been established(Turner, 2005): “H3K9ac”, **H3** corresponds to the modified histone, **K9** describes the amino acid lysine and its

position in the protein and **ac** define the chemical modification (ac: acetylation, me: mono-methylation, me2/3: di- or –tri-methylation, ub: ubiquitination, ph: phosphorylation, *et caetera*).

i. Acetylation

Acetylation is the first described HPTM(Allfrey, 1966), but it also targets other proteins(Choudhary et al., 2009). **H**istone **A**cetyl **T**ransferase (HAT) can mediate N-acetylation of lysine residues(Kouzarides, 2007). It uses Acetyl-CoA as acetyl donor to catalyse the reaction. Acetylation suppresses the lysine's positive charge and diminishes their interaction with the negatively charged DNA. It leads to chromatin decondensation and is generally associated with transcriptionally active genes(Spotswood and Turner, 2002). The reverse reaction is called deacetylation and is catalysed by histone deacetylase, to restore the lysine's positive charge.

HAT are usually divided into two classes. Type A HAT are located in the nucleus, where they acetylate chromatin-associated histones. Type B HAT are located in the cytoplasm and are responsible for the acetylation of newly synthesised histones.

Type A HAT are subdivided into three families: i) **G**CN5-related **N**-**A**cetyl **T**ransferase (GNAT) that includes **G**eneral **C**ontrol of **N**utrition protein 5 (GCN5) and **p**300/**C**BP-**A**ssociated **F**actor (PCAF), ii) MYST family (**M**OF - **Y**bf2/Sas3 - **S**as2 - **T**ip60) and iii) others like **C**REB-**B**inding **P**rotein (CBP) and p300(Kimura et al., 2005).

GCN5 also know as Lysine (**K**) **A**cetyl**T**ransferase 2A (KAT2A) is responsible for H3K14 acetylation alone but it is also included in two different complexes named **A**da **T**wo **A**-**C**ontaining (ATAC) and **S**pt-**A**da-**G**cn5 **A**cetyltransferase (SAGA). Both complexes stimulate its acetylation activity, especially ATAC(Riss et al., 2015). In such protein complexes, GCN5 is found to acetylate both H3K9 and H3K14 but on distinct genomic positions. ATAC associated GCN5 markedly targets enhancer regions(Krebs et al., 2011; Nagy et al., 2010). Whether incorporated within the SAGA complex, GCN5 also acetylate H3K18 and H3K23 residues(Rodríguez-Navarro, 2009).

In the context of nuclear receptor activation, GCN5/PCAF and CBP/p300 have been shown to be recruited on the promoter of NR target genes and to be responsible for H3K9ac and H3K18/27ac respectively(Jin et al., 2011). E2F Transcription factors also activate transcription by contacting GCN5 and by bringing it onto promoters. Moreover, *E2F1* itself is activated by GCN5/E2F1(Chen et al., 2013; Lang et al., 2001). *GCN5* could also be regulated by E2F1(Yin et al., 2015).

HDAC are the enzymes responsible for the removal of acetyl group from lysine residues. They are divided in four classes because of their structural differences, using Zn^{2+} as cofactor except for the Sirtuin class (Sirt) that uses NAD^{+}(Thaler and Mercurio, 2014). HDAC have low specificity and need to be regulated. As example, Sirt1 requires TFIIH to contact and deacetylate the cofactor PGC-1α(Traboulsi et al., 2014).

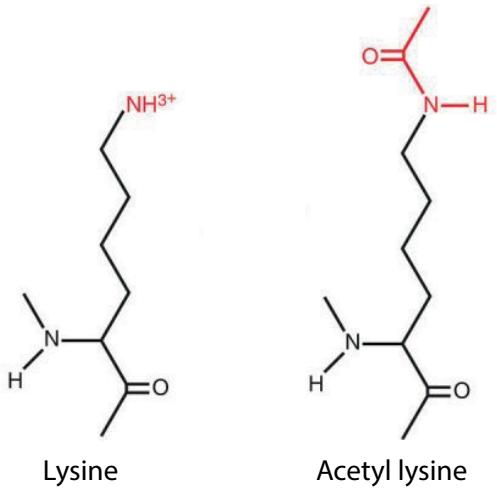


Figure 7: Lysine acetylation

ii. Methylation

Histone methylation can have various effects depending on the targeted residue and its position. Methylation targets both lysine and arginine residues but do not alters their charge or the histone structure. Methylated residues seem to serve as a platform for the recruitment of different proteins, able to alter chromatin state or transcription status. Lysine can be mono-, di-, or tri-methylated with variable consequences, but arginine can only be mono- or di-methylated(Bedford and Clarke, 2009; Greer and Shi, 2012). The enzymatic reaction is performed by **H**istone **M**ethyl-

Transferase (HMT), which transfers methyl group from S-adenosyl methionine to the targeted residue.

Three families of enzymes have been discovered. The SET-domain-containing proteins and the DOT1-like proteins can both acetylate lysine residues. **P**rotein a**R**ginine **M**ethyl **T**ransferases (PRMT) acetylate arginine residues. H3K4 is notably tri-methylated by SET1/COMPASS complex around the promoter of active genes. H3K36me marks are also enriched along the gene body of highly transcribed genes.

Two families of demethylase are characterised, depending on their active domain; amine oxidases and jumonji-C-domain containing, iron dependant dioxygenases (jJmjC).

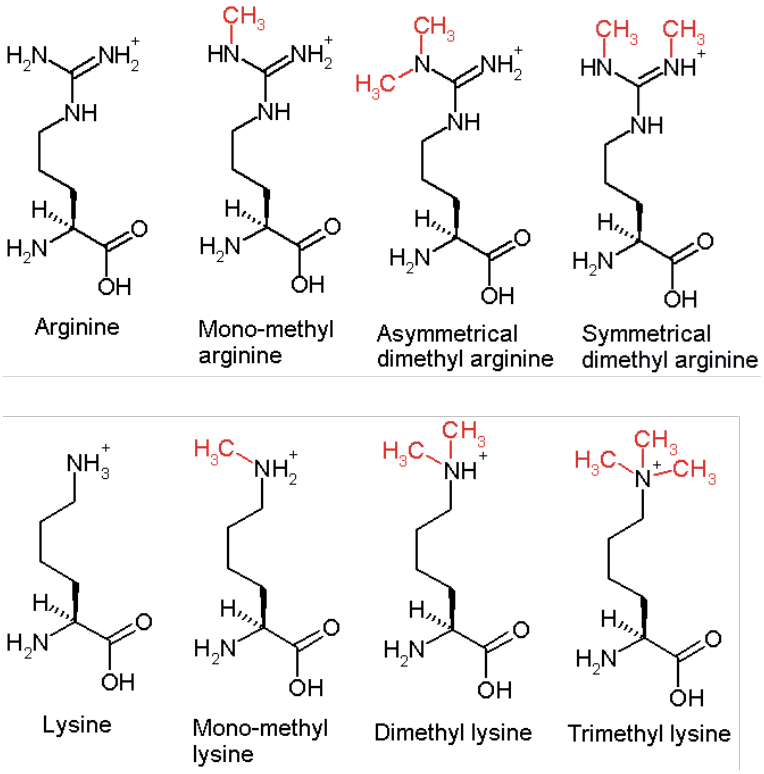


Figure 8: Arginine and Lysine methylation

2.d. Regulation cross-talk

Importantly, it has recently been shown that numerous HPTM may influence one another. For example, mono-ubiquitination of H2B promotes methylation of H3K4 and H3K79. Otherwise, repressive H3K9me3 marks need to be actively removed before the active H3K9ac marks to be added by p300 in CD4+ T cells(Ghare et al., 2014).

Histone variants influence the deposition of histone marks. Indeed, H2A.Z containing nucleosomes are more prone to H3K9ac and H3K4me3 marks deposition(Bártfai et al., 2010). DNA methylation status can also influence HPTM. MBD proteins are notably associated with various HDAC(Bird and Wolffe, 1999). DNA methylation is also thought to promote H3K9 methylation.

These interrelations between histone post-translational modifications, histone variants and DNA methylation status constitute a supplemental level of transcription regulation.

3. Chromatin rearrangement

Gene expression can also be regulated through very large distance. Rearrangement corresponds to large chromatin loops that allow to bring closer a specific gene and a distant regulatory element, either upstream or downstream, in order to regulate transcription. When this element is located on the same DNA molecule, it is called cis-regulatory element. On the contrary, when the regulatory element is located on another chromosome, it is called trans-regulatory element.

This has first been enlightened by the discovery of cis-regulation of the β-globin gene over a 150 kb distance(Vernimmen et al., 2007). Two kinds of loops have then been described. The first one allows physical rapprochement between enhancers and promoters(Krivega and Dean, 2012) while the second one brings close the promoters and the terminators(Lainé et al., 2009). These loops have also been described to be highly dynamic. For example, they can be induced under hormonal stimulation(Tan-Wong et al., 2008).

These loop conformations are maintained by a large number of protein. Among many one, CCCTC-binding protein (CTCF) is the strongest and most evolutionary conserved. Two copies of the protein can bind **C**TCF-**P**air-**D**efined domain (CPD) sequences to gather them. CTCF is implicated in gene activation or repression but also different functions like enhancer blocking, X-chromosome inactivation or genome imprinting. Furthermore, it was found to interact with Dnmt1 to inhibit DNA methylation. CTCF is thought to organise the three-dimensional conformation of the genome and to direct a large transcription regulation system.

II - DNA damage response

All along the life, genomic DNA is exposed to a large number of genotoxic attacks, either endogenous or exogenous. Indeed, biological by-products, UV rays, cigarette smoke and environmental chemicals are able to alter the genomic integrity and to create DNA lesions. These issues have to be properly managed and repaired in order not to perturb essential processes like DNA transcription and replication, which are both crucial for cells. The lack of DNA damage response may lead to cell transformation or death.

Different repair mechanisms have been selected through evolution. They are differentially activated depending on the lesion type, in order to restore the genomic integrity. It includes **N**ucleotide **E**xcision **R**epair (NER) that is responsible for the removal of UV-induced lesions and bulky DNA adducts, **B**ase **E**xcision **R**epair (BER) that is in charge of single strand breaks, the **N**on-**H**omologous **E**nd **J**oining (NHEJ) and **H**omologous **R**epair (HR) pathways, which are responsible for DNA double strands breaks religation and **M**is **M**atch **R**epair (MMR) that corrects replication issues.

1) Nucleotide excision repair

NER pathways are engaged in the removal of cyclobutane-pyrimidine dimers (CPD) and 6-4 pyrimidine pyrimidone photoproducts (6-4) PP that are induced by UV rays. They are also responsible for the removal of bulky DNA adducts induced by polycyclic aromatic hydrocarbon or anti-cancerous chemicals(Mocquet et al., 2007).

Two different NER pathways have been discovered, depending on the activating signal: the **G**lobal-**G**enome NER (GG-NER) and the **T**ranscription-**C**oupled NER (TC-NER).

1. GG-NER

As its name indicates, GG-NER acts throughout the genome to search for lesions in order to initiate their reparation.

1.a. Lesions recognition

DNA lesions are recognised all over the genome by the **X**eroderma **P**igmentosum complementation group **C** protein (XPC) in complex with its partners **H**omologue **R**ad 23 **B** (HR23B) and Centrin2 (CETN2)(Araki et al., 2001; Masutani et al., 1994). These two proteins are necessary to initiate NER. Indeed, if XPC-HR23B duplex is sufficient for *in vitro* NER reaction, CETN2 strongly increases *in vivo* reaction(Araújo et al., 2000; Nishi et al., 2005).

Chemical adducts or photoproducts induce DNA distortions that are recognised by XPC. Nevertheless, only the abnormal DNA structure is recognised by XPC, the presence of a chemical lesion has to be subsequently confirmed(Sugasawa et al., 2001, 2002). Along with this recognition mechanism, less distorting-DNA lesions are not easily identified by XPC/HR23B/CETN2 complex. In such cases, the damage is first recognised by **D**amage specific **D**NA **B**inding protein 2 (DDB2/XPE) in complex with DDB1. This DDB complex enhances the DNA disruption and thus helps to recruit XPC(Fitch et al., 2003). In such case, the DDB-associated Cul4 protein ubiquitinates both DDB2 and XPC to allow NER to initiate(Sugasawa et al., 2005).

1.b. DNA opening

When bound to the damage, XPC slightly distorts the DNA and recruits the TFIIH complex(Volker et al., 2001). As mentioned above, TFIIH is a dual complex. Besides its role in transcription, it is an essential protagonist of NER pathways. Indeed, both XPB and XPD are required for DNA opening around the lesion. The mechanism is not yet fully understood but it seems that XPD helicase activity is responsible for double helix unwinding and opening, when XPB allows the proper TFIIH anchorage on DNA(Oksenych et al., 2009). The CAK sub-complex seems to interfere with the overall repair mechanism and needs to be removed from TFIIH during NER.

When DNA strands are separated around the lesion, the protein RPA binds the undamaged strand to protect it from nuclease activities and to stabilize the open structure(Volker et al., 2001). This step allows the arrival of XPA and the release of the CAK(Coin et al., 2008). RPA is notably acetylated by GCN5/PCAF. This

acetylation promotes XPA/RPA interaction and increases the retention of XPA to the damage site(He et al., 2017; Zhao et al., 2017).

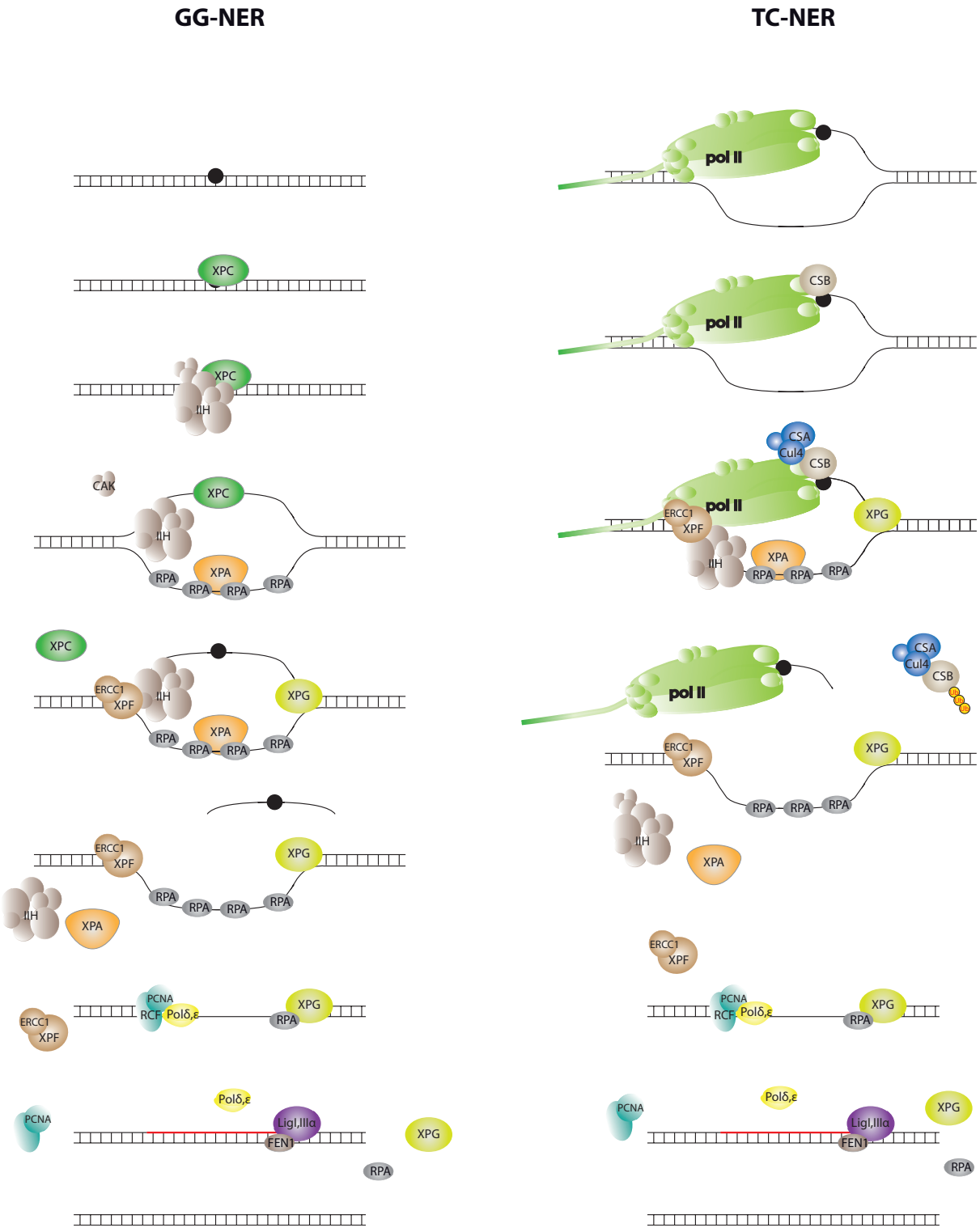


Figure 9: The two NER pathways

1.c. DNA incision

XPA promotes the recruitment of XPF-ERCC1 endonuclease and XPG is recruited on TFIIH, leading to the removal of XPC. XPG incises the damaged strand at the 3' extremity of the lesion bubble and induces the 5' end cleavage by XPF-ERCC1. A short 22-30 bp long oligonucleotide is released(Riedl et al., 2003). XPA is then removed from the lesion.

1.d. Synthesis

After incision, only RPA stays on the undamaged strand to protect it from nuclease activities. The double incision allows the recruitment of PCNA and RCF. XPF is then released and the remaining RPA, XPG, RCF and PCNA allow the recruitment of DNA polymerase δ or ε. After the release of XPG and RPA, synthesis of the complementary strand is performed using the undamaged strand as template(Aboussekhra et al., 1995; Moser et al., 2007).

1.e. Ligation

After the synthesis of the complementary strand, a ligase protein performs the necessary ligation of DNA. The implicated Ligase, as well as the responsible DNA polymerase, seems to depend on the cell cycle status. Indeed, the couple Pol δ/Ligase IIIα is in charge of synthesis/ligation throughout the cell cycle while Pol ε and Ligase I are in charge of synthesis/ligation in quiescent cells(Moser et al., 2007).

2. TC-NER

TC-NER is dedicated to the repair of the transcribed strand(Hu et al., 2015). It rapidly arises when the RNA polymerase II is stalled on DNA because of a lesion, during the transcription process. It seems particularly active around promoter and enhancer. Contrary to GG-NER, it removes a specific lesion to ensure the proper termination of a crucial ongoing function. In this context, XPC, XPE and DDB proteins are absolutely dispensable.

2.a. Stalled Pol II

The arrest of Pol II on UV-induced DNA damage strengthens its interaction with Cockaine Syndrome complementation group B protein (CSB)(van den Boom et al., 2004). CSB is further stabilized by its de-ubiquitination by USP7(Schwertman et

al., 2013). From now on, it strongly binds DNA and alters the Pol II/DNA interaction(Beerens et al., 2005). CSB then induces the recruitment of the CSA complex, which notably contains Cul4(Fousteri et al., 2006). CSB also induces the recruitment of TFIIH and the other NER factors RPA, XPA and the endonucleases XPF-ERCC1 and XPG.

2.b. Re-initiation

When NER machinery is set up, the regular repair pathway allows the proper removal of the lesion. After the restoration of genomic integrity, CSA complex re-ubiquitinate CSB to allow its degradation by the proteasome(Groisman et al., 2006). This last step is required for normal transcription re-initiation.

2.c. Pol II role in TC-NER

The role of Pol II during the process of TC-NER is not completely understood. On one hand, the presence of Pol II at the lesion does not prevent dual incision from occurring *in vitro*(Tremeau-Bravard et al., 2004). On the other hand, it has been shown that Pol II can be ubiquitinated, in order to be degraded(Lee et al., 2002). Finally, CSA also helps the recruitment of chromatin remodelling factors, including HMGN1, which is responsible for H3K14 acetylation. It has been proposed that local chromatin remodelling would help Pol II to get back from the lesion, in order to allow NER(Hanawalt and Spivak, 2008).

2) NER associated diseases

Mutations in NER coding genes are responsible for different syndromes, associated with photosensitivity. Indeed, the genetic mutations lead to partial or total impairment of UV-induced NER pathways (Table 4).

1. Xeroderma pigmentosum

Xeroderma pigmentosum (XP) is the first pathology that has been related to DNA repair deficiency(Cleaver, 1968). Patients were originally described for photosensitivity and for their high rate of fatal skin cancer, with hereditary transmission. Indeed, they have a 1000 increased risk rate for skin but also eye and tongue cancer. Twenty per cent of them also develop progressive neurological degeneration. Furthermore, they suffer from a slightly higher incidence of internal

cancer. There are seven complementation groups, named from XPA to XPG. They are accountable to mutations in related protein coding genes. XPC and XPE patients have a deficient GG-NER while the five other complementation groups endure both TC-NER and GG-NER deficiencies.

More recently, an eighth complementation group has been described, XPV. It arises from mutations in the gene coding for Pol η. This DNA polymerase is usually responsible for error-free replication of UV lesions(Johnson et al., 1999). It also results in photosensitivity and skin cancer, tethering DNA replication.

2. Cockaine syndrome

Patients are photosensitive but show no predisposition to skin cancer. Cockaine syndrome (CS), also known as progeroid dwarfism, is also characterised by microcephaly, defective growth and progressive neurological disorder. There are several levels of severity ranging from the moderate type I CS to the most severely affected patient showing pre-natal onset(Laugel, 2013). There are two complementation groups: CSA and CSB.

Furthermore, few mutations in XPB, XPD XPG and ERCC1 have been found to give rise to dual XP/CS syndrome, associated with a particular propensity to skin cancer.

3. Trichothiodystrophy

Trichothiodystrophy (TTD) patients exhibit intellectual disability with hereditary transmission. Scattered and brittle hairs as well as brittle nails mostly characterize patients. The “tiger-tail” of hair extremity is diagnostic of the disease. They can also manifest photosensitivity, icthiosis, short stature and decreased fertility. The photosensitive patients bear mutations in XPB, XPD or p8/TTDA. XPD mutations have also been reported to be responsible for dual XP/TTD patients. Few TFIIE mutations was also shown to be responsible for non-photosensitive TTD(Kuschal et al., 2016).

Table 4: NER factors and associated pathologies

Protein	Activity	Associated pathologies
XPA	Damage sensor	XP
XPB	3'-5' ATP dpdt helicase	XP; XP/CS; TTD
XPC	Damage sensor, chromatin remodelling	XP
XPD	5'-3' ATP dpdt helicase; 5'-3' ATP dpdt translocase	XP; XP/CS; TTD; XP/TTD
XPE	Damage sensor	XP
XPF	Endonuclease	XP
XPG	Endonuclease	XP; XPC/CS
XPV	DNA polymerase η	XP
CSA	Ubiquitin ligase	CS
CSB	DNA dpdt ATPase of SWI/SNF family	CS
p8/TTDA	Involved in NER	TTD

4. Others

Mutations in NER factors have also been associated with XFE progeroid syndrome and cerebello-oculo-facio-skeletal syndrome (COFS), both associated with photosensitivity and NER defect.

5. Aetiology

Photosensitivity and increased risk of skin cancer are notably explained by the DNA repair defect associated with the causative mutations. Mutations in NER factors strongly impair both TC-NER and GG-NER. Nevertheless, strong clinical features like neurodevelopmental defect, intellectual disabilities, progeria or dwarfism also accompany Xeroderma pigmentosum, Cockain syndrome and Trichothiodystrophy. Such phenotypic features are much more difficult to explain by a single DNA repair defect. While first characterized as DNA repair syndromes, recent works tend to suggest that some features arise from transcription impairments(Brooks, 2013; Compe and Egly, 2016). Moreover, several NER factors have recently been found to participle in transcription, especially *via* chromatin remodelling.

3) NER in transcription

As mentioned earlier, transcription of protein coding genes is a fundamental process of living cell that is carried out by a dedicated enzyme, the RNA polymerase,

with the help of several complexes, including chromatin remodellers, transcription factors, co-factors and general transcription factors.

Thus, the DNA template requires to be carefully maintained as any damage could alter the process of transcription. Every lesion has to be repaired when transcription stalled on it. These conclusions suggest a functional link between transcription and DNA repair.

TFIIH was the first reported complex to be implicated in both transcription and DNA damage response. It was first identified as part of the general transcription machinery(Matsui et al., 1980) before to be characterised as the helicase of nucleotide excision repair(Schaeffer et al., 1993). TFIIH is notably implicated in transcription initiation, promoter escape and elongation but also in NR phosphorylation(Drané et al., 2004; Rochette-Egly et al., 1997).

Mediator was recently suggested to allow the switch from transcription to DNA repair by facilitating the recruitment of repair factors in the context of DNA lesions(Eyboulet et al., 2013; Kikuchi et al., 2015).

Aside from TFIIH, other NER factors have also been implicated in the process of transcription. As mentioned earlier, the endonuclease XPG stabilizes the TFIIH complex in the context of transcription. It is notably required for nuclear receptors-associated transcription activation(Ito et al., 2007). A recent study also suggests that a XPG-TFIIH complex could play a role as elongation factor(Narita et al., 2015).

XPG was also shown to be required for the DNA demethylation while GADD45α induced gene activation(Barreto et al., 2007; Schmitz et al., 2009). Both XPG and XPF endonucleases have been shown to be required for promoter-terminator loop formation and optimal transcription. Indeed, XPG promotes DNA breaks and DNA demethylation at gene promoters to allow the recruitment of CTCF and gene looping(Le May et al., 2012).

XPC was shown to regulate the stem state of embryonic stem (ES) cells in the context of a trimeric complex called **S**tem **C**ell **C**oactivator (SCC), together with HR23B and CETN2. Indeed, Oct4/Sox2 recruits XPC through a direct interaction on the promoter of *Nanog* and *Oct4* genes as well as most Oct4/Sox2 bound genomic

regions, in order to orchestrate the ES cell-specific gene expression(Fong et al., 2011). In this context, XPC-containing SCC seems required for the maintenance of either natural or induced pluripotency(Cattoglio et al., 2015).

A recent study showed XPC to regulate TDG-dependent DNA demethylation in both somatic and stem cells but also in the context of cell reprogramming to generate induced-pluripotent stem cells(Ho et al., 2017).

Recently, our laboratory uncovered a major role of XPC during transcription(Le May et al., 2010a), in absence of any genotoxic attack. XPC was found to bind the promoter of activated gene after the recruitment of the transcription machinery. The arrival of XPC at the promoter of activated genes engendered the sequential recruitment of other NER factor including XPA, RPA, XPG and XPF-ERCC1.

Moreover, these NER factors were found to be necessary for optimal transcription. Indeed, they seem required for DNA demethylation as well as active histone PTM like H3K9/14 acetylation and H3K4/K9 methylation around the promoter. Deficiencies in NER factor and notably XPC were found to impair the recruitment of the other NER factors and to alter gene transactivation. Therefore, the comprehension of the exact role of XPC and the other NER factors in such processes is of great interest, both fundamentally and clinically.

III - Mediator complex

Manuscript in preparation

The Mediator complex and associated human genetic disorders!

BIDON Baptiste[∞], DONNIO Lise-Marie[∞], COIN Frederic & EGLY Jean-Marc*

Département de génomique fonctionnelle et Cancer, CNRS, INSERM, Université de Strasbourg, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 1 rue Laurent Fries, 67404 Illkirch-Graffenstaden, France.

* To whom correspondences should be address:

egly@igbmc.fr

∞ Should be regarded as joint First Authors

ABSTRACT

Mediator (MED) is a multi-subunit complex that plays a central role in the regulation of protein coding genes transcribed by RNA polymerase II (Pol II). MED conveys essential information from DNA-bound transcription factors to the basal Pol II transcription machinery. Although many studies have been engaged on MED, very few are known concerning the function of each subunit. This last decade, genetic mutations in MED subunits were shown to cause genetic diseases with common symptoms like congenital malformation and intellectual disability. After a brief description of the basic features of MED, the review will describe the mutation-associated syndromes and discuss their aetiology.

Key words: Mediator, Transcription, Genetic disorders & X-linked Intellectual Disability.

INTRODUCTION

Mediator (MED) is an evolutionarily conserved multi-protein complex that is a key regulator of protein-coding genes. MED is composed of 25 subunits in yeast and more than 30 subunits in higher eukaryotes. They are organized into three core modules (Head, Middle and Tail) and a dissociable Kinase module (Figure 1). MED is mainly implicated in transcription initiation and functions as a “molecular bridge” that conveys essential information from transcription factors bound at upstream DNA responsive elements to the basal transcription machinery, formed by RNA Polymerase II (Pol II) and the General transcription Factors (GTFs). By means of this, MED appears to be implicated in enhancer/promoter loop formation but also in the regulation of other transcriptional events such as elongation(Conaway and Conaway, 2013), termination(Mukundan and Ansari, 2013), mRNA processing(Huang et al., 2012) as well as chromatin remodelling(Hsieh et al., 2015; Kagey et al., 2010; Zhu et al., 2011). From a larger perspective, MED is strongly implicated in the transcriptional regulation of developmental processes(Yin and Wang, 2014). These key roles in gene regulation might thus explain why mutations in human are often associated with developmental disorders or cancers(Schiano et al., 2014a, 2014b).

Discovery of the Mediator complex

Before 1990, although the basic transcription mechanism was deciphered and the essential protagonists were identified, very little was known about gene regulation. Only scarce information was available about co-factors (either activators or repressors). But unconsciously, people were already working on some MED subunits, like Gal11 described to be required in galactose responsive genes and later considered as MED15(Nogi and Fukasawa, 1980) or RGR1 implicated in the response to glucose and corresponding to MED14(Sakai et al., 1988) as well as Sin4 required for the production of tRNAs which corresponds to MED16(Stadelmann et al., 1986).

In 1990, the Kornberg lab proposed the term ‘‘Mediator’’ to describe the recruitment of a yeast protein complex required to “mediate” signal transduction from Gal4-VP16 chimeric activators to Pol II, in an in vitro transcription assay(Kelleher et

al., 1990). In the followings years, they initiate the identification of the MED composition and delineate its function(Flanagan et al., 1991; Kim et al., 1994).

The human counter-part of yeast Mediator was identified in HeLa cells and named TRAP (Thyroid hormone receptor-associated protein), for its ability to associate with the Thyroid hormone Receptor α (TRα) and to potentiate (TRα)-regulated *in vitro* transcription(Fondell et al., 1996). Subsequently, different Mediator related complexes were identified for their abilities to activate transcription, like ARC (activator-recruited cofactor)/DRIP (vitamin D receptor-interacting proteins)(Näär et al., 1999; Rachez et al., 1998), CRSP (cofactor required for Sp1 activation)(Ryu et al., 1999) and PC2 (positive cofactor 2)(Malik et al., 2000). MED was thus considered as a global regulator of transcription activation.

Multidimensional Protein Identification Technology (MudPID) later stated that all these proteins are representative of the same Mediator complex, enlightening the composition of the minimal 26 subunits core complex and the associated 4 subunits kinase module(Sato et al., 2004). Therefore, a universal nomenclature for each of the 30 MED subunits was admitted(Bourbon et al., 2004). Three of the kinase module coding genes were found to have undergone duplications that generate the paralogs pairs MED12/MED12L, MED13/MED13L and CDK8/CDK19. Mass spectrometry analysis has recently shown that the presence of one paralog of the pair is mutually exclusive with the other one(Daniels et al., 2013).

Composition and Structure of the Mediator

Considering the various MED-like complexes that have been discovered, some questions arise. Are there several possible combinations of subunits, as occurs for SAGA, ATAC or TFIID complexes? Is Mediator composition or architecture reorganized during various steps of transcription? In the light of the chromatin environment, we can also question on potential post-translational modifications like phosphorylation, ubiquitination or acetylation and their effects in the composition and the structure of MED, as a function of the gene and the response element binding factors.

However, its size (up to 1.5 MDa), its intrinsic flexibility, and its apparent composition heterogeneity engender technical difficulties to obtain large amounts of

highly purified complex, making the high-resolution structure of the complete MED, a challenging topic. The first structure obtained at 40Å resolution, showed an extended protein complex that envelop Pol II. It also showed the complex conservation between yeast and mice(Asturias et al., 1999). During the following decade, several low resolution electron microscopy (EM) structures were successively released(Davis et al., 2002; Dotson et al., 2000; Meyer et al., 2010; Taatjes et al., 2004), establishing the conservation in eukaryotes and enlightening the modular organisation. They mostly showed that Head and Middle modules interact with Pol II while Tail module contacts activators. Each activator was shown to induce a specific MED conformational change.

In parallel, high-resolution X-ray structures were determined for smaller entities e.g. Cyclin C (CCNC)(Hoeppner et al., 2005), MED7/MED21(Baumli et al., 2005; Koschubs et al., 2009; Larivière et al., 2006) or MED8/MED18/MED20. Larger module structures have later been characterized with high resolution, such as the crystal structure of *S. pombe* or *S. cerevisiae* head module(Imasaki et al., 2011; Larivière et al., 2012) as well as the partial structure of *S. cerevisiae* Middle module(Larivière et al., 2013), showing their internal organisation and their interactions with the general transcription machinery.

More recently, an accurate electron microscopy structure of the yeast MED (yMED) has been obtained(Tsai et al., 2014). The localization of head and middle yMED subunits has been determined in the structure. By comparison, they also proposed a comprehensive map of the human MED. Cross-linking approach with endogenous complex have next refined the subunits organisation(Robinson et al., 2015).

Along with its role in gene activation, Mediator has soon been considered as part of a larger Pre-Initiation Complex (PIC)(Elmlund et al., 2006; Robinson et al., 2015). First structures containing both MED and Pol II start to emerge(Cai et al., 2009). Recently, high resolution structure of the complete PolII/GTFs/Mediator PIC have been released(Plaschka et al., 2015) and decipher precise MED architecture and dynamics in the context of transcription initiation(Tsai et al., 2017). Despite its large number of subunits and contrary to smaller protein complexes like TFIIH, MED

harbours only one enzymatic activity. The kinase activity of CDK8 was notably shown to control MED association with the rest of the PIC(Robinson et al., 2016).

Understanding the global role and functioning of MED is a real challenge that can only be solved by the combination of two approaches: the fundamental study of the protein complex and the analysis of the consequences of disease-causing mutations found in MED genes.

MEDIATOR COMPLEX IN TRANSCRIPTION INITIATION

Structural studies have enabled the emergence of a global view of the transcription initiation complex with the resolution of the various components such as Mediator, GTFs (TFIID, TFIIH) and Pol II (see below). The head and middle modules of MED constitute a compact structure that interacts with Pol II and GTFs, while the more apart tail module contacts distant transcription factors (TFs). Therefore, MED requires a constant composition and structure to integrate the basal transcription machinery, but it also might vary depending on the gene to adapt on specific DNA binding factors. The subunit MED14 plays the central role to nucleate the three core modules(Plaschka et al., 2015) and is thus essential for the large MED structural rearrangement(Tsai et al., 2017). The kinase module transitory binds to middle module and regulates transcription initiation through its enzymatic activity. All these events ultimately allow the MED to deliver outputs that range from repression to maximal activation of genes as well as modulation of the basal transcription.

RNA polymerase II

Early works showed that yMED associates closely with Pol II and some of the GTF in a build up complex termed Pol II holoenzyme. This complex have been described to assemble without DNA and to be directly load on the promoter(Kim et al., 1994; Koleske and Young, 1994; Thompson et al., 1993).

Consistent with these findings, MED subunit deletion or loss-of-function mutant have been found to compromise both Pol II and MED complex loading on the promoters of either induced or constitutively active genes(Ansari et al., 2009; Cantin

et al., 2003). We have also shown earlier this year that various mutations in MED12 subunit, which segregate with intellectual disability (ID), strongly affect the recruitment of Pol II and MED12 all over the genome(Donnio et al., 2017).

EM studies suggested that several subunits of Pol II, including Rpb1, Rpb2, Rpb3, Rpb6, Rpb11 and Rpb12, contact the middle or head modules of MED(Davis et al., 2002). An *in vivo* photo-crosslinking approach complemented by genetic analysis has identified a direct contact between Rpb3 and Med17, an interaction essential for Pol II genome-wide recruitment(Soutourina et al., 2011). Recent structural studies have revealed interaction between the head and Rbp4/7 or between the sub-module MED18/20 and Rbp3/11. The interaction is strengthened by contacts between the middle module subunits MED4/9 and Rpb1 and strongly depends on the hinge of the middle module. Indeed, mutation in MED21 does not alter MED integrity but strongly disturbs its interaction with Pol II(Nozawa et al., 2017; Plaschka et al., 2015; Sato et al., 2016; Tsai et al., 2017). Finally, the head module of MED is thought to bind the Pol II C-terminal domain (CTD) to facilitate its phosphorylation by CDK7(Kim et al., 1994; Robinson et al., 2016, 2012).

The head module of yMED stably associates with a Pol II-TFIIF complex, but not with the polymerase or TFIIF alone(Takagi et al., 2006). This difference suggests a combinatorial interaction but may also be the result of the conformational change induced by TFIIF(Forget et al., 1997). Moreover, a cryo-EM analysis showed that the presence of an activator (like VP-16) together with Mediator, is required to stably orient Pol II-TFIIF on the promoter, prior to transcription initiation(Bernecky and Taatjes, 2012; Bernecky et al., 2011). MED also overcome the inhibition of Pol II imposed by Gdown1, through TFIIF(Cheng et al., 2012; Hu et al., 2006; Jishage et al., 2012). Concordant with the MED-Pol II dissociation during transcription initiation, the head module of yMED was shown not to interact with Pol II-TFIIF when associated with DNA:RNA duplex(Takagi et al., 2006).

TFIIA, TFIIB and TBP/TFIID

Template commitment experiment first revealed a physical interaction of TFIID with MED(Koleske et al., 1992). Indeed, TFIID acts in collaboration with MED in the context of transcription activation(Guermah et al., 1998, 2001). Using immobilized

template assays and extract depleted or supplemented with purified factors, MED was revealed to coordinate TFIID binding to promoter(Black et al., 2006; Johnson et al., 2002). MED also facilitate TFIIB recruitment during the PIC assembly(Baek et al., 2006). Moreover, a purified TFIIA/TFIID/MED bound to promoter DNA was found to serve as a platform that supports active levels of PIC assembly and transcription(Johnson and Carey, 2003). The Head module of MED is structurally found to contact TFIIB through its ribbon part, and stabilize the initiation complex(Plaschka et al., 2015).

TFIIH and TFIIE

Different studies demonstrated that the tail module subunit MED15 binds stably to TFIIE and TFIIH(Sakurai and Fukasawa, 1997, 2000), whereas the head module MED11 subunit was found interact with the Rad3/XPD subunit of TFIIH(Esnault et al., 2008). Mediator has been shown to enhance *in vitro* phosphorylation of the Pol II CTD by TFIIH in a yeast reconstituted transcription system containing Pol II and basal factors, increasing its transcriptional activity(Kim et al., 1994). On the other hand, the phosphorylation of TFIIH subunit Cyclin H (CCNH) by MED subunit CDK8 seems to block CCNH/CDK7 kinase activity(Akoulitchev et al., 2000). This activity have been established to be a key event for promoter clearance and disruption of MED-Pol II interaction(Søgaard and Svejstrup, 2007). We also found few years ago a link between MED23 and TFIIH subunit CDK7. What is more, non-syndromic ID associated-MED23 mutation was shown to disturbed the recruitment of both TFIIH-CDK7 and MED-CDK8 on the promoter of activated genes(Hashimoto et al., 2011), enlightening the strong contact between TFIIH functions and the Mediator.

DNA binding factors

EM studies of human Mediator complexes revealed marked conformational changes of MED upon TFs binding. The structural comparison of complexes purified using either the SREBP-1a activator, the VP16 activation domain or a FLAG-tagged MED26 showed substantial differences in size and shape between them(Ebmeier and Taatjes, 2010). Different MED subunits were shown to interact with either the C-terminus domain or the activation domain of p53, leading to different consequences

on the overall PIC structure. Only MED associated with the p53 activation domain elicits the conformation able to activate a stalled Pol II into a productive Pol II(Meyer et al., 2010). Beside, CMT2B2 associated-MED25 mutation was shown to disturb its proline rich interaction domain, thus affecting the range of possible interacting SH3 domain proteins(Leal et al., 2009), with dramatic consequences.

Different Mediator subunits can work in synergy to regulate some genes. For example, the glucocorticoid receptor (GR) has been reported to use its ligand-dependent activation domain to target MED1, while its ligand-independent N-terminal activation domain targets MED14. By consequence, the expression of some GR-target genes requires MED14 but not MED1, while expression of other genes requires both(Chen et al., 2006). Similarly, MED23 is essential for expression of *Egr1* (Early Growth Response protein 1) gene in mES (mouse Embryonic Stems) cells, but is dispensable for its expression in MEF (Mouse Embryonic Fibroblast) cells(Balamotis et al., 2009).

Elongation and termination

MED seems to function as a “molecular bridge” that conveys essential information from transcription factors bound at upstream responsive elements to Pol II transcription machinery, suggesting a role in enhancer/promoter loop formation. Indeed, the role of MED in gene looping is established by the discovery of super-enhancers, which consist of clusters of enhancers. These domains are strongly bound by master transcription factors Oct4, Sox2 and Nanog, which recruit MED to control stem cell identity(Whyte et al., 2013). It is further reinforced by Mediator interaction with a cohesin complex containing NIPBL, Smc1a and Smc3, but not CTCF, to facilitate enhancer-promoter loops at stem cell associated-genes like Nanog, Oct4, Phc1 or Lefty1(Kagey et al., 2010). Aside, MED was also found to interact with certain ncRNA-a (non-coding RNA-activating) to regulate local gene expression through chromatin looping(Lai et al., 2013).

In addition to play a central role in PIC assembly, MED contributes to others steps of transcription. An *in vitro* study demonstrated that purified Mediator complex could stimulate transcription elongation by overcoming the DSIF-induced Pol II proximal pausing(Malik et al., 2007). Furthermore, MED26 interacts first with TFIID in

Pol II initiation complex and then exchanges TFIID for elongation complexes containing ELL/EAF and P-TEFβ, in order to facilitate transition of Pol II from a stalled state into an elongation state(Takahashi et al., 2011).

Mediator has also been reported to be important for transcription termination. MED18 was shown to be required for the recruitment of cleavage and polyadenylation factors while its absence leads to accumulation of Pol II near the 3' end of genes(Mukundan and Ansari, 2011). It was next demonstrated to facilitate the recruitment of cleavage factor 1 (CF1) complex at the 3′ end of genes by interacting with it(Mukundan and Ansari, 2013). These studies also highlight the role of MED in promoter/terminator loop formation and the link between such loops and transcription termination. Indeed, MED18 deletion strongly affects the loop formation and the termination process.

GENETIC DISORDERS ASSOCIATED WITH MEDIATOR

These last years, more and more studies have shown that mutations in MED subunits are associated with a wide range of human genetic disorders leading to congenital malformation and/or intellectual disability, enlightening the major role of Mediator in developmental processes (Table 1 and 2, Figure 2).

Understanding the defects that lead to such disorders at the molecular level might deepen our knowledge on Mediator as a major regulatory element of development, in addition to providing an explanation to the disorder itself.

Kinase module subunits

MED12-associated syndromes

Except *MED14*, *MED12* is the only Mediator subunit coding gene to be located on the sexual chromosome X. Considering the central role of MED14, it is not surprising that no mutations have been detected in this subunit when only one copy is available in male individuals. Such mutations would probably be too deleterious. *MED12* is located at Xq13.1 and codes for a 230kDa Mediator subunit part of the kinase module, along with MED13, CDK8 and CCNC or one of their paralogs. Three

independent but closely related syndromes, all presenting mild to severe ID, have been associated with *MED12* variations. With the raise of genetic diagnostic, *MED12* mutations have also been found to be responsible for other non-syndromic ID (Figure 3).

Opitz-Kaveggia syndrome (also known as FG syndrome; MIM #305450) was initially described in 1974 by Opitz and Kaveggia in a family of five males affected by intellectual disability (ID), macrocephaly, hypotonia and imperforate anus with X-linked inheritance(Opitz and Kaveggia, 1974). This syndrome has been linked to a recurrent missense mutation (c.2881C>T, p.R961W) in MED12 gene(Risheg et al., 2007). Hitherto, this mutation has been found in ten families, all sharing intellectual disability (ID), macrocephaly and hypotonia with variable digestive or genito-urinary anomalies (for a total of 23 affected males)(Clark et al., 2009). Another MED12 missense mutation (c.2873G>A, p.G958E) has been reported in a family with three cousins affected by Opitz-Kaveggia syndrome(Rump et al., 2011).

Few years later, Lujan and Fryns independently described another X-Linked Intellectual Disability (XLID) syndrome (commonly called Lujan-Fryns or Lujan syndrome, MIM #309520) also characterized by intellectual disability, macrocephaly, hypotonia and genito-urinary anomalies. It further includes dysgenesis of the corpus callosum, characteristic facial anomalies and behavioural disturbance(Fryns and Buttiens, 1987; Lujan et al., 1984). Later on, Schwartz team discovered a mutation (c.3020A>G, p.N1007S) in MED12 gene as responsible for Lujan syndrome(Schwartz et al., 2007).

Ohdo syndrome (MIM #300895) comprises a heterogeneous group of disorders characterized by intellectual disability and blepharophimosis (narrowing of the eye opening)(Ohdo et al., 1986). The Maat-Kievit-Brunner type (OSMKB or X-linked Ohdo syndrome) was initially distinguished from the other types of Ohdo syndrome due to its X-linked inheritance(Verloes et al., 2006). The facial characteristics also include prominent cheeks, nose with a rounded tip and a narrow mouth. As people with the condition get older, these characteristics become more pronounced and the face becomes more triangular (Table 1). Exome sequencing performed in two families with the OSMKB type revealed two different MED12 missense mutations segregating with the phenotype (c.3443G>A, p.R1148H, or

c.3493T>C p.S1165P). Subsequent analysis of a cohort of nine males with Ohdo syndrome, revealed another de novo missense mutation (c.5185C>A p.H1729N) in MED12(Isidor et al., 2014; Vulto-van Silfhout et al., 2013). Certain differences are noted among patients, even between patients bearing the same mutation (pR1148H).

Finally, Martinez team simultaneously published two studies presenting a *MED12* mutation (c.887G>A, p.R296Q) and enlightening the overlapping phenotypes of MED12-related patients(Caro-Llopis et al., 2016; Martínez et al., 2017). Indeed, the affected male was characterized as Ohdo syndrome. However, the authors noted symptoms not previously reported in other OSMKB patients, further expanding Ohdo syndrome and overlapping with other MED12-related symptoms.

MED12 Non-syndromic XLID

Along with these syndrome-associated mutations, numerous alterations of *MED12* have been identified, as the sequencing techniques were becoming more widely used, in non-syndromic X-linked Intellectual Disability (XLID) patients. Indeed, fourteen genetic alterations, mostly missense mutations, were characterised along the gene, in patients with intellectual disability (Table 1; Figure 3). Only few of them perfectly feat with one the three syndromes, but others are sharing most characteristic symptoms. In addition to ID, they often present delayed development with speech difficulties, Micro or Macrocephaly, various cranio-facial abnormalities (long narrow face, high forehead, altered philtrum), genito-urinary malformation and feeding or digestive troubles, in childhood and sometimes longer. In such conditions, some authors also defined a fourth and a fifth *Med12*-related condition(Narayanan and Phadke, 2017; Prontera et al., 2016).

For the three major syndromes, all reports only presented male patients along with mutations-carrying females. As *MED12* gene is encoded on X-chromosome, the authors often explained the absence of symptoms in female by the chromosomal compensation, also supported sometimes by skewed inactivation in favour of the wild allele. But in 2013, unlike the other *MED12*-missense mutations previously described, cognitive impairment was also noted in one heterozygous female with frameshift mutation. Indeed, sequencing of all X-chromosome exons in a large family with profound XLID, allows the identification of a frameshift mutation in *MED12*

(c.5898insC frameshift, p.S1967Qfsx84)(Lesca et al., 2013). The truncating mutation seems to have a more severe effect on MED12 function than previously described missense mutations. Nevertheless, two affected female-containing families have then been described with *MED12* missense mutations (c.1562G>A, p.R521H and c.2312T>C, p.I771T)(Fieremans et al., 2016; Prontera et al., 2016). Recently, three new missense mutations (c.617G>A, p.R206Q; c.2692A>G, p.N898D and c.3884G>A, p.R1295H) have been found in patients with XLID, bearing redundant symptoms with both definite syndromes and non-syndromic XLID. Among patients, women were also found to be partially affected by mutations (Figure 2)(Donnio et al., 2017).

MED12 amino-acid changes responsible for FG syndrome and Lujan syndrome are very close (within 50a.a.), thus it is not so surprising to observe such clinical signs overlay. But although both syndromes show a common X-linked inheritance and their main symptoms to overlap (Table 1), neither one of them was considered in the differential diagnosis of the other. The clinical consequences of the different genetic variations that affect *MED12* have first been described through the overlapping FG, Lujan and Ohdo syndromes, but more and more mutations are discovered in XLID patients. With regular intra-syndromes variability and observable resemblance between all described individuals, they could most likely be reconsidered as affected by a unique "*MED12* syndrome" with major common symptoms and some peculiar signs depending on the position of the mutation (Table 1). While the prevailing approach was to assign a unique variant as the genetic origin of a disease, it would be meaningful to consider the overall gene sequence to originate symptoms of seemingly related patients.

MED12 functions

Considering our current knowledge on the cellular functions of MED12, we can unveil potential origin for certain clinical signs of the patients. The amino acid sequence of MED12 reveals two different domains in its C-terminal part: the PQL domain, a domain rich in proline, glutamine and leucine, which is involved in proteins interaction and an OPA domain, a domain rich in glutamine. MED12 has been shown to interact via its PQL domain with Amyloid Precursor Protein Intracellular Domain (AICD)(Xu et al., 2011). AICD translocate into the nucleus and activate different

genes implicated in cellular processes relevant to Alzheimer disease. AICD was shown to recruit the Mediator complex through MED12 interaction on AICD-responsive promoters. This might be the first suggestion of the role of MED12 in neuron maintenance.

Sonic hedgehog (Shh) signalling pathway is strongly implicated in neuronal cell lineage and brain development(Yao et al., 2016), and is controlled by MED(Zhou et al., 2006). Indeed, MED12 interaction with Gli3 abolishes the repression of Shh targeted genes by the MED complex. In FG and Lujan patients, it was found that the MED12 mutations disrupt this MED repression thereby leading to altered Sonic hedgehog pathway(Zhou et al., 2012).

Mediator is involved in a protein network required for extra neuronal gene silencing. Indeed, MED12 has been demonstrated to link REST (RE1-silencing transcription factor) with G9a in order to silence REST-target genes, in non-neuronal cells. G9a is a histone methyltransferase, which catalyses histone H3K9 repressive mono- and di-methylation(Ding et al., 2008). Several MED12 mutations have been shown to compromise its ability to mediate REST-direct recruitment of G9a and to disturb the expression of neuronal genes(Ding et al., 2008; Vulto-van Silfhout et al., 2013). The same mutations, plus some others, have also been found to disturb the expression of Immediate Early Genes (EIGs). These genes are important for brain development and neuronal plasticity. They also code for transcription factors that control the expression of REST. MED12 mutations were found to disturb the expression of REST but also MMP-3, implicated in neuronal and synaptic plasticity and REST-regulated SYN1, implicated in axonogenesis and synaptogenesis(Donnio et al., 2017).

The Shh and REST signalling pathways are important for neuronal cells as well as brain organisation and their abnormal regulation undoubtedly participate to brain malformation and malfunction in affected patients. The etiological basis of Med12 associated disorders, while not fully resolved, is also suggested by studies that implicate MED12 in critical aspects of development, e.g. genitourinary malformation(Moghal, 2003). Furthermore, Med12-deficient zebra fish embryos show defects in the development of brain, neural crest and ear, among other organs(Hong et al., 2005; Rau et al., 2006; Wang et al., 2006). In these models, Med12 has been

shown to play an important role in the production of mono-aminergic neurons and cranial sensory ganglia through selective regulation of neuronal gene expression(Wang et al., 2006). This may also produce some explanations to the neurological features observed in the MED12-related patients. Finally, another work has identified a role for Med12 during endoderm development which may originate the craniofacial characteristics and the digestive defects observed in Med12-patients(Shin et al., 2008).

Considering the roles of MED12 in brain development, it is not surprising to found some polymorphisms to be associated with psychiatric diseases. Indeed, an increased risk of schizophrenia in people with northern European ancestry has been associated with a particular polymorphism in the MED12 gene, known as the HOPA (12bp) polymorphism. This variation is an insertion of four additional amino acid residues (QQHQ) in the OPA domain of MED12(DeLisi et al., 2000; Philibert et al., 2007). A second rare deletion polymorphism within the MED12 OPA domain (HOPA−15 bp) appears to be related with psychosis(Beyer et al., 2002).

Med13L syndrome

Searching for genes associated with congenital heart defects, a *MED13L* gene interruption was reported to be responsible for transposition of the great arteries (TGA; MIM #608808) in patient presenting both ID and TGA(Muncke et al., 2003). In 2011, a new study has found a *MED13L* missense mutation to be responsible for mild ID in a consanguineous family with no reported heart defects(Najmabadi et al., 2011). A *MED13L* splicing abnormality and a 2 bp deletion was also reported to be associated with autism disorder(Codina-Solà et al., 2015; Iossifov et al., 2012).

In 2013 first appears the concept of MED13L syndrome with gene dosage changes. Several patients with chromosomal deletion or duplication combine both ID and congenital heart defects, with specific morphological features(Asadollahi et al., 2013). The next years, several studies strengthened the association between *MED13L* issues (either a disruption or genetic mutation) and intellectual disability(Gilissen et al., 2014; Hamdan et al., 2014; Redin et al., 2014; Utami et al., 2014), with no regards for eventual heart defect. Later, two patients were reported with ID and other characteristic features but only one of them was affected by heart

problem, a persistent foramen ovale(van Haelst et al., 2015), showing that MED13L is not necessarily associated with cardiac issues.

In 2015, Adegbola *et al.* published a large patient study with either mutational or chromosomal issues of *MED13L* gene that allow to define a new syndrome called Mental Retardation And Distinctive Facial Features With Or Without Cardiac Defects (MRFACD; MIM #616789). The syndrome combines mild to severe ID with specific facial abnormalities such as bulbous nose, irregular nasal bridge and up slanting palpebral fissure, associated with hypotonia and speech delay in most of the cases. Some patients also bear congenital heart defects. It can be TGA, tetralogy of Fallot, ventricular septal defect or persistent foramen ovale(Adegbola et al., 2015; Asadollahi et al., 2017; Cafiero et al., 2015; Caro-Llopis et al., 2016; Mullegama et al., 2017; Yamamoto et al., 2017). All MED13L patients are summurized in Table 1 of Asadollahi et al., 2017.

A seven years old child with moderate intellectual disability and craniofacial abnormalities has been found to carry a 800-Kb deletion of chromosome 17, notably encompassing the *MED13* gene(Boutry-Kryza et al., 2012). As mentioned earlier, MED13L is mutually exclusive with its paralog MED13(Daniels et al., 2013). Further research will be required to explain their respective roles in development. Nonetheless, MED13/MED13L are implicated in the interaction of the Kinase module with the core MED(Davis et al., 2013). Considering the role of MED in development and the specific role of the Kinase module in transcription regulation, one can speculate that any mutation or abnormal protein level of MED13/MED13L would alter the highly regulated programme of transcription controlling organismal development.

Although symptoms are not precisely overlapping those related to MED12, similarities may be considered with MED13/MED13L patients. Indeed, they present mild to severe ID, often associated with hypotonia and specific facial features like slanting palpebral fissure or bulbous nose that are found altogether in MED12 related syndromes.

CDK19 and mild mental retardation

The female probant with a unique combination of symptoms is affected by bilateral congenital retinal folds, nystagmus, microcephaly and mild intellectual

disability. FISH analysis performed on patient lymphocytes revealed chromosomal inversion between 6q12.1 and 6q21 on one chromosome. The breakpoint 6q21 disrupts the *CDK19* gene. qPCR revealed a 50% reduction of the transcript(Mukhopadhyay et al., 2010). If CDK19 is yet not well characterised, its paralog CDK8(Daniels et al., 2013) is already know to be implicated in dendritic branching(Kaufmann and Moser, 2000) and neuronal wiring of the visual system as well as neuronal dendritic proliferation(Berger et al., 2008; Loncle et al., 2007).

Head module subunits

Med17 and infantile cerebral and cerebellar atrophy

Few years ago, a specific form of microcephaly within the Caucasus Jewish community has been associated with mutation in *MED17*(Kaufmann et al., 2010). This association was discovered through the study of five infants from four unrelated families who manifested post-natal progressive microcephaly, spasticity, epilepsy and severe developmental retardation (MIM #613668). Brain scans revealed cerebral and cerebellar atrophy with severe myelination defect, small thalami and a thin brainstem. A homozygous missense mutation in MED17 (c.1112T>C, p.L371P) was found to segregate with the disease state.

In 2016, another family of two siblings have been described(Hirabayashi et al., 2016). Although both were born at term without complication, the boy encountered post-natal microcephaly while the girl suffered from cerebellar atrophy associated with myelination delay. They also quickly develop nystagmus and sudden opisthotonic posturing and subsequently became hypotonic with choreiform movement. Whole exome sequencing revealed a combination of two heterozygous mutations in *MED17* (c.1013-5A>G, p.S338Nfs*15 and c.1484T>G, p.L495T), inherited from their heterozygous father and heterozygous mother respectively.

MED17 belong to the head module and has a central role in Mediator architecture and function. It is critical for head module assembly(Imasaki et al., 2011; Takagi et al., 2006) and its contact with the rest of Mediator, by forming the largest contact between the Head and Middle modules(Tsai et al., 2014). Moreover, as described in introduction, the interaction of Med17 with Pol II subunit Rpb3 is essential for genome-wide Pol II recruitment *in vivo*(Soutourina et al., 2011). This

may explain transcription impairment and lethality of Med17 inactivation in yeast and Drosophila(Boube et al., 2000; Linder et al., 2006; Thompson and Young, 1995). Additionally, MED17 is an established physical target of the transcription factors p53 and NF-κB(Ito et al., 1999; van Essen et al., 2009) and also of the DNA repair proteins Rad2/XPG(Eyboulet et al., 2013).

Further studies will be required to understand how Med17 mutations might impact these critical functions and lead to infantile cerebral and cerebellar atrophy. L371P missense mutation has a stronger phenotype while the two others mutations require to be associated, as heterozygous parents behave normally. Considering the apparent normal prenatal development of patients and the decisive postnatal issues, the authors proposed that MED17 mutations disrupt a critical function of MED17 in oligodendrocyte development, a process beginning only after birth in human.

Med20 and infantile-onset neurodegenerative movement disorder

Two sisters have been described for infantile movement disorder associated syndrome(Vodopiutz et al., 2015). When they were born from uneventful pregnancies, they rapidly exhibited delayed psychomotor development with spasticity. They subsequently started to loose acquired skills and their head circumference decreased to a characterised cerebral and cerebellar atrophy. By whole exome sequencing, the two siblings were found to carry a mutation in head subunit MED20 (c.341G>C, p.G114A) that segregate with the disease.

While the specific roles of MED20 subunit are not well described, it appeared to form a movable jaw during transcription initiation in association with MED18(Larivière et al., 2012). Importantly, it can be noted that MED20 patients are shearing similar symptoms with MED17 patients, while both subunits belong to the head module.

LGS and CDL associated with a large chromosomal deletion including Med30

Comparative genomic hybridization array (aCGH) on a malformed foetus revealed a large 2.88-Mb deletion on chromosome 8, notably encompassing *Med30*. Malformations, especially the cranio-facial features, were recognisable as Langer-Giedion syndrome and Cornelia de Lange syndrome (LGS/CDL; MIM #150230 and

#614701) both associated with intellectual disability and microcephaly(Chen et al., 2015).

Tail module subunits

Med15 and DiGeorge syndrome

The DiGeorge syndrome (DGS)/velocardiofacial syndrome (VCFS) (MIM #188400) also know as 22q11.2 deletion syndrome is associated with variable symptoms which often includes congenital heart defect, developmental delay and characteristic facial features, notably palpebral abnormalities. A significant number of patients are affected by intellectual disability. Among several gene, the deleted region encompassed *MED15*. A study in mice revealed its presence and possible roles in the development of structures affected in DGS/VCFS(Berti et al., 2001). It have also been proposed to be associated with schizophrenia(De Luca et al., 2003).

Med23 and non-syndromic intellectual disability

A recent work in our lab has uncovered a direct link between a non syndromic intellectual disability and *MED23*, a tail subunit(Hashimoto et al., 2011). This link was established through genetic analysis of an Algerian family where five of eight children, born to healthy consanguineous parents, exhibited inability to read or to write but no malformations and normal metabolic screening. The variation (c.1850 G>A, p.R617Q) in *MED23* gene was revealed to segregate with the disease and was not found in control chromosomes (MIM #614249). In 2015, a pair of brother was subsequently discovered to bear heterozygous *MED23* mutations (c.3656A>G, p.H1219R and c.4006C>T, p.R1336X) in a non-consanguineous family(Trehan et al., 2015). The siblings are affected by profound ID with globally delayed development, brain anomalies (EEG and myelination), spasticity and congenital heart disease.

Med23 was originally identified as a genetic suppressor of hyperactive ras phenotype in *C. Elegans*(Singh and Han, 1995) and to mediates the response of IEG gene to serum mitogens, notably *EGR1*(Balamotis et al., 2009; Stevens, 2002; Wang et al., 2005, 2013). In line with such a function, our group found that the mutation in MED23 alters the interaction between enhancer-bound transcription factors and MED, leading to altered expression of mitogen-responsive IEG. These genes,

important for brain development and functioning, have been show to be deregulated by the two other mutations.

In our hands, Mediator was shown to interact with TCF4(Hashimoto et al., 2011). *De novo* mutations (deletions, frameshift, nonsense, splice site or missense mutations) of TFC4 coding gene caused Pitt-Hopkins syndrome (PTHS, MIM # 610954), an autosomal dominant disorder characterized by severe intellectual disability, distinctive facial features and breathing anomalies. In addition, half of PTHS patients develop a postnatal microcephaly(Peippo and Ignatius, 2011), like patients affected by *MED23* mutations. TCF4 is strongly implicated in the regulation of IEGs transcription, which are dysregulated in different Mediator related neurological disorder(Donnio et al., 2017; Hashimoto et al., 2011).

Med25 and CMT2B2

Charcot-Marie-Tooth (CMT) disease, also known as hereditary motor and sensory neuropathy, encompasses a group of clinically and genetically related disorders, affecting the peripheral nervous system. This disease, one of the most common inherited neurological disorders, is characterized by muscle wasting, weakness and sensory loss across various parts of the body (Reilly et al., 2011). Among all forms of CMT, the axonal form is very rare (CMT2B2; MIM #605589) and at this time three causative genes have been identified: Lamin, GDAP1 and MED25. The association of the MED25 subunit with ARCMT2 was established through investigation of an extended Costa Rican family with Spanish and Amerindian ancestor (Leal et al., 2009). Affected members presented chronic symmetric sensory-motor neuropathy and primary axonal degenerative process with mild myelin impairment. A homozygous missense mutation (c.1004C>T, p.A335V) in MED25 gene was identified as the cause of the disease(Leal et al., 2009).

Another MED25 genetic disruption has been associated with ID. Seven adults from a large consanguineous family in Northeastern Brazil were found to carry a mutation in MED25 that segregate the disease (c.418C>T, p.R140W). They present moderate to severe ID accompanied with facial characteristics like tall high foreheads or prognatism. They are totally dependent for daily life tasks and only able to speak few words. Behavioural problems are present for some of tem(Figueiredo et al.,

2015). Another study presents seven patients from four unrelated families affected by eye-intellectual disability syndrome. They were found to carry MED25 mutation that segregates with the disease (c.116A>G, p.T39C). They share various phenotype such as microcephaly, craniofacial malformations, cardiovascular defects and different eye issues like hypertelorism, strabismus or cataract (MIM #616449)(Basel-Vanagaite et al., 2015).

MED25 contains several domains that allow it to interact with multiple proteins, such as the histone acetyltransferase CBP and RAR (retinoic acid receptor) in a ligand-dependent manner. These interactions are important for the recruitment of MED complex to retinoic acid (RA)-responsive genes(Lee et al., 2007), that are involved in the maintenance of adult neurons(Maden, 2007). If mutations alter the structure and interactome of MED25, proper regulation could be compromised and partially explain the axonal degeneration symptom observed in CMT disease and other neurological issues.

CONCLUSION

Although recent studies have greatly expanded our knowledge on the Mediator complex, we are only beginning to understand the diversity of its role in transcription. Further characterization of Mediator will be necessary to improve our comprehension of the complex mechanisms that regulate the expression of protein coding genes. Beside to biochemical characterisation, clinical data provided us with capital insights into Mediator functioning and its roles in development. Considering the number of neurological and neurodevelopmental disorder related to its subunits, we could consider the Mediator as a major player of the brain development. Advancing our comprehension on transcription mechanism will also help us to better understand the aetiology of patients bearing Mediator mutations.

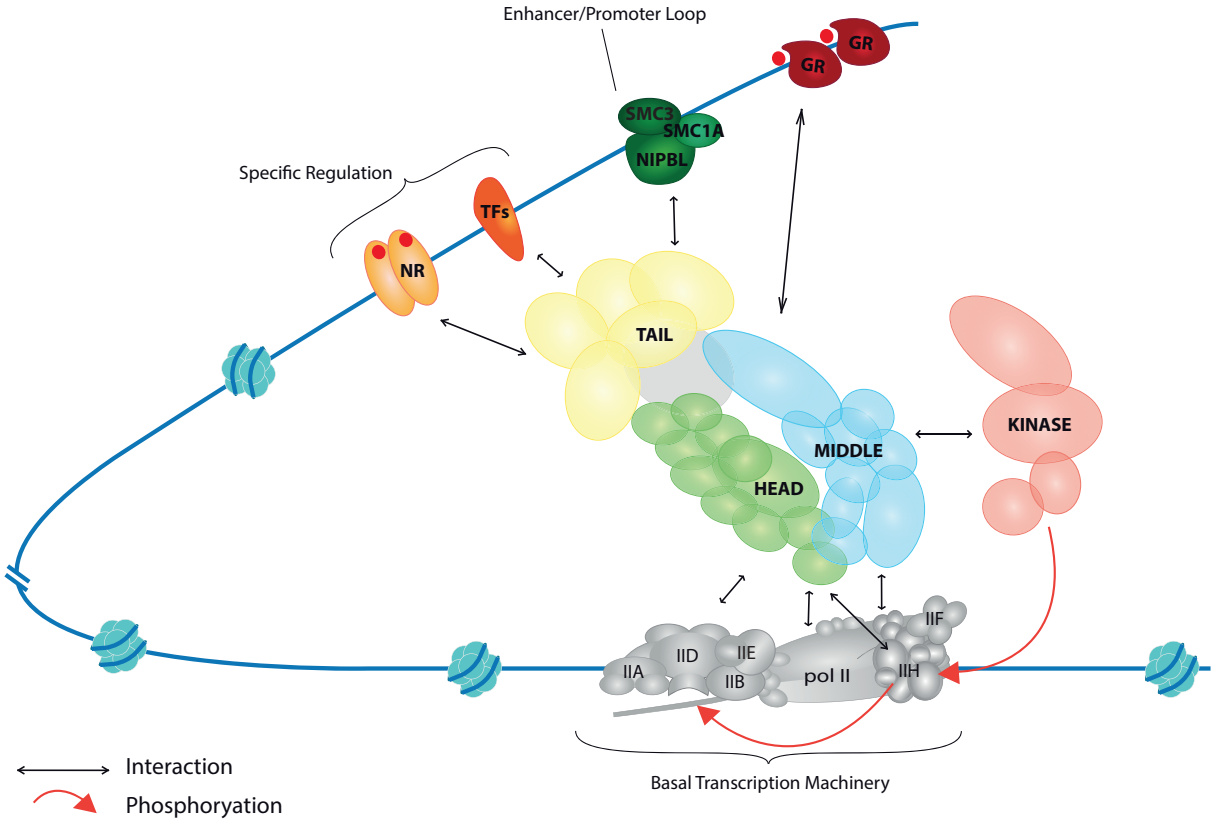


Figure 1: Schematic representation of the Mediator complex in the context of transcription initiation

A Schematic representation of the Mediator complex in the context of transcription initiation showing the interactions between the different MED modules and the basal transcription machinery (RNA polymerase II/general transcription factors) as well as specific transcription factors/nuclear receptors or the NIPL complex. MED modules and subunits are specified on Figure 2.

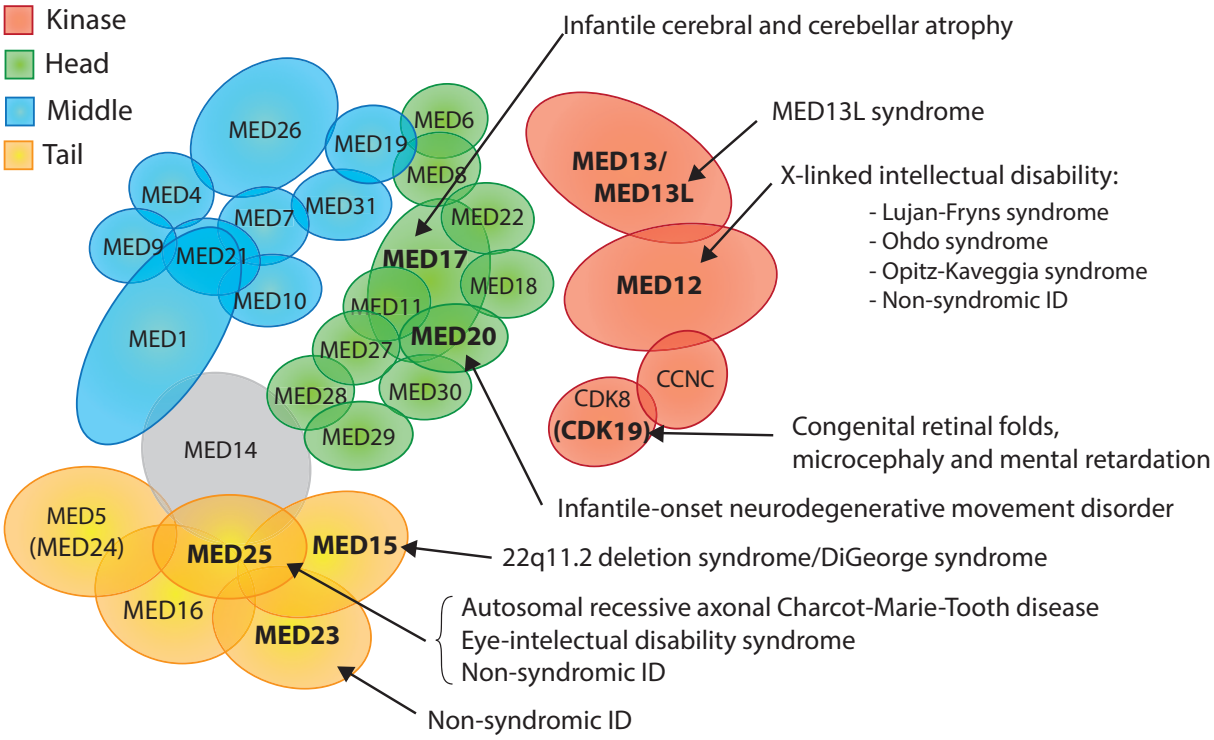
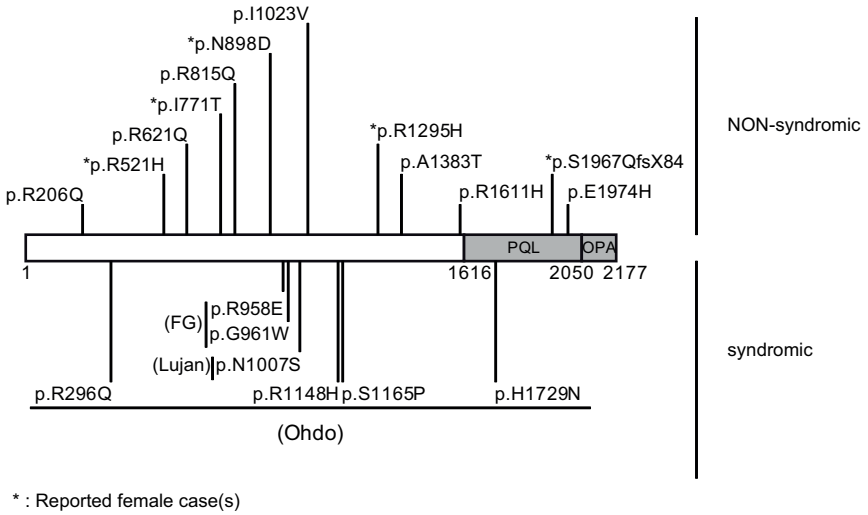


Figure 2: Schematic representation of the Mediator complex and associated genetic diseases

A schematic representation of the Mediator showing the different modules as well as their subunit composition. The subunit MED14, depicted in grey, plays the central role to nucleate the three core modules. Subunits that can be mutated are highlight in bold and the related diseases are specified.



* : Reported female case(s)

Figure 3: Schematic overview of the MED12 protein and the reported mutations associated with intellectual disability

The amino-acid sequence of MED12 is depicted with its different domains. All reported MED12 variants associated with a specific syndrome are reported on the lower part. All other MED12 variants, that are not associated with a specific syndrome, are reported on the upper part.

Table 1: MED12 mutations and associated symptoms

	R206Q	R296Q	R521H	R621Q	I771T	R815Q	N898D	G958E R961W	N1007S	I1023V	R1148H S1165P H1729N	R1295H	A1383T	R1611H	S1967QfsX84	E1974H	Xq12_q13
Associated-syndrome		Ohdo						FG	Lujan		Ohdo						
OMIM		# 300895							#309520		# 300895						
Refs	(Donnio et al., 2017)	(Martínez et al., 2017; Patil et al., 2017)	(Fieremans et al., 2016)	(Prescott et al., 2016)	(Prontera et al., 2016)	(Tzschach et al., 2015)	(Donnio et al., 2017)	(Lyons et al., 2009; Risheg et al., 2007; Rump et al., 2011)	(Schwartz et al., 2007)	(Yamamoto and Shimojima, 2015)	(Isidor et al., 2014; Vulto-van Silfhout et al., 2013)	(Callier et al., 2013; Donnio et al., 2017)	(Langley et al., 2015)	(Narayanan and Phadke, 2017)	(Lesca et al., 2013)	(Bouazzi et al., 2015)	(Kaya et al., 2012)
Growth																	
Stature	Tall	Short							Tall			Tall (1/3)					
Cephalic characteristic	Macro	Brachy and micro			Micro		Macro	Macro	Macro			Macro					Micro
Hearing loss				+													
Neurological																	
Developmental delay	+	+	+	+			+	+	+	+	+	+	+	+			+
Intellectual disability	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Agenesis of corpus callosum								+	+			1/3					+
Hypotonia		hypertonia		+			+	+	+		+	+	+				+
Behavioural disturbance	+		+		+		+	+	+	+	+	1.3			+	+	+
Speech abnormality	+		+	+	+				+	+	+	+	+	+		+	+
Craniofacial																	
Long narrow face					+				+	+		+			+		
Tall prominent forehead				+	+		+	+	+	+	+	+			+	+	
Triangular face				+							+	1/3	+				+
Blepharophimosis		+		+							+		+	+			
Down slanting palpebrae	+			+			+	+	+		+	+	+	+			
Nasal bridge		Split tip and depressed root			High nasal root								Narrow and prominent	Broad	High	High nasal root	Flattened
Bulbous nose		+			+								Pointy				
Eyebrows		Arched	Sparse lateral							Arched			Thin	Arched/ sparse			
Eyes problem	Strabismus	Strabismus	Deep set eyes	Gaze paresis, exotropia, astigmatism and hypermetropia			Astigmatism	Strabismus	Strabismus		Strabismus	Strabismus (1/3)	Strabismus				Epicanthic fold
Hypertelorism		+					+	+				1/3	+				+/-
Ears		Abnormal	Rotated/Small	Low set/Small	Large		+	+	Abnormal	Large	+	1/3	Rotated		Large		
Philtrum		Long/flat			Short		Short		Short		Long	Long	Flat		Short		
Maxillary hypoplasia				+	+		+	+	+			2/3	+			+	
Micrognathia		+ (Retro)		+				+	+		+				Pro/retro	+	
High narrow palate				+			+	+	+		+	+					
Mouth				Small	Large		Open	Open	Open		Open		Small		Open (+/-)	Open	
Dental anomalies			+	+			+	+	+	+		+					

Extremity																	
Foetal finger pads							+			+							
Digits		Absent digital triradius	Clinodactyly	Clinodactyly				Syndactyly									Syndactyly
Thumbs/toes		Overlapping 4th	Short	Short/proximal			Broad	Broad	Broad								
Horizontal palmar crease								+									
Long hyper-extensible digits				+			+		+		+	1/3					
				Forefoot adduction													
Cardiovascular																	
Congenital heart defect	LVH		ASD				Spontaneous closure ASD										
Gastro-intestinal																	
Constipation				+			+	+		+	+	2/3	GER				
Anal anomalies							+	+				1/3					
Genito-urinary																	
Genital anomalies	+	+					+	+	+		+	+	+				+
Others	Extra-nipples		Short upper limbs	Cervical block (C3-C4), scoliosis		No data						Extra-nipples/Thoracic kyphosis	Gastronomy tube				

No clinical data was available for the patient bearing the mutation p.R815Q (Tzschach et al., 2015)

LVH: Left Ventricular Hypertrophy; ASD: Atrial Septal Defect; GER: GastroEsophageal Reflux;

Table 2: MED subunits mutations and associated symptoms

	MED15	MED17	MED17	MED20	MED23	MED23	MED25	MED25	MED25	CDK19	MED13	MED30
Mutations	22q11.2 deletion	L371P	Heterozygote S338Nnfs*15 /L495W	G114A	R617Q	Heterozygote H1219R/R1336X	T39C	R140W	A335V	Inversion	800Kb Deletion	2.88Mb Deletion
Refs	(De Luca et al., 2003)	(Kaufmann et al., 2010)	(Hirabayashi et al., 2016)	(Vodopiutz et al., 2015)	(Hashimoto et al., 2011)	(Trehan et al., 2015),	(Basel-Vanagaite et al., 2015)	(Figueiredo et al., 2015)	(Leal et al., 2009)	(Mukhopadhyay et al., 2010)	(Boutry-Kryza et al., 2012)	(Chen et al., 2015)
Associated syndrome	DiGeorge syndrome	Cerebral and cerebellar atrophy	/	Infantile-onset neurodegenerative movement disorder	/	/	Eye-intellectual disability syndrome	/	CMT22B2	/	/	Langer-Giedeon syndrome and Cornelia de Lange
OMIM	#188400	#613668	/	/	# 614249	/	/	/	#605589	/	/	#150230 and #614701
Growth												
Cephalic characteristic		Micro		Micro			Micro			Micro		
Neurological												
Intellectual disability	+	+	+	+	+	+	+	+		+	+	LGS/CDL
Developmental delay		+	+	+				+		+	+	
Hypotonia		Spasticity	+	Spasticity	N.A	Axial hypotonia+spasticity						
Behavioural disturbance								+				
Speech abnormality			+	+	N.A	+		+			+	
Craniofacial												
Long narrow face								+				
Down slanting palpebrae	Fissure											
Abnormal nasal bridge												+
Bulbous nose	+							+			Short	+
Abnormal eyes		+		+			+			Retinal fold		
Hypertelorism	Telecanthus						+/-					
Abnormal ears	+						+/-				+	+
Philtrum	Short											+
Micrognathia	+							Prognathia				+
Mouth	Small										Small	
Extremity												
Digits							+/-		+			
Thumbs/toes							+/-		+			
Foot									+		+	
Cardiovascular												
Congenital heart defect	+					+	+					
Gastro-intestinal												
Constipation							+/-					
Genito-urinary												
Genital anomalies							+					
Others	Thymus and parathyroid hypoplasia	Epilepsy					Nevus flammeus simplex on the forehead, receding frontal hairline		Muscle wasting, weakness and sensory loss	Congenital retinal fold syndrome		

Results

I - The nucleotide excision repair factor XPC

Manuscript in preparation

Cooperation between E2F1 and XPC to regulate H3 histone acetylation by GCN5 acetylase

BIDON Baptiste∞, ILTIS Izarn∞, SEMER Marissa, NAGY Zita, LARNICOL Annabel, COIN Fréderic, EGLY Jean-Marc* & LE MAY Nicolas*

Département de génomique fonctionnelle et Cancer, CNRS, INSERM, Université de Strasbourg, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 1 rue Laurent Fries, 67404 Illkirch-Graffenstaden, France.

* To whom correspondences should be address:

nlemay@igbmc.fr; egly@igbmc.fr

∞ Should be regarded as joint First Authors

ABSTRACT

Nucleotide excision repair (NER) factors including XPC, XPA, XPG and XPF/ERCC1 were found to be recruited at the promoter of activated genes in absence of genotoxic attacks(Le May et al., 2010a), to contribute to several chromatin-remodelling events such as histone post-translational modifications (PTMs), DNA breaks and DNA demethylation. Here, we focused on XPC, known to initiate GG-NER, and dissect its functions in transcription. Genome-wide analysis revealed that XPC is mainly recruited at promoters, colocalising with Pol II. We first show that XPC is specifically involved in histone modifications and plays a role in the regulation of a certain set of genes. Interestingly, we showed that XPC mediates together with E2F1, a transcription factor, the recruitment of GCN5 histone acetylase within the ATAC complex, to favour the acetylation of H3K9. *In vitro* acetylation assays further showed that TFIIH, a transcription/DNA repair factor target the XPC/E2F1/GNC5 complex to enhance H3 acetylation.

Key words: XPC, Histone post-translational modifications, Gene expression, GG-NER, GCN5, E2F1, TFIIH, H3K9ac, Xeroderma Pigmentosum.

INTRODUCTION

Transcription is a highly orchestrated event that requires several hundreds of proteins, including the basal transcription machinery, the Mediator, the co-activators and the chromatin remodelers. Transcription initiation also requires the recruitment of certain DNA repair proteins. The link between DNA repair and transcription was first established with the discovery of the multi-protein complex TFIIH, which is both a basal transcription and a nucleotide-excision repair (NER) factor(Feaver et al., 1993; Schaeffer et al., 1993). This connection was then strengthened by the discovery that other NER factors (CSB, XPC, XPA, XPG, XPF/ERCC1) also participate to transcription(Barreto et al., 2007; Le May et al., 2010a; Schmitz et al., 2009). These factors were first characterized as part of NER, a DNA repair pathway that can eliminate a wide variety of DNA lesions, originated by endogenous or exogenous genotoxic attacks like UV irradiation(Friedberg et al., 2006). While Cockayne syndrome group A and B (CSA and CSB) proteins target RNA polymerase II (Pol II) stalled in front of a DNA lesions to engage transcription coupled-NER (TC-NER), *Xeroderma pigmentosum* group C (XPC) protein, in complex with HR23B, recognizes a wide variety of DNA damages located all over the genome, to initiate the global genome-NER (GG-NER)(Sugasawa, 2010). XPC was also shown to regulate the DDB2 complex, which mediate chromatin decondensation through ubiquitination around DNA damages(Luijsterburg et al., 2012; Takedachi et al., 2010).

Our group has shown earlier that NER factors are associated with the RNA polymerase II (Pol II) transcription machinery and are sequentially recruited (XPC/CSB followed by RPA/XPA and XPG/XPF) at the promoter of retinoic acid-activated RAR*ß2*(Le May et al., 2010a). In such conditions, the recruitment of the NER factors is necessary to achieve optimal chromatin remodelling, including histone post-translational modifications (PTMs) as well as active DNA demethylation, DNA breaks induction and gene looping(Barreto et al., 2007; Le May et al., 2010a, 2012; Schmitz et al., 2009). More recently, a DNA repair complex containing XPC has been characterised to function as a coactivator of Oct4/Sox2 in ES and iPS cells(Cattoglio et al., 2015; Fong et al., 2011).

Understanding the functions of NER factors is particularly important since mutations in the genes give rise to the human autosomal recessive disorder

Xeroderma pigmentosum (XP). Biochemical defect in XP-C represents the most frequent NER defective group. Clinically, XPC is characterized by an extreme photosensitivity and a high susceptibility to develop tumours (melanoma, squamous cell carcinoma) on sunlight-exposed areas of the skin(Cleaver, 2005), accompanied by an increased susceptibility for lung, breast and colorectal cancers. However, these XPC individuals rarely exhibit neurological disorders or developmental defects.

In the present study, we have investigated the roles of XPC in chromatin remodelling upon transcription. Our genome-wide scale data have revealed that XPC is mainly recruited with Pol II at promoter regions. A specific set of gene is down regulated when XPC is absent of their promoter. The absence of XPC coincided with deregulated enrichment of Pol II and altered euchromatin marks, including H3K9ac, at the promoter. The absence of H3K9ac on such promoters led us to identify interaction between XPC and the Histone Acetyl-Transferase (HAT) General Control of Nutrition 5/K Acetyl-Transferase 2A (GCN5/KAT2A), as part of the ATAC complex. An E2F1 signature was found to characterize most of the genes regulated by XPC/GCN5. Finally, we describe the recruitment of a complex including XPC/HR23B, E2F1 and GCN5 on promoter of activated gene. This complex acts cooperatively with the general transcription TFIIH to increase the GCN5 HAT activity. Altogether, the results presented here provide new insights into the transcriptional role of XPC and the molecular aetiology of XP-C.

RESULTS

XPC deficiency disturbs gene expression

To investigate the involvement of XPC in transcription, we analysed *RARß2* transactivation in HeLa cells constitutively expressing shRNA directed against either XPC (shXPC), XPA (shXPA) or scrambled (shCtrl), after all-trans retinoic acid (ATRA) treatment. We also used fibroblasts derived from two XP-C patients presenting severe and mild clinical features and bearing p.R579st (XPC/R579st) and p.P334H (XPC/P334H) mutations respectively as well as XPC rescued fibroblasts (XPC/WT)(Bernardes de Jesus et al., 2008; Gozukara et al., 2001; Li et al., 1993). At

6h post ATRA treatment, we observed a correlation between *RARß2* expression (Figure 1A1 and 1B1) and the recruitment of the transcriptional machinery (including Pol II, RAR and TFIIH kinase CDK7) but also the NER factors (XPC, XPA, XPG) at *RARß2* promoter *in* shCtrl, XPC/WT as well as in XPC/P334H cells (Figure 1A2 and A5, B2 and B5, B4 and B7). However, such correlation was abolished in shXPC and XPC/R579st cells, (Figure 1A3 and A6, B3 and B6) both characterized by the absence of XPC (Figure S1A). Surprisingly, in shXPA cells, we observed a reduced *RARß2* mRNA induction without deregulation of the recruitment of XPC or the transcriptional machinery (Figure 1A4 and A7).

We next sought to identify the correlation between XPC and Pol II at the genomic scale. Chromatin-immunoprecipitation (ChIP) followed by high-throughput DNA sequencing (ChIP-seq) yielded 2,191 and 16,440 binding events for XPC and Pol II respectively, 6 hours post ATRA treatment in XPC/WT cells (Figure 1C). Interestingly, we identified 1,797 XPC/Pol II common binding events, representing 82% of XPC peaks and covering 1529 genes. According to HOMER annotation, one third (32%) of the XPC binding events appeared at promoter region (Figure 1D). By combination of promoter, 5'-UTR and exon annotated peaks (mostly located in the first one), the proportion of XPC peaks located close to the transcription start site (TSS) rises to 44%. Interestingly, 94% of the XPC-bound promoters also showed a Pol II enrichment (Figure 1D).

Collectively, these findings established an important enrichment of XPC around promoters also occupied by Pol II in XPC proficient cells, for up to 1500 genes.

XPC specifically controls H3K9 acetylation by recruiting GCN5

Following gene activation, the recruitment of the basal transcription machinery and NER factors is accompanied by chromatin remodelling events like histones post-translational modifications (PTMs), DNA demethylation, DNA breaks and gene looping(Barreto et al., 2007; Fong et al., 2013; Le May et al., 2010a; Schmitz et al., 2009). We observed that XPC recruitment at the *RARß2* promoter was crucial for preparing accurate transcription: (i) BioChip assay which measures the incorporation of biotinylated dUTP within broken DNA, evidenced DNA cleavages in XPC/WT and XPC/P334H but not in shXPC, and XPC/R579st (Figure S1B). (ii) Unmethylated DNA

immunoprecipitation (UnMeDIP) showed that in XPC deficient cells, the *RARß2* promoter region remained methylated contrary to what occured in wild type cells (Figure S1C). (iii) Similarly, the q3C assay did not show any stable interaction (loop formation) between the promoter and the terminator of the *RARß2* in shXPC and XPC/R579st cells (Figure S1D).

We next showed that upon ATRA treatment, an increase of H3K9ac and H3K4me3 active histone marks was detected around *RARß2* promoter in shCtrl, XPC/WT and XPC/P334H cells, whereas this enrichment was not observed in shXPC and XPC/R579st cells (Figure 2A). Note that the silencing of XPC and XPA do not disturbed the steady state level of these histone PTMs (Figure S1E). Interestingly, in shXPA cells, where the normal recruitment of XPC was detected (Figure 1A), the H3K9ac as well as H3K4me3 marks were not disturbed (Figure 2A, left panel), while the DNA breaks, DNA demethylation and DNA loop were indeed absent(Le May et al., 2010a, 2012). These results suggested that chromatin remodelling events were specifically related to the presence of XPC, acting independently of the other NER factors. ChIP-seq analysis next showed that among the 1,797 XPC/Pol II binding events detected in XPC/WT cells, 98% (1756) were enriched for the H3K9ac mark and 78% (1385) with both H3K9ac and H3K4me3 marks (Figure 2B). Histones modifications are regulated by various enzymes such as lysine methyltransferases (KMT), lysine-demethylases (KDM), histone acetyltransferases (HAT) as well as histone deacetylases (HDAC), also found around the promoter of activated genes. Given the role of GCN5 in H3K9 acetylation during NER(Guo et al., 2011), we investigated whether the variation of H3K9ac at the activated *RARß2* promoter could be mediated by XPC. We detected the presence of GCN5 concomitantly to the increased acetylation of H3K9 at the activated *RARß2* promoter in shCtrl, XPC/WT, XPC/P334H as well as in shXPA cells but not in XPC deficient cells (Figure 2A). Note that the steady state level of GCN5 was not disturbed by the absence of XPC or XPA (Figure S2A). ChIP-reChIP (XPC/GCN5 or GCN5/XPC) experiments next showed that XPC and GCN5 co-occupied the *RARß2* promoter in shCtrl cells but not in shXPC cells (Figure 2C). ChIP-seq performed in ATRA treated XPC/WT cells yielded 19,758 genomic locations for GCN5 (Figure 2D). Among the 1797 XPC/Pol II co-occupied positions, 53% (948) were overlapping with GCN5 binding events. As expected, almost all of the common XPC/Pol II/GCN5 binding events are enriched for

the H3K9ac mark (99%, Figure S2B). HOMER annotation further indicated that almost 50% of them were located close to the TSS (Figure S2C).

All together our results underlined the involvement of XPC in the regulation of histones modifications at promoters of ATRA activated genes. They also indicated a significant correlation between XPC and GCN5 recruitment, associated with promoter specific H3K9 acetylation.

XPC interacts directly with GCN5

We next investigated the sequence of events that made possible GCN5 recruitment. In cells lacking GCN5 (si GCN5) (Figure 3A), *RARß2* expression was decreased compared to cells transfected with scrambled siRNA (si Ctrl) (Figure 3B). ChIP experiments next showed that silencing GCN5 did not alter the enrichment of both Pol II and XPC at *RARß2* promoter but prevent the acetylation of H3K9 compared to si Ctrl cells (Figure 3C and 3D). It seemed that the H3K9 acetylation at *RARß2* promoter required GCN5. Indeed, this later joined the promoter upon the recruitment of the basal transcription machinery.

To further investigate the relationship between XPC and GCN5, we designed several experiments. Using nuclear extracts from XPC/WT cells, we found that XPC and GCN5 co-immunoprecipitated (Figure 3E, lanes 5 and 7); such complex was not found in XPC/R579st cells (lanes 6 and 8). We next produced and purified the recombinant tagged heterodimer XPC/HR23B and GCN5 (Figure 3F, left panel). We found an interaction between XPC and GCN5 (Figure 3F, right panel). Searching for XPC/GCN5 interacting domains, we designed and purified several truncated recombinant proteins (Figure 3G)(Bernardes de Jesus et al., 2008). GCN5 was found to interact with all the XPC variants except with XPC/Q474st. This further delineates the 474-579 XPC domain as required for its interaction with GCN5.

Collectively, these data showed that the regulation of H3K9 acetylation by XPC is mediated by GCN5.

GCN5 as part of the ATAC complex is recruited by XPC

RNA-seq analysis was performed in parallel to ChIP-seq in ATRA treated XPC/WT and XPC/R579st cells. Among the 1529 genes that were covered by

XPC/Pol II binding events (Figure 1C), RNA-seq allowed the identification of 283 genes that were significantly down regulated (Table S1). We also observed that in our experimental conditions, some of them (around 200) were up regulated (see below). ChIP-seq comparison between XPC/WT and XPC/R579st cells highlighted a correlation between the level of PTMS in the TSS surrounding and the level of gene expression (Figure S3A). Indeed, down-regulated genes failed to be surrounded by H3K9ac and H3K4me3 in XPC/R579st cells while these PTMs remained higher for the up-regulated genes. Interestingly, the comparison of GCN5 ChIP-seq data between these two cell lines showed the presence of GCN5, on the promoters of down-regulated genes, correlating with increased H3K9ac (Figure S3A). Surprisingly, GCN5 was almost absent at the promoter of all the up-regulated genes in XPC/WT cells while H3K9 was acetylated. Expression of representatives of up regulated and down regulated genes such as the cyclin D1 (CCND1) and the death associated protein kinase 1 (DAPK1) respectively were analysed. CCDN1 is overexpressed in XPC/R579st cells compared to XPC/WT and XPC/P334H cells as opposed to what occurs with DAPK1 in XPC/R579st cells compared to XPC/WT as well as XPC/P334H (Figure 4A1 and A2). ChIP experiments further showed that the recruitment of XPC and Pol II (Figure 4A3 and A4) as well as the presence of acetylated H3K9 and H3K14 paralleled CCDN1 and DAPK1 expression (Figure 4A5 and 4A6) in the corresponding cells.

GCN5 is found as part of two functionally distinct coactivator complexes, SAGA (Spt Ada GCN5 Acetyltransferase) and ATAC (Ada Two A Containing); both included GCN5 or its closely related paralog PCAF as well as distinct ADA family proteins. We thus investigated which GCN5 containing complex (SAGA and/or ATAC) could be associated with XPC. Immuno-precipitations (IP) from XPC/WT nuclear extracts showed that XPC co-precipitate with ZZZ3 and WDR5 subunits of ATAC as well as with TRRAP and SPT7L subunits of SAGA respectively (Figure 4B and 4C, lanes 3). Further IPs using antibodies directed against either TRRAP or ZZZ3 subunits showed that XPC could be detected distinctly in both complexes (Figure 3C, lanes 4 and 5). Interestingly, PCAF was not immunoprecipitated by XPC although this histone acetyl-transferase HAT was detected in both SAGA and ATAC (Figure 4C).

We then were wondering whether the recruitment of such complexes was XPC dependent. While H3 acetylation parallel gene expression, ChIP experiments detect

a significant and specific enrichment of ZZZ3 (ATAC) that parallel the recruitment of GCN5 at *DAPK1* in XPC/WT and XPC/P334H cells (Figure 4A10). No enhancement of the TRRAP (SAGA complex) was observed in these ATRA treated cells. In XPC/R579st cells, ATAC as well as SAGA complexes were not recruited. Note that all the up-regulated genes as exemplified by *CCDN1* were unable to recruit GCN5 containing complexes while H3K4 and H3K14 acetylation occurred in XPC deficient ATRA treated cells (Figure 4A5 and A9, and Fig S3A). Interestingly, we detected the enrichment of ZZZ3 (but not TRRAP) together with GCN5 at *RARß2* promoter in ATRA treated shCtrl and shXPA cells while this pattern was lost in shXPC cells, indicating that the ATAC recruitment specifically depended on XPC (Figure S3B).

Knowing the dual role of XPC in transcription and DNA repair, we next determined whether GCN5 contribution to GG-NER, after UVC irradiation, involved SAGA or ATAC presence at damage sites. We then followed the coming and going of factors by ChIP/Western assays upon UV irradiation of Hela cells(Coin et al., 2008). We observed that XPC presence decreased 5min after UV irradiation (Figure 4C), once having attracted TFIIH (as indicated by the detection of its XPB subunit). At 15min post UV, we visualized a transient arrival of GCN5 together with WDR5 subunit of ATAC as well as a significant increase of H3K9 acetylation (Figure 4C, lane 3). Similarly using XPB antibodies, we also observed 15min post UV, the transient presence of GCN5 together with the ATAC subunit (Figure 4D, lane 3). To be noted also the arrival of the XPF endonuclease starting at 5min post UV, indicating that the removal of the DNA damage generated by UV irradiation was on going (Figure 4D, lanes 2-4).

Altogether, our findings indicated that H3 acetylation resulted from the recruitment of GCN5 as part of the ATAC complex by XPC, a situation also observed in DNA repair.

E2F1 cooperate with XPC for recruiting GCN5

We next investigated whether DNA binding elements could be selective in recruiting the ATAC complex in a XPC dependent manner at specific genomic locations(Krebs et al., 2011; Spedale et al., 2012). Gene Ontology (GO) indicated that most of the down-regulated genes (254/283; p-value 1,99.10-12) contained

E2F1 binding site around their TSS (Table S2). Moreover, E2F1 ChIP-seq data from ENCODE (www.encodeproject.org) highlighted a stronger E2F1 binding factor around TSS from the down regulated genes (Figure 5A).

It has been described that E2F1 can directly interact with GCN5 to stimulate transcription of its target genes(Lang et al., 2001). Antibodies directed towards E2F1 precipitated both XPC and GCN5 from XPC/WT nuclear extracts (Figure 5B). Moreover, using purified recombinant proteins (Figure 5C, right panel), we observed that antibody directed towards GCN5 precipitated XPC and E2F1 (Figure 5C, left panel, lanes 7, 9 and 10)(Singh and Dagnino, 2016), underlining a partnership between these three components.

ChIP experiments next showed the recruitment of E2F1 at the promoter of *DAPK1*, that possess an E2F1 binding site, in ATRA treated XPC proficient cells (Figure 5D). Such enrichment was lost in XPC/R579Stp cells (lacking XPC), while it was still detected in XPC/P334H cells (Figure 5D). E2F1 was not recruited at *CCDN1* that did not contain an E2F1 responsive element. We next evaluate the impact of E2F1 on CCDN1 and DAPK1 by transiently transfecting XPC/WT cells with siRNA targeting either E2F1 or GCN5 (Figure 5E). Silencing E2F1 abolished *DAPK1* transactivation, a situation also observed when silencing GCN5 (Figure 5F). On the contrary, silencing E2F1 and GCN5 did not affect transactivation of genes such as CCND1 (Figure 5F). In si E2F1 and si GCN5 cells, Pol II as well as XPC were recruited at *DAPK1* promoter (Figure 5G). In si E2F1, XPC as well as GCN5 were absent while in si GCN5 cells, E2F1 was present (Figure 5H) suggesting that E2F1 was involved in the positioning of XPC at the *DAPK1* promoter. The absence of either GCN5 or E2F1 did not allow histone acetylation around *DAPK1* promoter (Figure 5I). Silencing either E2F1 or GCN5 did not disturb *CCND1* expression (Figure 5F). In such case, as expected, Pol II, TFIIH/XPB as well as XPC were recruited and H3K9 was acetylated in the three cell lines (Figure 5G, 5H and 5I). Interestingly, ChIP-seq patterns clearly showed the presence of XPC, Pol II, GCN5, E2F1 as well as H3K9ac and H3K4me3 around the TSS of *DAPK1*, the expression of which was down regulated in XPC/R579st cells (Figure S3C). On the contrary, the absence of XPC did not prevent the recruitment of Pol II and HPTMs that was even increased around *CCND1*; in this later case we noticed that neither GCN5 nor E2F were present at the promoter (Figure S3C).

TFIIH enhances GCN5 enzymatic activity

The above data underlined a connection between XPC and H3K9 acetylation. To further evaluate the influence of XPC on histone acetylation, we set up an acetylation assay. Note that a complete *in vitro* assay including the entire transcription machinery with the E2F1 binding site in a chromatinized template was not available. We first observed that GCN5 is able to acetylate H3K9 (Figure 6A, lanes 1-2); Moreover, neither XPC nor E2F1 that co-localized and interacted with GCN5 *in vivo* (Figure 5B) as well as *in vitro* (Figure 5C, lanes 7, 9 and 10) were able to improve such H3K9 acetylation (Figure 6A, lanes 6, 7 and 12). Interestingly, the transcription/DNA repair factor TFIIH stimulates the H3K9 acetylation by GCN5 (Figure 6A, lanes 5 and 8). Moreover, addition of XPC and/or E2F1 did not enhance GCN5 HAT activity (figure 6A, lanes 13, 14 and 16). In our experimental conditions, TFIIH fraction that did not contain a HAT activity (Figure 6A, lane 5) was not found to phosphorylate GCN5 (data not shown).

Using purified recombinant proteins, we observed that antibody directed towards GCN5 precipitated TFIIH, as shown by XPB and CDK7 (Figure 6B, lanes 8 and 13). Antibodies directed towards GCN5 are also able to precipitate XPC and E2F1, in addition to TFIIH (Figure 6B, lane 19). Moreover, antibodies directed towards TFIIH (XPB) precipitated both GCN5 and XPC from XPC/WT nuclear extracts (Figure 6C, lanes 2). Interestingly, the interaction between GCN5 and XPB was lost in absence of XPC (Figure S5, right panel). In addition, antibodies against GCN5 precipitated XPB and XPC while antibodies directed towards XPC precipitated XPB and GCN5, from XPC/WT nuclear extracts (Figure 6C, lanes 1 ad 3).

ChIP experiments next showed the recruitment of TFIIH at the promoter of *DAPK1* and *CCND1*, in ATRA treated XPC proficient cells (si Ctrl) (Figure 6D). Silencing E2F1, which abolished *DAPK1* transactivation (Figure 5F), decreases the recruitment of TFIIH at the promoter of *DAPK1*; a situation also observed when silencing GCN5 (Figure 6D). Interestingly, silencing E2F1 and GCN5 did not affect the recruitment of TFIIH at the promoter of CCND1 (Figure 6D).

Altogether, these results indicated that TFIIH stimulated H3 acetylation by interacting with GCN5.

DISCUSSION

Having shown NER factors together with the transcription machinery targeting promoter of activated genes for the formation of the preinitiation complex(Le May et al., 2010a), we were next focusing on their role in the RNA synthesis process. Previous investigations have indicated the sequential arrival of the NER factors in the following order: XPC, XPA, and XPG, XPF/ERCC1, once the transcription machinery was positioned around the promoter of a given gene. Here we document the role of XPC in regulating histones PTMS at the transcription initiation level.

ChIP-seq experiments demonstrated that in ATRA treated cells, XPC together with RNA pol II was mainly localized around TSS (representing around 45% of XPC binding events), of up to 1500 activated genes (Figure 1C and 1D). XPC was also found located on other DNA regulatory regions (enhancers) such as those involved in the pluripotency regulation of ES cells(Fong et al., 2011). Focusing on fibroblasts from patients and rescued ones, we discovered that defect in XPC disrupt the expression of a certain set of genes. We show that XPC, the first NER factor that joined the transcription machinery(Le May et al., 2010a), is involved in the chromatin remodelling process, and particularly in histone post-translational modifications (Figure 2), as also observed in yeast where the ATP-chromatin remodelling complex SWI/SNF targeted XPC after UV irradiation(Gong et al., 2006). It seems that part of the chromatin remodelling process, is conditioned by the presence of XPC. Indeed, in absence of XPA (as occurred in shXPA cells), that abrogated the recruitment of all the NER factors except XPC, histone H3 acetylation around the promoter of activated gene was present(Le May et al., 2010a, 2012). However, DNA breaks, DNA demethylation and DNA loop, a role devoted to the other NER factors (Le May 2012), were absent (Figure S1B-1D**).**

Our data suggest that XPC is specifically required for the chromatin modification around the promoter of a certain set of genes, such as *DAPK1*. In this case, the H3K9/H3K14 acetylation occurs when XPC leads to the GCN5 histone acetylase recruitment (Figures 4A8 and S3), which in fact is not sufficient. This recruitment also requires E2F1. Indeed, (i) 89% of the XPC dependent genes contains E2F1 consensus binding site(s); (ii) the recruitment of both XPC and GCN5 is abrogated in E2F1 silenced cells as shown for DAPK1 (Figure 5F4 and F6); (iii)

neither E2F1 nor GCN5 are found at the promoter of activated genes in XPC-deficient cells (Figures 4A, 5D and S3). Moreover, E2F1 recruit GCN5 within ATAC complex instead of SAGA although both contain GCN5(Lang et al., 2001), highlighting a potential role of the ATAC associated proteins in discriminating expression of specific (or selected) genes. ATAC was shown to be preferentially recruited to both promoters and enhancers, while SAGA can principally be found at promoters(Krebs et al., 2011).

In addition to dissect the role of XPC during transcription, this work highlights the uniqueness of the expression of each gene. We indeed show how components (or signal) are required to activate a certain set of genes. Transcription is initiated following (1) ATRA ligand induction; this later promotes the recruitment of the corresponding nuclear receptor and the co-activator complexes. The general transcription factors as well as Pol II next targets the promoter; (2) some of these genes encompass an E2F1 responsive element and likely its cognate factor; (3) XPC is recruited after the basal transcription machinery(Le May et al., 2010a); (4) both E2F1 and XPC serve as a platform for the recruitment GCN5 within the ATAC complex. GCN5 will then proceed to H3 histone acetylation (an essential step that will allow chromatin opening) and allowed further RNA synthesis. This might explain how among the 1500 genes targeted by both XPC and Pol II after ATRA treatment, only 283 genes (as shown for *DAPK1* but also for *HOXB13* and *LRRC11*, Figure S6) were down regulated in XPC deficient cells. Some other genes such as *CCND1* were also recruiting XPC and follow another scenario in which GCN5 is not involved. We also have observed that in absence of XPC, the H3K4 methylation was defective as well as the recruitment of its enzyme partner, SET1 methyl transferase (Figure S7). The *SET1* coding gene was one of the genes down regulated in absence of XPC (Table S1). Whether or not this histone methylation modification is consecutive with the histone acetylation process will require further investigations. Moreover, it cannot be excluded that additional players such as components of ATAC, other histone modification complexes, TBP associated factors (TAFs, know to interact with DNA repair factors)(Chatzinikolaou et al., 2017; Kamileri et al., 2012) might also participate in chromatin modification. It should be noted the parallel between transcription and NER, in which GCN5 as well as E2F1 were shown to be involved in the regulation of

H3K9 acetylation(Guo et al., 2010, 2011; Kakumu et al., 2017), underlining how both processes are connected and is common role of XPC.

It seems that XPC, together with E2F1, are involved in the recruitment of GCN5 but not in the regulation of its activity. Our minimal in vitro system established in absence of both chromatinized DNA and components of activated transcription, showed that XPC and E2F1 either alone or in combination do not modify the level of H3 acetylation (Fig. 6A). However we observed that H3K9 acetylation was stimulated by TFIIH, which is not surprising knowing that: (i) intact TFIIH is required for proper histones PTMs(Singh et al., 2015); (ii) binding of TFIIH to activated genes as well as DNA lesions is impaired when silencing E2F1 (Figure 6D). Although we knew that the phosphorylation of E2F1 at DNA lesions was required to promote GCN5 and NER factors recruitment, we failed to identify phosphorylation of either GCN5 or E2F1 by the CDK7 kinase of TFIIH. It is also possible that the phosphorylation of GCN5 required for its location at DNA lesions(Guo et al., 2010, 2011), could be performed by other kinase such as the PKA. Indeed, during fasting, PKA phosphorylates GCN5, thereby increasing GCN5 acetyl transferase activity(Sakai et al., 2016).

Then arise questions about the clinical features of XPC patients that mainly display skin sensitivity to sunlight with a 1,000-fold increased susceptibility to developing skin cancer. Among all the XP patients, XP-C ones rarely exhibit visible neurological disorders or developmental defects. Interestingly, Gene Ontology point out that a large proportion of genes was related to chromatin structure regulation and present oncogenic and immunologic signatures (Table S1). Indeed, it has been demonstrated that upon genotoxic attack, XPC can induce the expression of cytokine like interleukin-6 (IL-6), that have pro-inflammatory effects in lung fibroblasts(Schreck et al., 2016). Moreover, multi-omic analyses had identified factors and pathways implicated in the cellular response to UV-induced DNA damages in relation with the immune system (Boeing et al., 2016). For example, *DAPK1* is an inhibitor of RIG-I signalling which is necessary to induce the production of type I interferon(Willemsen et al., 2017). It was also shown to act as a tumour suppressor in multiple cancer types(Kissil et al., 1997; Raveh and Kimchi, 2001).

It is clear that further identification and analysis of the deregulated genes would (i) help to determine relevant markers for an early and specific diagnosis, (ii)

anticipate/predict the cancer risk among the different symptoms within XPC patients. In addition, our work uncover the role of XPC, one of the NER factors, in the regulation of the expression of a certain set of genes through the regulation of Histone PTMs.

EXPERIMENTAL PROCEDURE

Cell Culture

HeLa Silencix cells (Tebu-Bio) including shCtrl, shXPC and shXPA cells as well as XP-C patients derived fibroblasts GM14867 (XPC/R579St) and GM02096 (XPC/P334H) and rescued XP-C (XPC/R579Stp + GFP-XPCwt)(Bernardes de Jesus et al., 2008) were cultured in appropriate medium and maintain at 37°C in a 5% CO_2 environment. 12 hours before ATRA treatment, cells were incubated in phenol red-free medium with charcoal treated-FCS and 40 mg/ml gentamycin. Cells were treated with 10µM all-trans retinoic acid (ATRA) (MP).

Antibodies

Antibodies towards His, Flag, RNA Pol II, RAR, TBP, XPA, GCN5, XPG, ZZZ3, TRRAP and tubulin were produced at the IGBMC. CDK7 (C-19), XPB (S-19), XPF (H300) and XPC (D-18) antibodies were purchased at Santa-Cruz Biotechnology. H3k4me3 (ab1012) and CTCF (ab70303) antibodies were purchased at Abcam and WDR5 (07-706) antibodies from Upstate. Antibodies against H3 (#4620), GCN5 (2676) and Spt7L were obtained from Cell signalling, Epigentek and Bethyl respectively. Antibodies against H3K9ac (61251) and H3K14ac (39599) were from Active Motif.

Reverse Transcription and Quantitative PCR

Total RNA was isolated from cells using a GenElute Mammalian Total RNA Miniprep kit (Sigma) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen). The quantitative PCR was done using a Lightcycler 480 (Roche). The primer sequences used in qPCR are indicated in Table S3. The mRNA expression of

a gene represents the ratio between values from treated and untreated cells after normalization against *GAPDH*.

Chromatin Immunoprecipitation (ChIP)

Cells were cross-linked at room temperature (RT) for 10 min with 1% formaldehyde. Chromatin was prepared and sonicated on ice for 30 min using a Qsonica Q800R as previously described[12]. Samples were immunoprecipitated with antibodies at 4°C overnight, and protein G Sepharose beads (Upstate) were added, incubated for 4 hours at 4°C, and sequentially washed. Protein-DNA complexes were eluted and decrosslinked. DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analysed by qPCR using primers as indicated in Table S3.

ChIP/Western blot on UV-irradiated cells

XPC/WT cells were seeded (10^6 cells per dish) 24h prior to the experiment, rinsed with PBS and UV-irradiated (20 J/m2). Cells were then cross-linked at room temperature for 30 min with 1% formaldehyde at indicated times post-UV irradiation and chromatin was prepared(Fousteri et al., 2006). ChIP/Western blot, using XPB or GFP antibodies, was performed as previously described(Coin et al., 2008). Briefly, the chromatin suspension was sonicated in buffer S (10 mM Tris-HCl [pH 8.0], 140 mM NaCl, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS) using the Qsonica Q800R in 20s pulse followed by 40 s cooling. Samples were sedimented (13 000 rpm, 15 min), and the supernatant that contained the cross-linked chromatin was frozen and stored at -80°C. In each assay, 600 µg of protein from cross-linked chromatin was immunoprecipitated with 1 µg of antibody in buffer S for 4h at 4°C. The immunocomplexes were collected by adsorption to protein G Sepharose beads (Upstate, Billerica, MA) overnight at 4°C. The beads were next washed three times with 5 volumes of buffer S and resuspended in 1 X Laemmlli SDS Buffer. Samples were incubated at 95°C for 90 min for crosslinking reversal prior electrophoresis.

RNA-seq analysis

Total RNA from XPC/R579St and XPC/WT cells were extracted before or 6 hours after ATRA treatment (10uM) using TRI REAGENT (MRC) and purified by

phenol-chloroform extraction. Libraries was prepared with TruSeq Stranded mRNA Sample Preparation kit following guide instruction and subsequently proceed on an Illumina Hiseq 4000 as single-end 50 base reads following Illumina's instructions. Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Reads were mapped onto the hg19 assembly of the human genome. Reads count was performed with HOMER v4.8.3(65) and differently expression was estimated with EdgeR. Genome ontology was performed with Genomic Region Enrichment Analysis Tool 3.0.0 (http://bejerano.stanford.edu/great/public/html/).

ChIP-seq analysis

Purified DNA fragments were prepared with the ChIP-IT High Sensitivity Kit (Active Motif) and the related antibodies. ChIP-seq was performed on an Illumina Hiseq 4000 as single-end 50 base reads following Illumina's instructions. Image analysis and base calling were performed using RTA 1.17.20 and CASAVA 1.8.2. Reads were mapped onto the hg19 assembly of the human genome. Peak detection was performed using MACS14 (http://liulab.dfci.harvard.edu/MACS/) under settings where the input fraction was used as negative control. Peaks detected were annotated using HOMER (http://biowhat.ucsd.edu/homer/ngs/annotation.html) as well as TSS protein enrichment comparison. As reference coordinates, we used the MACS-determined peaks of human genes as defined by RefSeq database.

Plasmids and purification of recombinant proteins

PCR products for the entire coding sequence XPC and the different variants were cloned into pDONOR-207 vector using the Gateway system (Invitrogen) and later cloned in bicistronic plasmid VEAP5317 with hHR23B kindly obtained from A. Poterszman. For recombinant Flag-GCN5 expression in Sf9 cells, the corresponding vector was kindly provided by L. Tora. PCR product for the entire coding sequence of E2F1 was also cloned into pDONOR-207 vector and later sub-cloned in pAC8 vector. Sf9 cells were infected with baculoviruses expressing a FLAG-tagged GCN5, Strep XPC/ His-tagged hHR23B or c myc-tagged E2F1 and the harvested recombinant proteins were purified as previously described(Singh et al., 2015).

GFP-ERCC3 (XPB), Flag-ERCC2 (XPD), GTF2H1 (p62), GTF2H4 (p52), GTF2H2 (p44), GTF2H3 (p34), Flag-CDK7, CCNH (Cyclin H), MNAT1 (MAT1), and

GTF2H5 (p8/TTD-A) subunits of TFIIH were produced as previously described(Dubaele et al., 2003).

Sf21 insect cells were infected with the different baculoviruses in order to separately obtain core-IIH, CAK and ERCC2. The different whole-cell extracts were incubated 4h at 4°C with anti-M2-Flag antibody bound to agarose beads. After extensive washings, the immunoprecipitated fractions were eluted. The recombinant TFIIH was made by mixing purified core-IIH, CAK, and ERCC2.

siRNA transfection

ON-TARGET plus smart pool siRNA control or targeting human GCN5 and E2F1 were purchased from Dharmacon and transfected in HeLa or XPC/WT cells at a final concentration of 100nM using X-tremeGENE siRNA transfection reagent (Roche) following manufacturer protocol.

Co-immunoprecipation

For *in vivo* co-IPs, nuclear extracts from XPC/WT and XPC/R579Spt cells were prepared as previously described (ref). After GFP-trap or GCN5, ZZZ3 or TRRAP Immunoprecipitation using the appropriate antibodies conjugated to protein G coated Dynabeads, followed by extensive washes (150mM NaCl) was carried out and the different co-precipitated proteins were detected using specific antibodies after immunoblotting.

For *in vitro* co-IPs, recombinant purified flag-GCN5 was then incubated with recombinant purified Hid-hHR23B/Strep-XPC full-length or variants before Flag immunoprecipitation was carried out. After washes, bound proteins were resolved by SDS–PAGE and detected by western blot.

Histone acetyltransferase assay

Recombinant purified GCN5 was incubated with recombinant purified proteins (XPC/hHR23B, E2F1 and TFIIH). The HAT of GCN5 was then measured using histone acetyltransferase assay as previously described (Di Cerbo et al., 2014). For the HAT assay, recombinant histones H3.3 and Octamer were incubated with GCN5 HAT and the other putative partners (XPC/HR32B, E2F1, TFIIH complex) in the

presence of acetyl-CoA (without in case of mock reaction) in HAT buffer (50 mM Tris-HCl pH8.0, 7% glycerol, 25 mM NaCl, 0.1 mM EDTA, 5 mM DTT) for 1 hour at 30°C. The reaction was then analysed by western blotting with specific antibodies (H3 and H3K9ac) and the activity of the enzyme was checked using antibodies against known specific targets.

Acknowledgements

We are grateful to Sylvain Daujat (IGBMC) for in vitro HAT assay, to A. Poterszman (IGBMC) for the VEAP5317 construct, to members of our team and to Didier Devys (IGBMC) for fruitful discussion. We thank IGBMC antibodies and cell culture facilities. Sequencing was performed by the IGBMC Microarray and Sequencing platform, a member of the “France Génomique” consortium (ANR-10-INBS-0009). This study was supported by l’Association de la Recherche contre le Cancer (ARC n°SL220130607082) and the National Research Foundation of Korea for International Collaboration. This study was supported by the Ligue contre Le cancer (CCIR-GE 2015). This study was supported by the grant ANR-10- LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d’Avenir ANR-10- IDEX-0002-02. B.B was supported by the fondation pour la recherche médicale (FDT20160736467). M.S was supported by le prix d'encouragement de la province Sud (Nouvelle Calédonie).

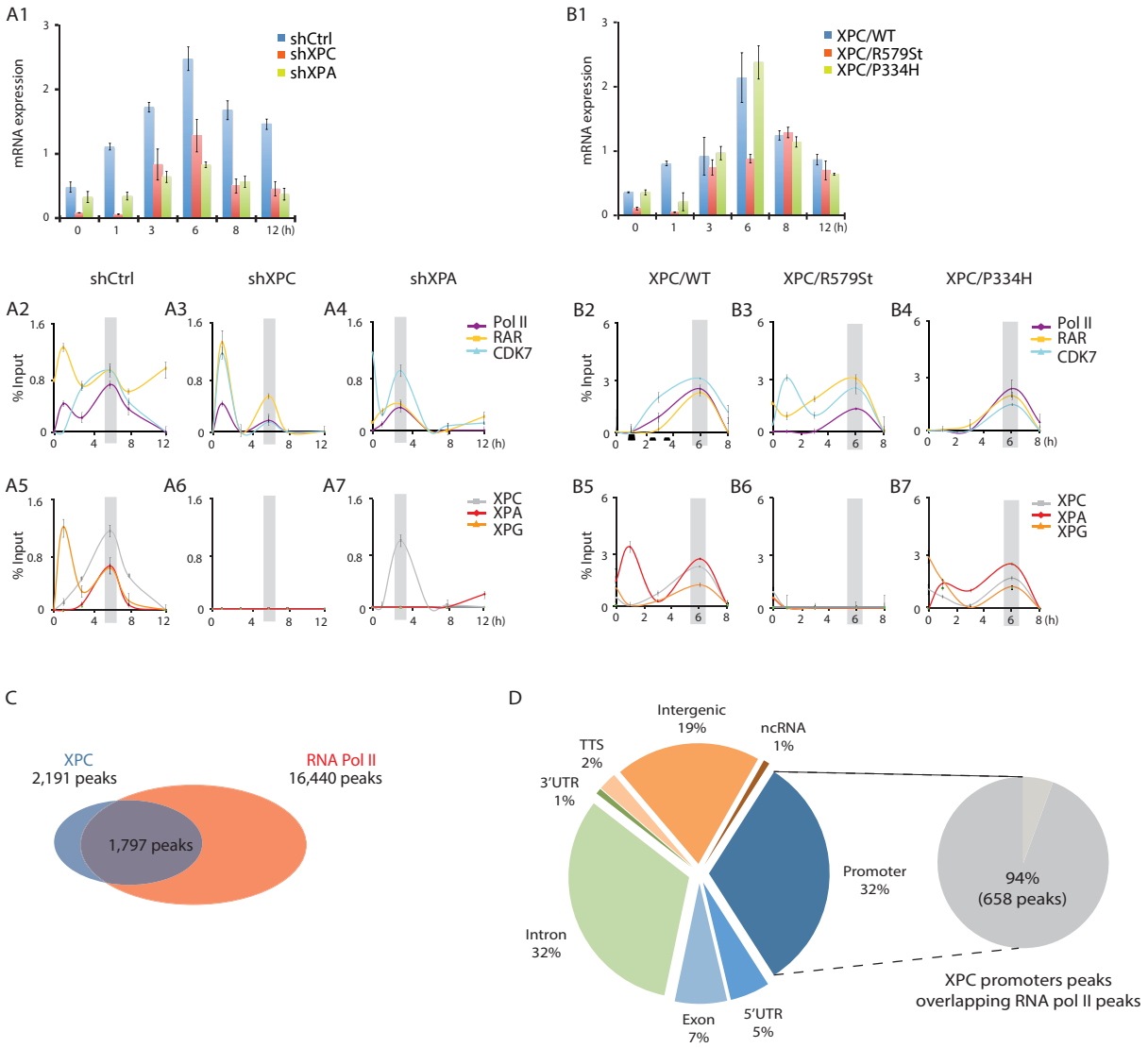


Figure 1: XPC deficiency disturbs gene expression

A-Relative mRNA expression of *RARβ2* (1) monitored by RTqPCR and respective recruitment of Pol II, RAR, CDK7, XPC, XPA and XPG at *RARβ2* promoter (2-7) monitored by ChIP in shCtrl, shXPC or shXPA HeLa cells after ATRA treatment, in 12 hours' time course experiment. B-Relative mRNA expression of *RARβ2* (1) monitored by RTqPCR and respective recruitment of Pol II, RAR, CDK7, XPC, XPA and XPG at *RARβ2* promoter (2-7) monitored by ChIP in fibroblasts XPC/WT, XPC/R579st and XPC/P334H after ATRA treatment, in 12 hours' time course experiment. C-Overlapping of MACS14 determined peaks for both Pol II and XPC in XPC/WT cells 6H after ATRA induction. XPC peaks correspond to recurrent peaks found in three independent ChIP-seq experiments. D-HOMER annotation of the 2,191 XPC peaks and proportion of promoter XPC peaks enriched in Pol II.

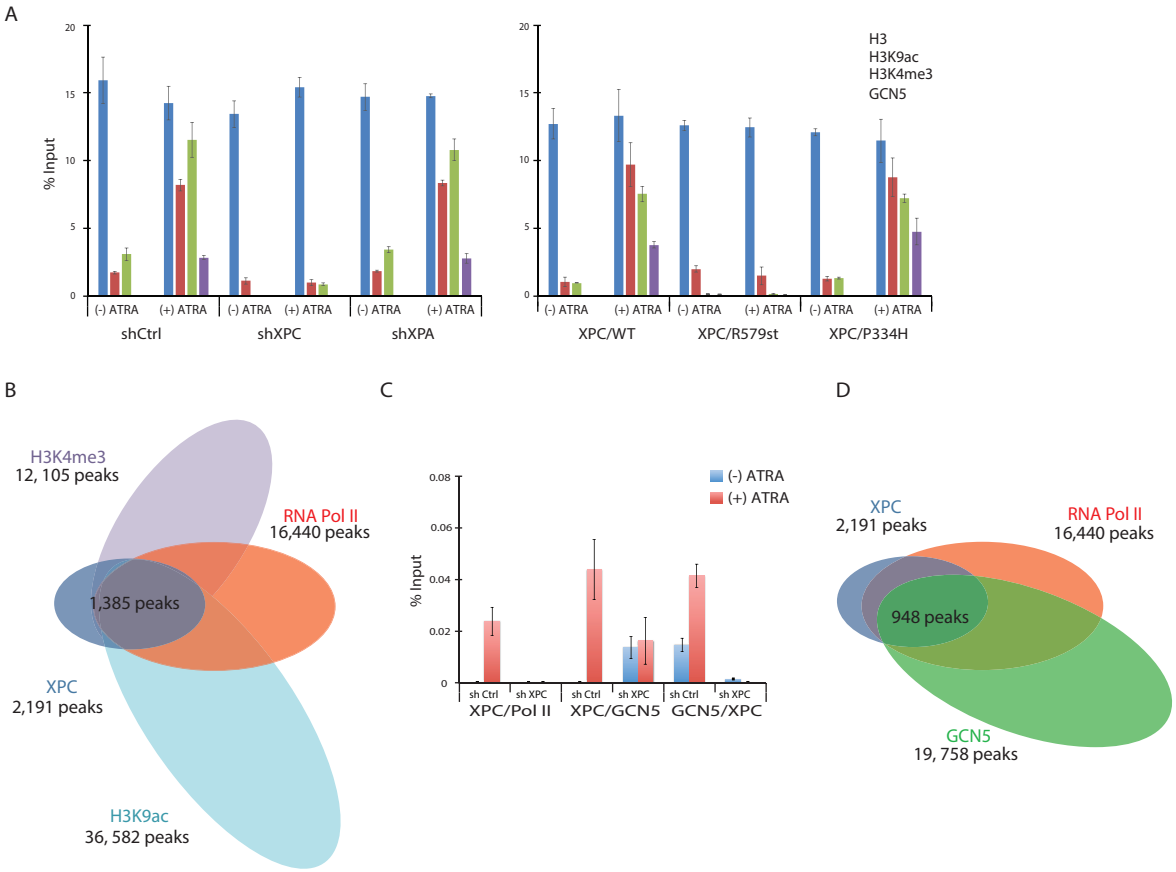


Figure 2: XPC specifically controls H3K9 acetylation by recruiting GCN5

A-Enrichment of Histone H3, H3K9ac, H3K4me3 and GCN5 at *RARβ2* promoter, monitored by ChIP in shCtrl, shXPC, shXPA HeLa cells and fibroblasts XPC/WT, XPC/R579st and XPC/P334H, before and 6H after ATRA induction. B-Overlapping of peaks for Pol II and XPC with PTMs H3K4me3 and H3K9ac in XPC/WT cells 6H after ATRA induction. C-ChIP/Re-ChIP experiment monitoring the co-occupancy of either XPC and Pol II or XPC and GCN5 in shCtrl and shXPC HeLa cells at *RARβ2* promoter upon ATRA treatment. D-Overlapping of Pol II, XPC and GCN5 peaks in XPC/WT fibroblasts 6H after ATRA induction.

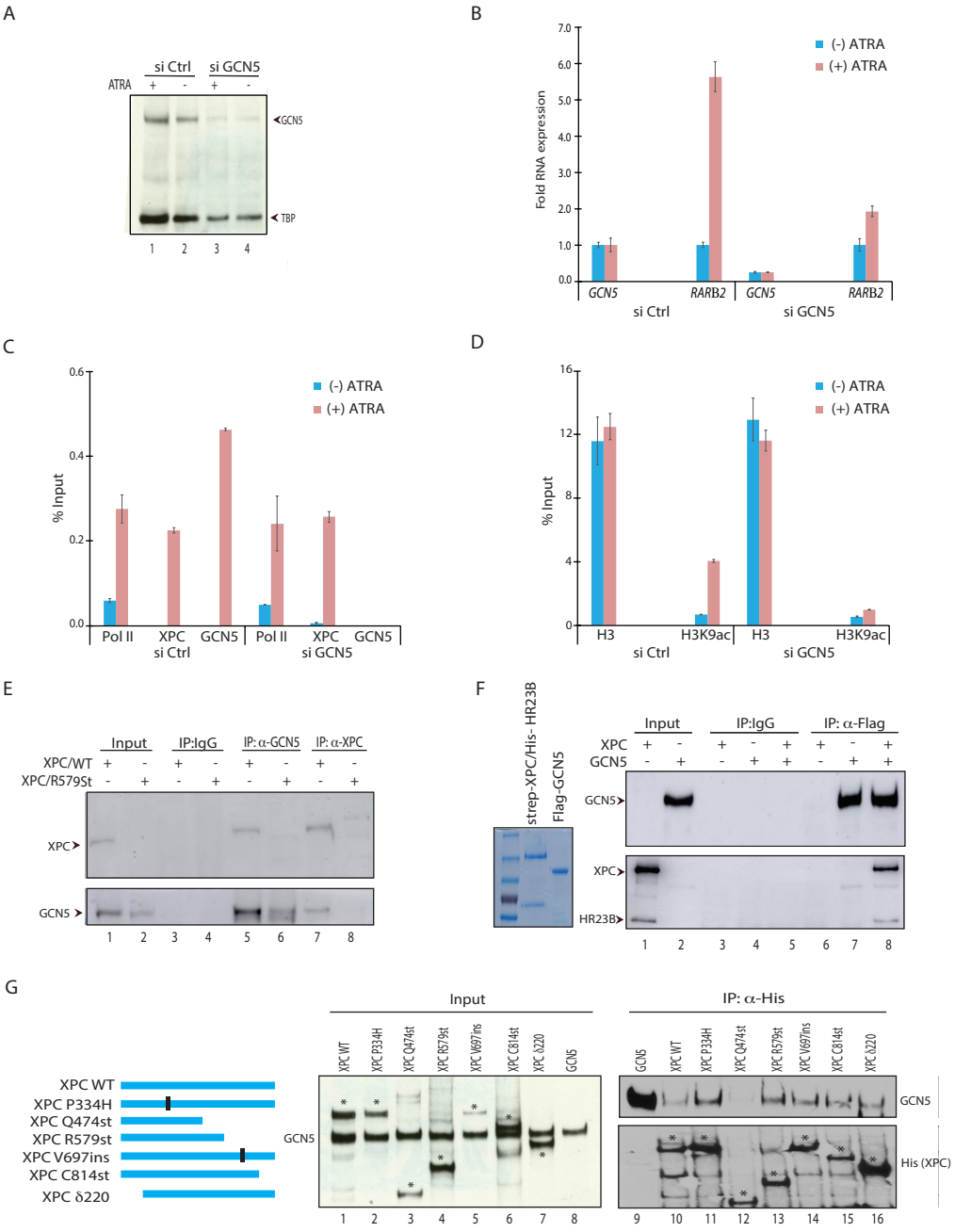

A
si Ctrl
si GCN5
ATRA
GCN5
TBP
1 2 3 4
B
(-) ATRA
(+) ATRA
Fold RNA expression
GCN5
RARB2
si Ctrl
si GCN5
C
% Input
Pol II
XPC
GCN5
si Ctrl
si GCN5
D
H3
H3K9ac
E
Input
IP:IgG
IP: α-GCN5
IP: α-XPC
XPC/WT
XPC/R579St
XPC
GCN5
F
strep-XPC/His- HR23B
Flag-GCN5
IP: α-Flag
HR23B
G
XPC WT
XPC P334H
XPC Q474st
XPC R579st
XPC V697ins
XPC C814st
XPC δ220
IP: α-His
His (XPC)

Figure 3: XPC interacts directly with GCN5

A-GCN5 and TBP protein expression monitored by Western Blot in HeLa cells treated with scrambled siRNA (si Ctrl) or siRNA targeting *GCN5* (si GCN5), before and 6H after ATRA treatment. B-Relative mRNA expression of *RARβ2* and *GCN5* in HeLa cells treated with si Ctrl or si GCN5, before and 6H after ATRA treatment. C-D-Occupancy of PolII, XPC and GCN5 as well as H3 and H3K9ac at *RARβ2* promoter in HeLa cells treated with si Ctrl or si GCN, before and 6H after ATRA treatment. E-Immunoprecipitation performed on nuclear extract from XPC/WT or XPC/R579st fibroblasts with antibody against XPC, GCN5 or IgG. F-Blue staining of recombinant Flag-GCN5 and duplex strep-XPC/His-HR23B (left panel). *In vitro* co-immunoprecipitation assay performed by antibodies against Flag tag and IgG with the recombinant protein Flag-GCN5 and the duplex strep-XPC/His-HR23B (right panel). G-Global overview of His-tag XPC full length protein (XPC FL) and several mutated proteins or truncated proteins, from ether the N-terminus or the C-terminus. *In vitro* co-immuno-precipitation assay performed with antibodies against His-tag with the different recombinant His-XPC and Flag-GCN5. INPUT (left panel) and IP (right panel) are revealed by WB against His and Flag. Asterisk represents the position of various forms of XPC.

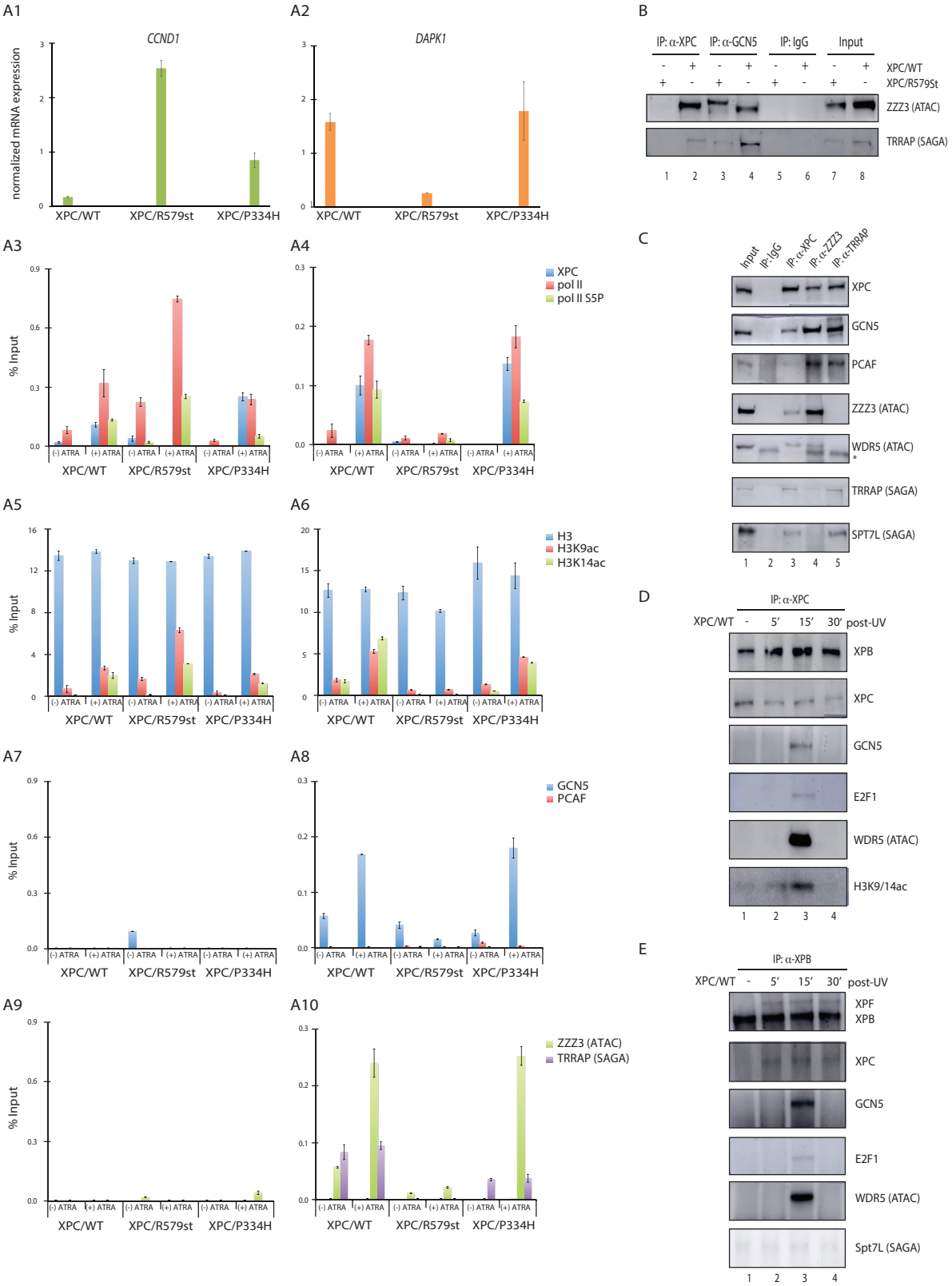

A1
CCND1
normalized mRNA expression
XPC/WT
XPC/R579st
XPC/P334H
A2
DAPK1
A3
% Input
(-) ATRA
(+) ATRA
A4
XPC
pol II
pol II S5P
A5
H3
H3K9ac
H3K14ac
A6
A7
GCN5
PCAF
A8
A9
ZZZ3 (ATAC)
TRRAP (SAGA)
A10
B
IP: α-XPC
IP: α-GCN5
IP: IgG
Input
XPC/WT
XPC/R579St
ZZZ3 (ATAC)
TRRAP (SAGA)
C
Input
IP: IgG
IP: α-XPC
IP: α-ZZZ3
IP: α-TRRAP
XPC
GCN5
PCAF
ZZZ3 (ATAC)
WDR5 (ATAC)
TRRAP (SAGA)
SPT7L (SAGA)
D
IP: α-XPC
XPC/WT - 5′ 15′ 30′ post-UV
XPB
XPC
GCN5
E2F1
WDR5 (ATAC)
H3K9/14ac
E
IP: α-XPB
XPC/WT - 5′ 15′ 30′ post-UV
XPF
XPB
XPC
GCN5
E2F1
WDR5 (ATAC)
Spt7L (SAGA)

Figure 4: GCN5 as part of the ATAC complex is recruited by XPC

A-Relative mRNA expression of *CCND1* (1) and *DAPK1* (2) in patient fibroblasts, before and 6H after ATRA treatment as well as the respective recruitment of XPC, Pol II, serine 2 phosphorylated Pol II (3-4), the relative presence of H3, H3K9ac, H3K14ac (5-6), the recruitment of GCN5, PCAF (7-8), ZZZ3 (ATAC) and TRRAP (SAGA) (9-10) monitored by ChIP at *CCND1* and *DAPK1* promoters. B-Immunoprecipitation performed on nuclear extract from XPC/WT and XPC/R579st fibroblasts with antibody against XPC, GCN5 or IgG. Western Blot was revealed with antibodies directed towards ZZZ3 (ATAC) and TRRAP (SAGA). C-Immunoprecipitation performed on nuclear extract of XPC/WT fibroblasts with antibody against TRRAP, ZZZ3, XPC or IgG. Western Blot was revealed with antibodies directed towards XPC, GCN5, PCAF, ZZZ3, WDR5 (ATAC), TRRAP and SPT7L (SAGA). D-XPC ChIP experiment and E-XPB ChIP experiment coupled to western Blot performed in XPC/WT fibroblasts UV-irradiated and harvested at indicated times. ImmunoBlot was revealed with antibodies directed against XPB, XPC, XPF, GCN5, E2F1, WDR5, SPT7L and H3K9/14ac.

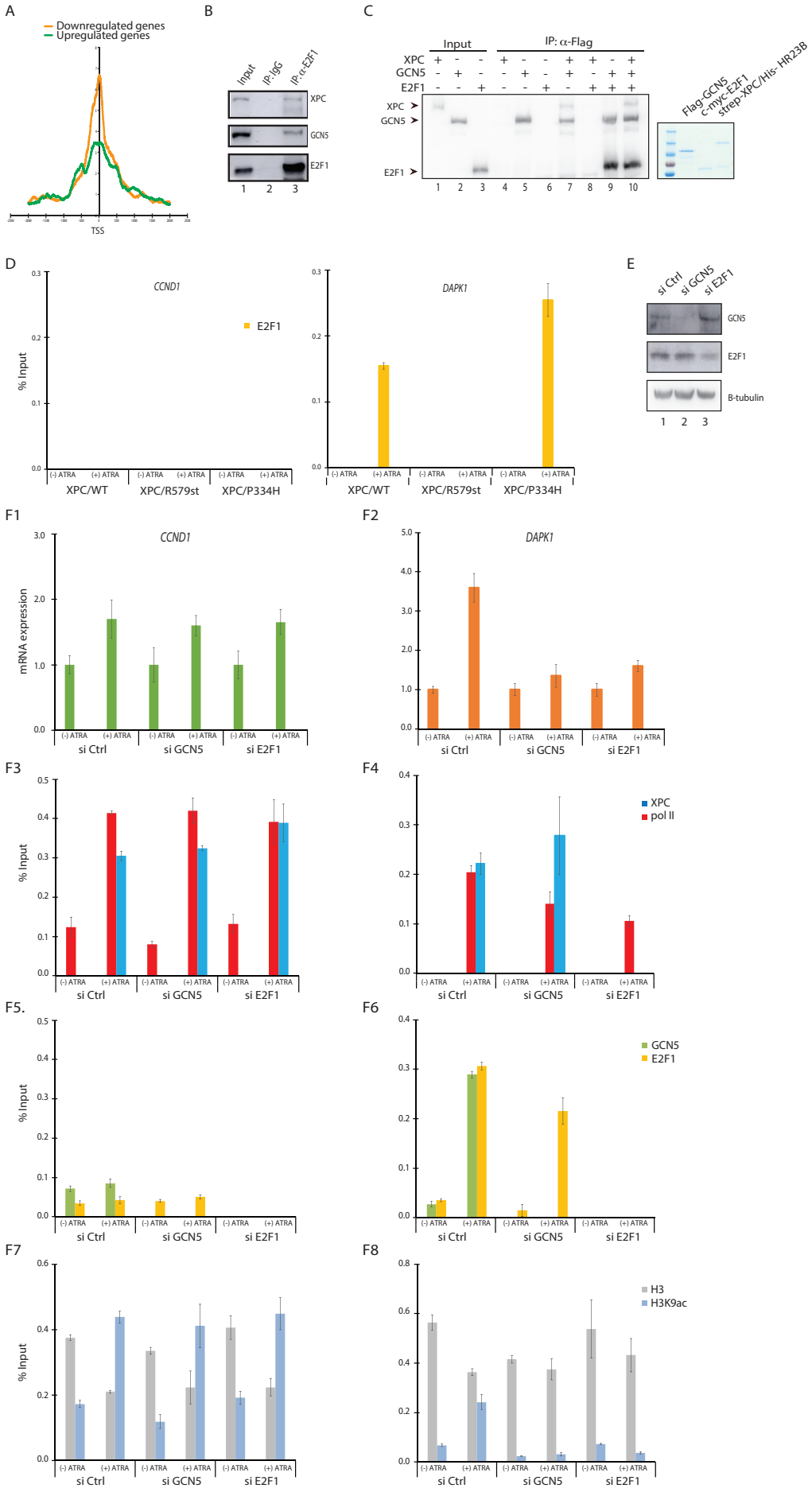

A
Downregulated genes
Upregulated genes
TSS
B
Input
IP:IgG
IP:α-E2F1
XPC
GCN5
E2F1
1 2 3
C
Input
IP: α-Flag
XPC + - - + - - + + - +
GCN5 - + - - + - + - + +
E2F1 - - + - - + - + + +
XPC
GCN5
E2F1
1 2 3 4 5 6 7 8 9 10
Flag-GCN5
c-myc-E2F1
strep-XPC/His- HR23B
D
CCND1
DAPK1
E2F1
% Input
(-) ATRA (+) ATRA
XPC/WT
XPC/R579st
XPC/P334H
E
si Ctrl
si GCN5
si E2F1
GCN5
E2F1
B-tubulin
1 2 3
F1
CCND1
mRNA expression
si Ctrl
si GCN5
si E2F1
F2
DAPK1
F3
% Input
F4
XPC
pol II
F5.
% Input
F6
GCN5
E2F1
F7
% Input
F8
H3
H3K9ac

Figure 5: E2F1 cooperate with XPC for recruiting GCN5

A-Diagrams representing the fragment depth of E2F1 ChIP-seq experiment (ENCODE project) arround up-regulated (green) and down-regulated (yellow) gene TSS, as they were previously determined in XPC/WT and XPC/R579st by RNAseq. B-Immunoprecipitation performed on nuclear extracts from XPC/WT fibroblasts with antibody against either E2F1 or IgG. Western Blot was revealed with antibodies directed against E2F1, XPC and GCN5. C-*In vitro* co-immunoprecipitation assay performed by antibodies against Flag tag with the recombinant protein Flag-GCN5, the duplex strep-XPC/His-HR23B and myc-E2F1 (left panel). Blue staining of recombinant Flag-GCN5, duplex strep-XPC/His-HR23B and myc-E2F1 (right panel). D-ChIP experiment investigating the occupancy of E2F1 at *CCND1* and *DAPK1* promoters, before and 6H after ATRA treatment in fibroblasts XPC/WT, XPC/R579st and XPC/P334H. E-Expression of GCN5, E2F1 and B-tubulin in XPC/WT cells treated with si GCN5, si E2F1 or si Ctrl, monitored by Western Blot. F-Relative mRNA expression of *CCND1* and *DAPK1*, before and 6H after ATRA treatment in si Ctrl, si GCN5 and si E2F1 treated XPC/WT (1-2) and the corresponding recruitment of XPC and Pol II (3-4), GCN5 and E2F1 (5-6) as well as the presence of histone H3 and H3K9ac (7-8) monitored by ChIP at promoter of the genes.

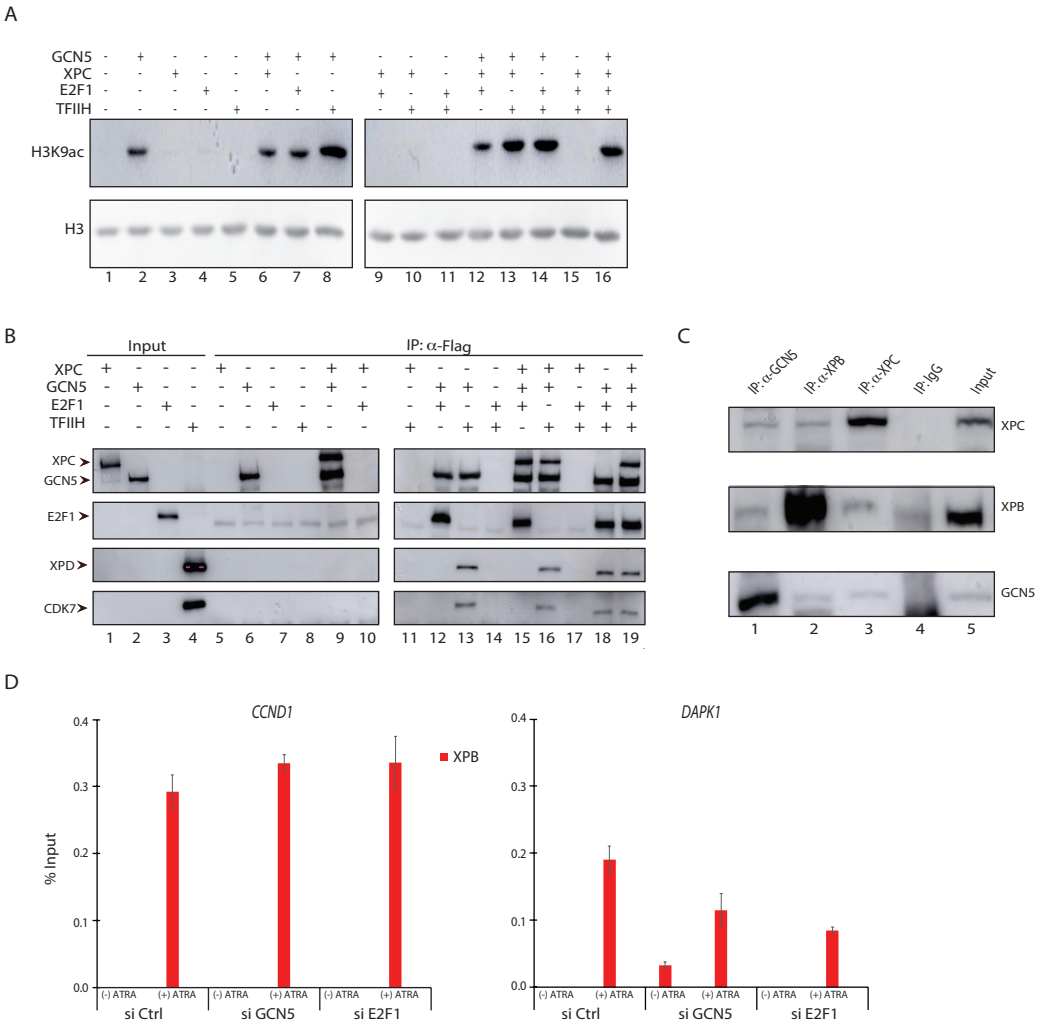
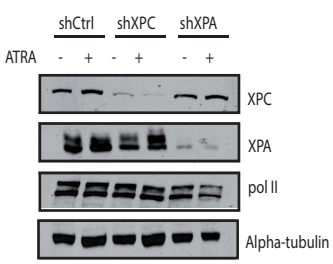
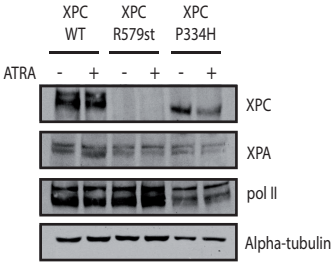


Figure 6: TFIIH enhances GCN5 enzymatic activity

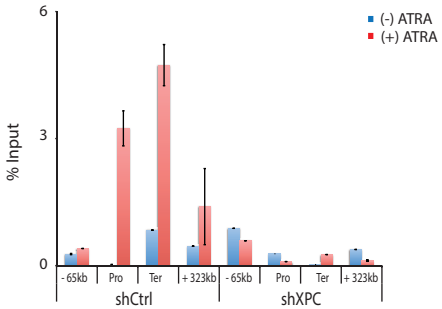
A-In vitro Histone Acetyl-Transferase (HAT) assay monitoring the acetylation of histone H3.3 by GCN5, in the presence of c-myc-E2F1, the duplex strep-XPC/His HR23B, and/or the complexe TFIIH. B-In vitro co-immunoprecipitation assay performed by antibodies against Flag tag with the recombinant protein Flag-GCN5, the duplex strep-XPC/His-HR23B, the recombinant c-myc-E2F1 and purified TFIIH. C-Immunoprecipitation performed on nuclear extracts from XPC/WT fibroblasts with antibody against GCN5, XPB, XPC or IgG. Western Blot was revealed with antibodies directed towards XPC, XPB and GCN5. D-ChIP experiment monitoring the recruitment of the XPB subunit of TFIIH at the promoter of CCND1 and DAPK1, in XPC/WT cells treated with si GCN5, si E2F1 or si Ctrl before and 6H after ATRA treatment.

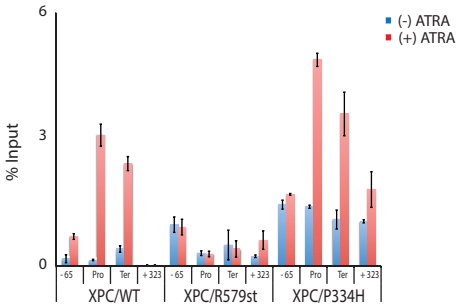
A



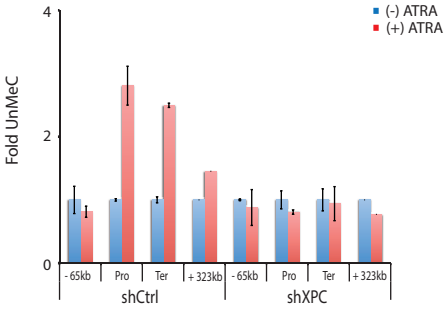


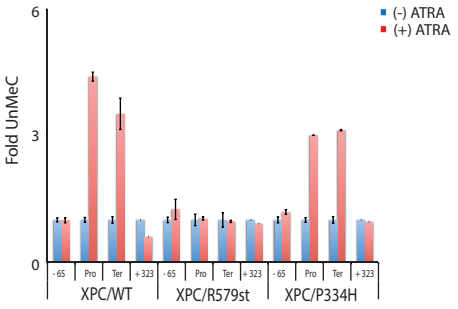
B-BIOCHIP



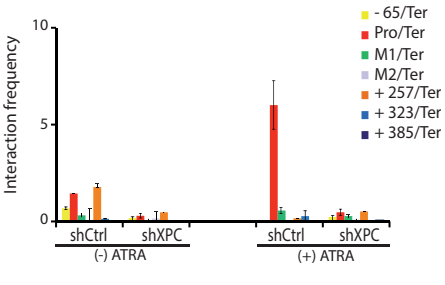


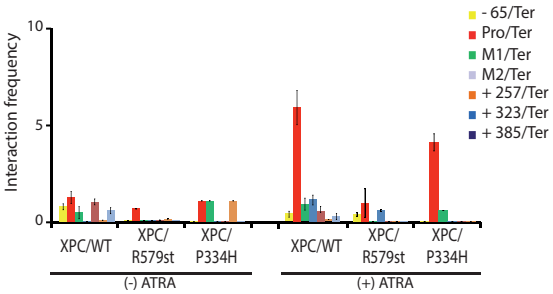
C-UnMEDIP



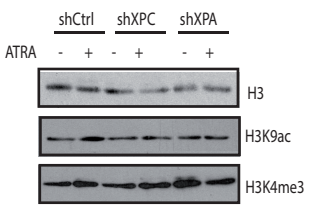


D-q3C





E



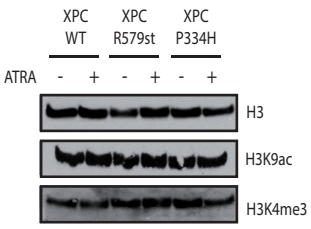


Figure S1:

A-Relative protein expression of XPC, XPA, Pol II and alpha-tubulin analysed by Western Blot in whole cell extract from shCtrl, shXPC or shXPA HeLa cells and XPC/WT, XPC/R579st and XPC/P334H fibroblasts, before and 6H after ATRA treatment. B-Detection of DNA breaks at -65kb, pro, ter and +323kb of the *RARβ2* locus in shCtrl or shXPC HeLa cells and XPC/WT, XPC/R579st and XPC/P334H fibroblasts, before and 6H after ATRA treatment by BioChIP. DNA breaks are detected through incorporation of Biotin-dUTP by terminal deoxynucleotidyl transferase followed by regular ChIP with Biotin antibodies. C-Analysis of unmethylated DNA by UnMedIP at -65kb, pro, ter and +323kb of the *RARβ2* locus in shCtrl or shXPC HeLa cells and XPC/WT, XPC/R579st and XPC/P334H fibroblasts, before and 6H after ATRA treatment. D-q3C assays were performed using crosslinked and HindIII-digested chromatin from shCtrl or shXPC HeLa cells and XPC/WT, XPC/R579st and XPC/P334H fibroblasts, before and 6H after ATRA treatment, as previously defined in Le May et al., 2012. E-Protein expression of histone H3 and presence of the H3K9ac or H4K4me3 forms analysed by Western Blot from whole cell extracts from shCtrl, shXPC or shXPA HeLa cells and XPC/WT, XPC/R579st and XPC/P334H fibroblasts, in before and 6H after ATRA treatment.

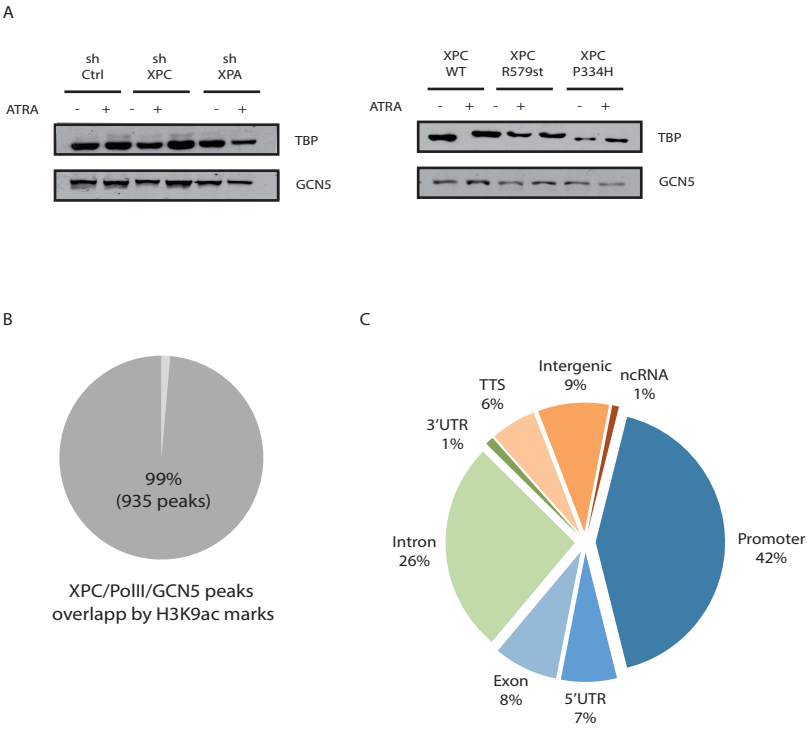


Figure S2:

A-Relative protein expression of TBP and GCN5 analysed by Western Blot in whole cell extract from shCtrl, shXPC or shXPA HeLa cells and XPC/WT, XPC/R579st and XPC/P334H fibroblasts, before and 6H after ATRA treatment. B-Proportion of XPC/Pol II/GCN5 common peaks enriched in H3K9ac mark. C-HOMER annotation of the XPC/Pol II/GCN5 common peaks.

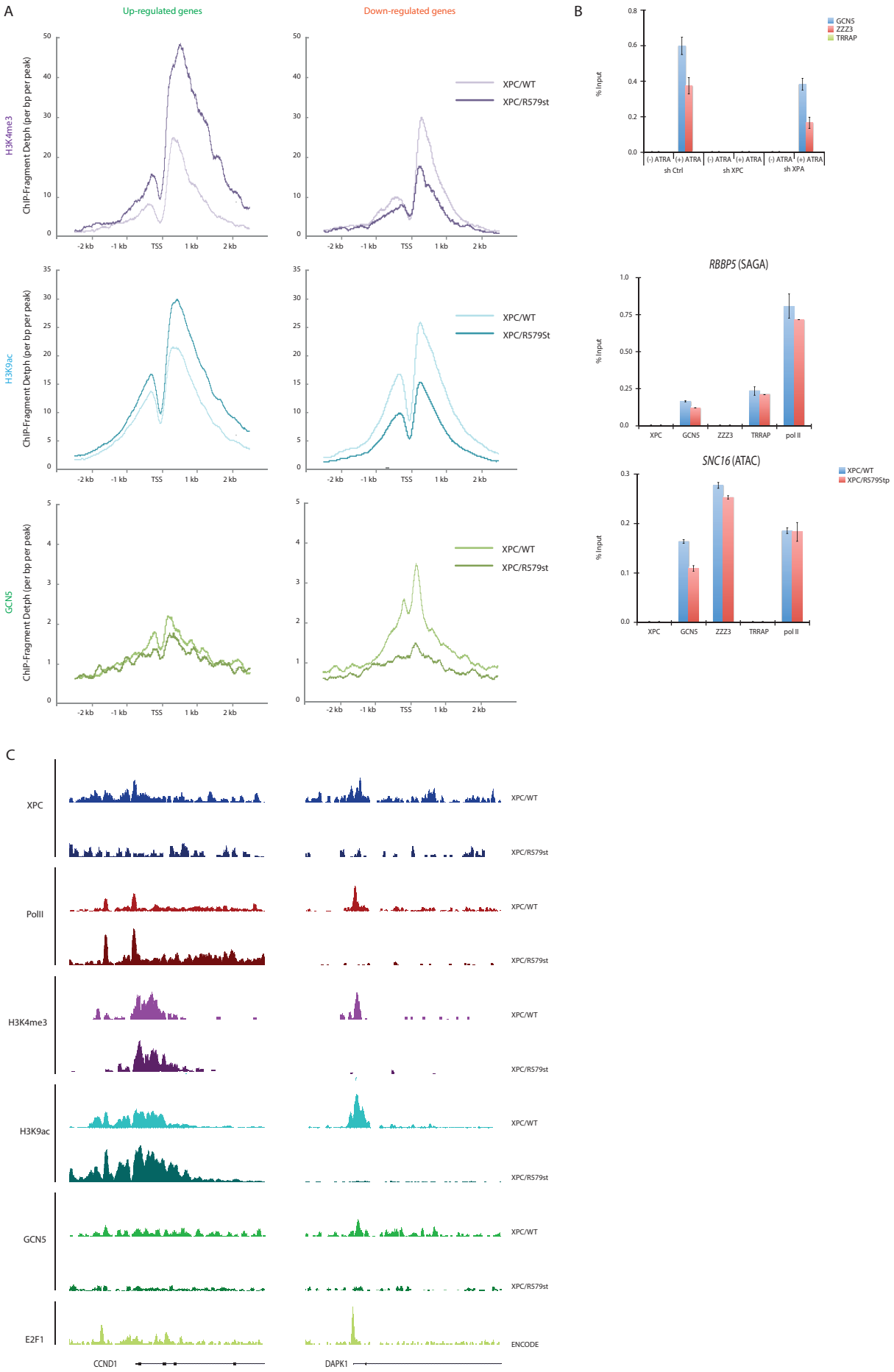

A
Up-regulated genes
Down-regulated genes
H3K4me3
ChIP-Fragment Detph (per bp per peak)
XPC/WT
XPC/R579st
H3K9ac
GCN5
-2 kb
-1 kb
TSS
1 kb
2 kb
B
GCN5
ZZZ3
TRRAP
% Input
(-) ATRA
(+) ATRA
sh Ctrl
sh XPC
sh XPA
RBBP5 (SAGA)
XPC
pol II
SNC16 (ATAC)
XPC/WT
XPC/R579Stp
C
XPC
PolII
H3K4me3
H3K9ac
GCN5
E2F1
ENCODE
CCND1
DAPK1

Figure S3:

A-Diagrams representing the fragment depth of H3K4me3, H3K9ac and GCN5 ChIPseq experiment for up-regulated and down-regulated genes, in XPC/WT and XPC/R579st. B-Enrichment at RARβ2 promoter, detected by ChIP, of histone H3, H3K9ac and H3K4me3 in shCtrl, shXPC or shXPA HeLa cells, before and 6H after ATRA treatment. ChIP experiment looking for the occupancy of XPC, GCN5, ZZZ3, TRRAP and PolII at RBBP5 (targeted by SAGA) and SNC16 (targeted by ATAC) promoters in XPC/WT and XPC/R579st fibroblasts. C-UCSC genome browser for XPC, Pol II, H3K4me3, H3K9ac and GCN5 ChIP-seq experiment in XPC/WT and XPC579st as well as E2F1 from ENCODE project, at the promoter of *CCND1* and *DAPK1*.

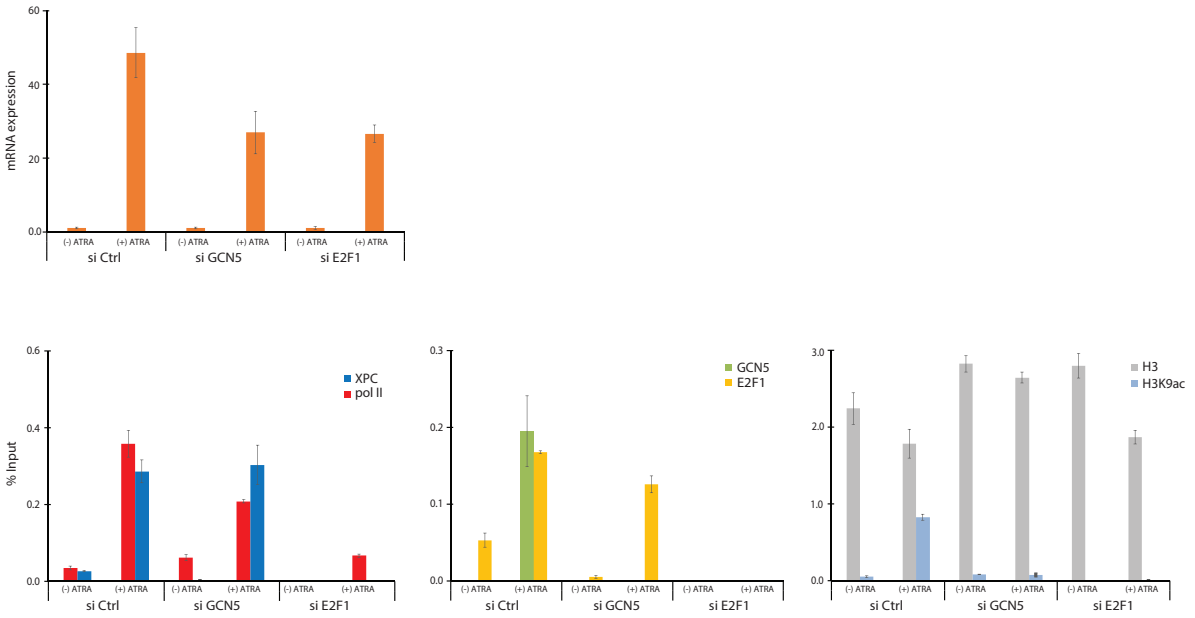


Figure S4:

Relative mRNA expression of *RARβ2*, before and 6H after ATRA treatment in si Ctrl, si GCN5 and si E2F1 treated XPC/WT and the corresponding recruitment of XPC and Pol II (left panel), GCN5 and E2F1 (middle panel) as well as the presence of histone H3 and H3K9ac (right panel) monitored by ChIP at promoter of *RARβ2*.

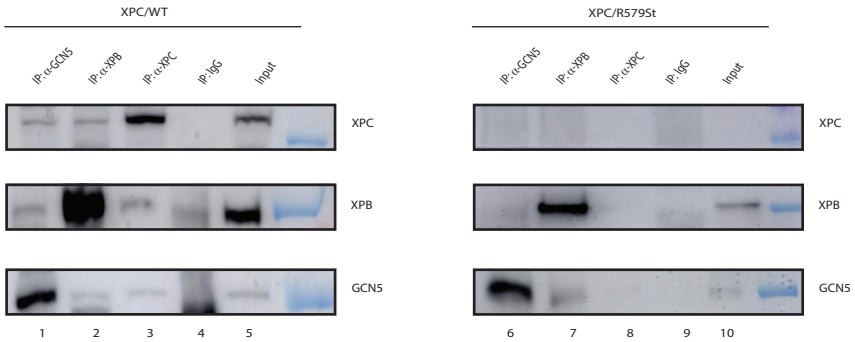


Figure S5:

Immunoprecipitation performed on nuclear extracts from XPC/WT and XPC/R579st fibroblasts with antibody against GCN5, XPB, XPC or IgG. Western Blot was revealed with antibodies directed towards XPC, XPB and GCN5.

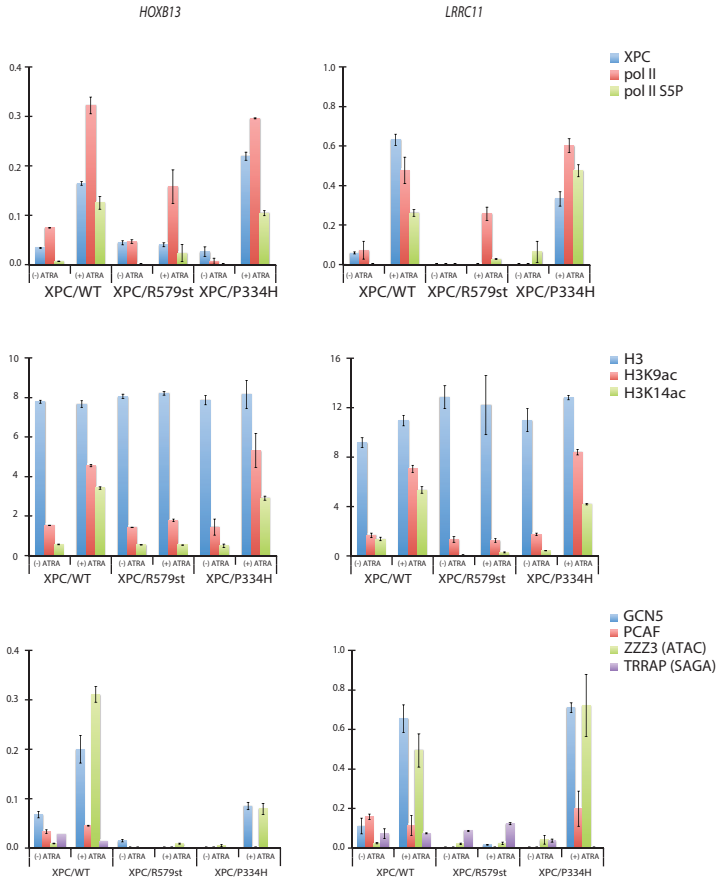


Figure S6:

Recruitment of XPC, Pol II, serine 2 phosphorylated Pol II (upper panels), the relative presence of H3, H3K9ac, H3K14ac (central panels) and the recruitment of GCN5, PCAF, ZZZ3 (ATAC) and TRRAP (SAGA) (lower panels) monitored by ChIP at *HOXB13* and *LRRC11* promoters in XPC/WT, XPC/R579st and XPC/P334H fibroblasts.

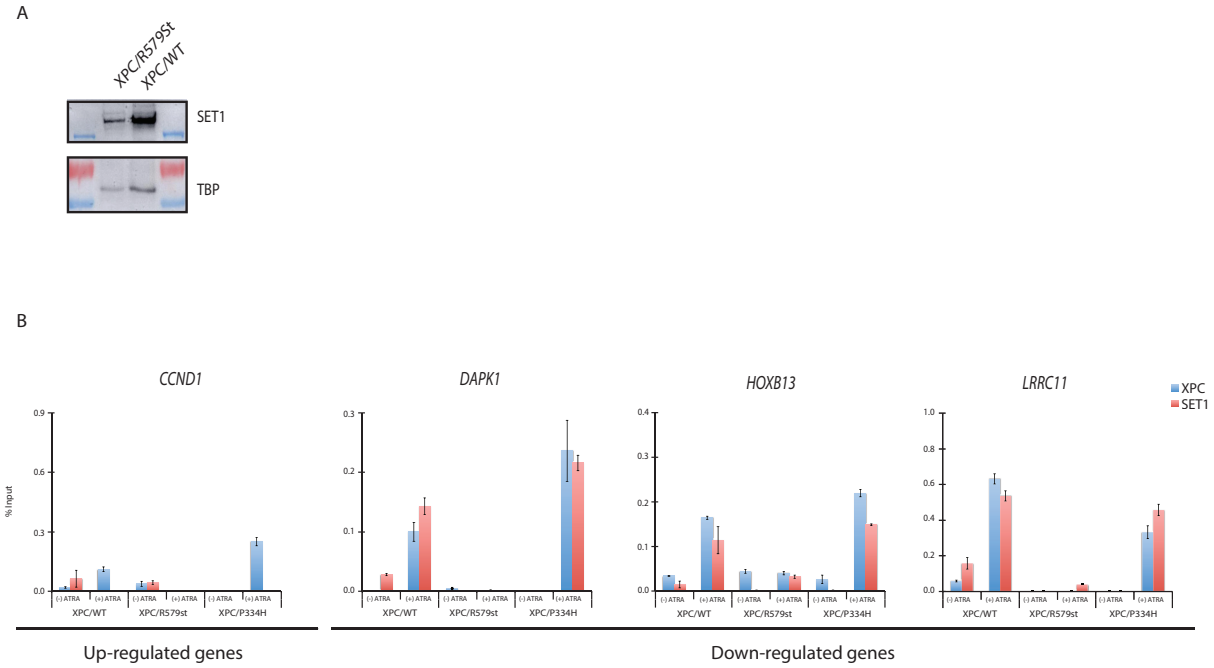


Figure S7:

A-Protein expression of SET1 and TBP analysed by Western Blot in whole cell extract from XPC/WT and XPC/R579st fibroblasts. B-ChIP experiment looking for occupancy of XPC and SET1 at CCND1, DAPK1, HOXB13 and LRRC11 promoters using chromatin extracts from XPC/WT, XPC/R579st and XPC/P334H fibroblasts.

II - The Mediator subunit MED12

MED12-related XLID disorders are dose-dependent of immediate early genes (IEGs) expression

DONNIO Lise-MARIE∞, BIDON Baptiste∞, HASHIMOTO Satoru∞, MAY Melanie, EPANCHINTSEV Alexey, RYAN Colm, ALLEN William, HACKETT Anna, GECZ Jozef, SKINNER Cindy, STENVENSON Roger E., de BROUWER Arjan P.M., COUTTON Charles, FRANCANNET Christine, JOUK Pierre-Simon, SCHWARTZ Charles E. & EGLY Jean-Marc*

Département de génomique fonctionnelle et Cancer, CNRS, INSERM, Université de Strasbourg, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 1 rue Laurent Fries, 67404 Illkirch-Graffenstaden, France.

* To whom correspondences should be address:

nlemay@igbmc.fr; egly@igbmc.fr

∞ Should be regarded as joint First Authors

OXFORD

Human Molecular Genetics, 2017, Vol. 26, No. 11 2062–2075

doi: 10.1093/hmg/ddx099
Advance Access Publication Date: 24 March 2017
Original Article

ORIGINAL ARTICLE

MED12-related XLID disorders are dose-dependent of immediate early genes (IEGs) expression

Lise-Marie Donnio[1,†], Baptiste Bidon[1,†], Satoru Hashimoto[1,2,†], Melanie May[3], Alexey Epanchintsev[1], Colm Ryan[1], William Allen[4], Anna Hackett[5], Jozef Gecz[6], Cindy Skinner[3], Roger E. Stevenson[3], Arjan P.M. de Brouwer[7], Charles Coutton[8], Christine Francannet[9], Pierre-Simon Jouk[8], Charles E. Schwartz[3] and Jean-Marc Egly[1,*]

[1]Department of Functional Genomics and Cancer biology, IGBMC, CNRS/INSERM/Université de Strasbourg, 67404 Illkirch-Graffenstaden, France, [2]Department of Clinical Pharmacology and Therapeutics Oita University Faculty of Medicine, Yufu city, Oita 879-5593, Japan, [3]Greenwood Genetic Center, Greenwood, SC 29649, USA, [4]Fullerton Genetics Center, Asheville, NC 28803, USA, [5]Hunter Genetics, Waratah, NSW 2298, Australia, [6]School of Medicine, and the Robinson Research Institute, The University of Adelaide, and South Australian Health and Medical Research Institute, Adelaide, SA 5000, Australia, [7]Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, 6525 HP, The Netherlands, [8]Département de Génétique et Procréation, Centre Hospitalier-Universitaire, Institut Albert Bonniot, CNRS/INSERM/Université Grenoble Alpes, 38000 Grenoble, France and [9]Service de Génétique Médicale, Centre Hospitalier-Universitaire, 63003 Clermont-Ferrand, France

*To whom correspondence should be addressed at: Department of Functional Genomics and Cancer Biology, IGBMC, 1 rue Laurent Fries, 67404 Illkirch-Graffenstaden, France. Tel: 0033 388653447; Fax: 0033 388653201; Email: egly@igbmc.fr

Abstract

Mediator occupies a key role in protein coding genes expression in mediating the contacts between gene specific factors and the basal transcription machinery but little is known regarding the role of each Mediator subunits. Mutations in MED12 are linked with a broad spectrum of genetic disorders with X-linked intellectual disability that are difficult to range as Lujan, Opitz-Kaveggia or Ohdo syndromes. Here, we investigated several MED12 patients mutations (p.R206Q, p.N898D, p.R961W, p.N1007S, p.R1148H, p.S1165P and p.R1295H) and show that each MED12 mutations cause specific expression patterns of JUN, FOS and EGR1 immediate early genes (IEGs), reflected by the presence or absence of MED12 containing complex at their respective promoters. Moreover, the effect of MED12 mutations has cell-type specificity on IEG expression. As a consequence, the expression of late responsive genes such as the matrix metalloproteinase-3 and the RE1 silencing transcription factor implicated respectively in neural plasticity and the specific expression of neuronal genes is disturbed as documented for MED12/p.R1295H mutation. In such case, JUN and FOS failed to be properly recruited at their AP1-binding site. Our results suggest that the differences between MED12-related phenotypes are essentially the result of distinct IEGs expression patterns,

[†]The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.
Received: January 23, 2017. **Revised:** March 3, 2017. **Accepted:** March 8, 2017

the later ones depending on the accurate formation of the transcription initiation complex. This might challenge clinicians to rethink the traditional syndromes boundaries and to include genetic criterion in patients' diagnostic.

Introduction

Mammalian mediator (MED) is an evolutionary conserved multi-protein complex that is a key regulator of gene expression involved in cell growth, homeostasis, development and differentiation (1–3). Mediator is composed of more than 30 subunits, arranged in four different modules (Fig. 1A). Upon gene activation, MED conveys essential information from transcriptional regulatory proteins bound at DNA responsive elements to RNA polymerase II (Pol II) machinery bound at the transcription start site (TSS) (4–6).

Dysfunction of the transcriptional machinery components has been shown to elicit a range of effects on cell states (proliferation or differentiation) giving rise to diverse pathologies, including cancers (7). Mutations in MED subunits are also associated with a wide range of genetic disorders, such as infantile cerebral and cerebellar atrophy (MIM: 613668) (*MED17, MIM: 603810*) (8), Charcot-Marie Tooth disease (MIM: 605589) (*MED25, MIM: 610197*) (9) and non-syndromic intellectual disability (MIM: 614249) (*MED23, MIM: 605042*) (10) and (MED13L, MIM: 608771) (11); most of them exhibiting neurological defects (12). Germ line mutations of MED12 (MIM: 300188) have already been found in several genetic disorders associated with X-linked intellectual disability (XLID) (13), such as Opitz-Kaveggia syndrome also named FG syndrome (MIM: 305450) (p.R961W) (14) and (p.G958E) (15), Lujan syndrome (MIM: 309520) (p.N1007S) (16) and Ohdo syndrome (MIM: 249620) (p.S1165P, p.R1148H and H1729N) (17,18). More recently, an important number of mutations have also been associated with non-syndromic XLID (p.S1967QfsX84) (19), (p.R521H) (20), (p.R621Q) (21), (p.I771T) (22), (p.R815Q) (23), (p.I1023V) (24) (p.A1383T) (25) and (p.E1974H) (26) (Fig. 1B). However, the underlying basis leading to phenotypically distinct syndromes from several genetic mutations of MED12 remains unclear.

MED12 gene is located at Xq13.1 and together with MED13, CDK8 and Cyclin C (CCNC) forms the Kinase module, that variably associates with the core Mediator (27). MED12 is required for the stable incorporation of CDK8/CCNC into Mediator and appears to initiate the CDK8 kinase activity (28) which can regulate transcription (29). For example, CDK8 phosphorylation of Pol II leads to the disruption of Mediator–Pol II interactions thus resulting in transcriptional inhibition (30). Similarly, CDK8 phosphorylates Cyclin H, a subunit of the general transcription/DNA repair factor TFIIH, and thus represses both the ability of TFIIH to activate transcription and its CTD kinase activity (31). MED12 has also been linked with the Notch, Wnt and Sonic hedgehog signalling pathways and control key aspects of brain development and function, from initial patterning to neuronal plasticity (32–37).

Here, we first describe a male patient with characteristics including intellectual and developmental delay (38). He was found to carry a maternally inherited missense mutation in *MED12* (MED12/p.R1295H). In an attempt to deepen our knowledge about the role of Mediator in the regulation of gene expression, we have deeply investigated the effect of this mutation on the transcriptional activation of key genes. We discovered that this MED12 mutation specifically modulates the expression of activated immediate early genes (IEGs) such as *JUN* (MIM: 165160) by disturbing the formation of the transcription complexes. Consequently, late response genes (LRGs) that are regulated by the AP1 complex (FOS/JUN dimer) are also disrupted. We then enlarged our study to seven different mutations and come across distinct deregulation of *JUN/FOS*(MIM:164810)/*EGR1*(MIM:164810) expression, possibly explaining the large clinical spectrum covered by MED12-related patients. This might challenge clinicians to rethink the traditional syndromes boundaries and to reconsider patients' diagnostic through genetic criterion.

Result

MED12/p.R1295H mutation disturb IEGs expression

Mediator is known to mediate the response of *IEGs*, involved in diverse processes such as brain development and neuronal plasticity (10,39). To evaluate the global impact of MED12 on *IEGs* expression, we used CRISPR/Cas9 system to Knock-Down *MED12* in human LAN-1 neuroblastoma cell line (MED12-KD). We first observed that silencing *MED12* does not impact the level of other Mediator subunits (Fig. 1C). We thus evaluated the expression of three *IEGs*: *JUN, FOS* and *EGR1* which code for DNA binding factors, 30 min after serum addition to wild type (WT) and MED12-KD cells. Quantitative RT-PCR on total mRNA extract clearly showed that *JUN, FOS* and *EGR1* expression is down regulated in MED12-KD cells (Fig. 1D) as confirmed by the decrease of their protein synthesis (Supplementary Material, Fig. S1).

We next focused on a *MED12* patient from whom skin fibroblasts had been obtained (Table 1; Supplementary Material, Fig. S2A and B; see 'Materials and Methods' section). The R1295H patient and his heterozygous mother carried a G-to-A missense substitution at nucleotide position c.3884 resulting in an arginine to histidine modification (p.R1295H) referred as R1295Hson and R1295Hmother, respectively (Fig. 1E). This patient is currently monitored for intellectual disability and delayed motor skills. He has a Marfan habitus (MIM: 154700) and was suspected to have X-linked Lujan-Fryns syndrome. His mother had the same skeletal and morphologic features as well as a mild intellectual disability (38). Sanger sequencing of the MED12 transcript revealed that ~80% of cultivated mother fibroblasts express the mutant allele, due to skewed X-inactivation. RT-qPCR and Western blot analysis on both patient and mother fibroblasts shows that the mutation does not significantly affect mRNA expression of MED12 and MED23 subunits of Mediator as well as the corresponding protein level (Supplementary Material, Fig. S3). However, we repeatedly observed that CDK8 subunit of the Kinase module was slightly less expressed in R1295H fibroblasts (Supplementary Material, Fig. S3).

In response to serum mitogens, *JUN* activation was defective in both R1295Hson and R1295Hmother fibroblasts while *FOS* and *EGR1* were similarly expressed when compared with WT cells (Fig. 1F). In order to gain confidence in the consequence of MED12/p.R1295H mutation in IEGs serum induction response, we performed clonal selection (cs) from the R1295H mother's fibroblasts and obtained cells expressing only one allele due to X-inactivation (see 'Materials and Methods' section). In csR1295H clone selected cells, we similarly observed that *JUN*

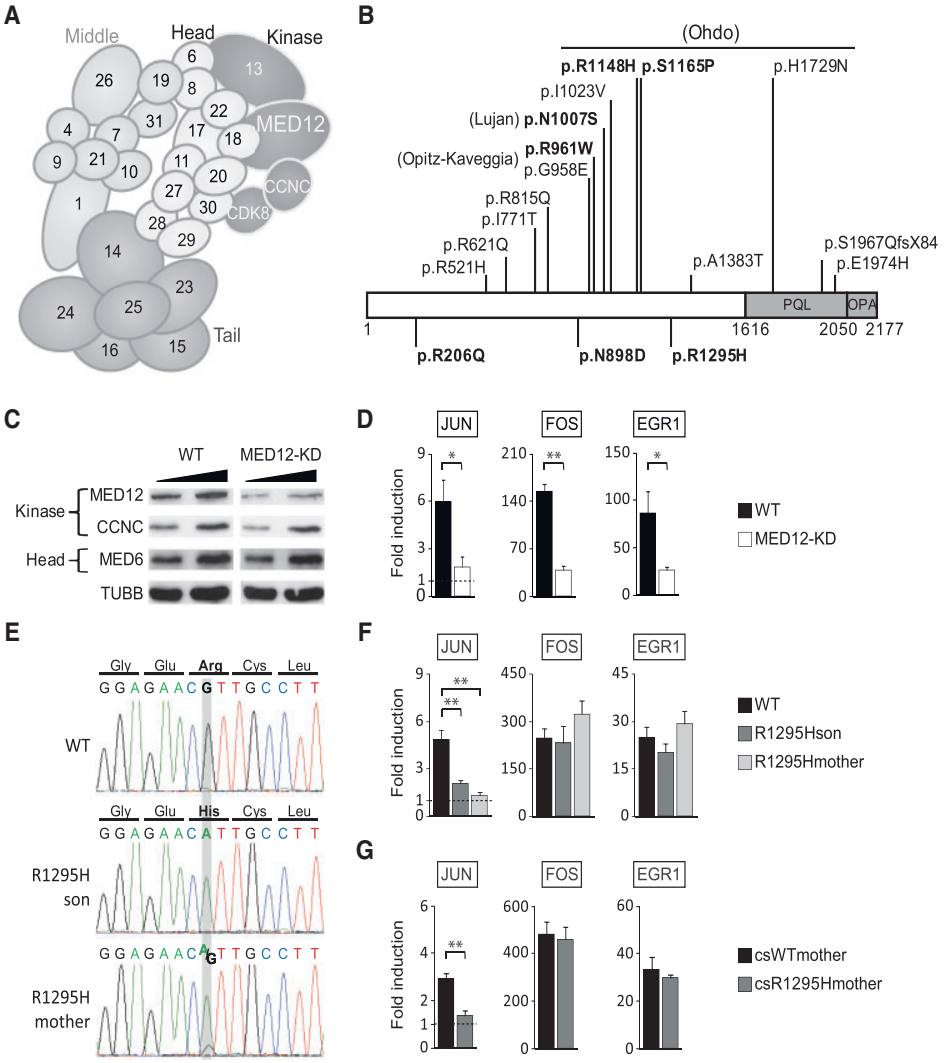


Figure 1. Human Mediator complex composition and MED12 mutations. **(A)** Global architecture of Mediator complex. **(B)** Schematic overview of MED12 protein with the PQL (proline-, glutamine- and leucine-rich) domain and the OPA (glutamine-rich) domain. The three newly identified amino acid changes are indicated (bottom), as well as previously published amino acid changes (top). Mutations in **bold** are investigated in this study. Mutations in *italic* are not associated with a specific syndrome. **(C)** WT and MED12-KD neuroblastoma whole cell extracts were resolved by SDS-PAGE and processed by Western Blot analysis using antibodies against MED12, as well as CCNC and MED6. **(D)** The neuroblastoma cells were treated with serum for 30min after serum starvation and expression of *JUN*, *FOS* and *EGR1* were monitored by RT-qPCR. **(E)** Electropherograms of the nucleotide variation c.3884 G>A of MED12 sequence in the patient (middle), his mother (bottom) and healthy control (top) fibroblasts. **(F)** The fibroblasts were treated with serum for 30 min after serum starvation and expression of *JUN*, *FOS* and *EGR1* were monitored by RT-qPCR. **(G)** Mother fibroblasts were submitted to successive dilution for cs to obtain cells expressing only one allele: csWT (expressing WT allele) and csR1295H (expressing mutated allele). Expression of *JUN*, *FOS* and *EGR1* were monitored in both csWT and in csR1295H. (*corresponds to $P < 0.05$; **corresponds to $P < 0.01$).

was down regulated compared with clone csWT while *FOS* and *EGR1* (according to *P*-value) were not modified (Fig. 1G). This result emphasized what was first observed with patient fibroblasts (Fig. 1F). Therefore we decided to use these cells as references in the following experiments.

Deregulation in the transcription complex formation at the JUN promoter

To further understand how a single point mutation disrupt the expression of a given gene, we set up Chromatin Immunoprecipitation (ChIP) experiments. We observed that the absence of *JUN* induction correlate with the defective recruitment of MED12 and CDK8 part of the kinase module as well as Pol II and TFIIB, a general transcription factor (Fig. 2A1–A4). *JUN* is regulated by several responsive elements, including one that is targeted by transcription factor 4 (TCF4), involved in neurological development and mutated in Pitt-Hopkins syndrome (40–42) Even though TCF4 was normally recruited at its distal responsive element (data no shown), its interaction with the basal transcription machinery via Mediator is strongly disturbed in R1295Hson cells when compared with WT cells (Fig. 2A5). Corresponding with the normal levels of *FOS* and *EGR1* expression, we observed a normal recruitment of MED12, CDK8, Pol II, TFIIB, as well as the phosphorylated form of ELK1 (ELK1-P) at their promoters (Figs. 2B1–B5 and 2C1–C5). *FOS* and *EGR1* expression are regulated by the cooperative binding of ELK1, ELK3 or ELK4 to serum response elements (SREs) in WT cells. These factors which belong to the ETS family have roles in various contexts, including long-term memory formation, drug addiction, Alzheimer's disease, Down syndrome and depression (43–45).

Gene activation is accompanied by important chromatin remodelling events. Euchromatin, which enables transcription to occur, is mainly characterized by acetylation of H3K9 histone

Table 1. Table summarizing the clinical features of MED12 patients and IEGs responses

	R206Q	N898D	FG R961WG958E	Lujan N1007S	Ohdo R1148HS1165P	R1295H
Growth						
Tall Stature	+	−	−	+	−	1/3
Macrocephaly	+	+	+	+	−	2/3
Neurological						
Intellectual disability	+	+	+	+	+	3/3
Agenesis of corpus callosum			+	+	−	1/3
Hypotonia		+	+	+	+	3/3
Behavioural disturbance	+	+	+	+	+	1/3
Speech abnormalities	+	−	−	−	+	3/3
Craniofacial						
Long narrow face			−	+	−	3/3
Tall prominent forehead		+	+	+	+	3/3
Triangular face			−	−	+	1/3
Blepharophimosis			−	−	+	−
Downslanting palpebrae	+	+	+	+	+	3/3
Eyes problem	Strasbismus	Astigmatism	Strasbismus	Strasbismus	Strasbismus	Strasbismus 1/3
Hypertelorim		+	+	−	−	1/3
Small ears		+	+	−	+	1/3
Philtrum		Short		short	long	Long
Maxillary hypoplasia		+	+	+	−	2/3
Micrognathia		−	+	+	+	−
High narrow palate		+	+	+	+	3/3
Open mouth		+	+	+	+	−
Dental anomalies		+	+	+	−	1/3
Extremity						
Foetal finger pads		−	+	−	−	−
Syndactyly		−	+	−	−	−
Broad thumbs/toes		+	+	+	−	−
Horizontal palmar crease		−	+	−	−	−
Long hyperextensible digits		+	−	+	+	1/3
Cardiovascular						
Congenital heart defect	LVH	Spontaneous closure ASD	−	−	−	−
Gastrointestinal						
Constipation		+	+	−	+	2/3
Anal anomalies		+	+	−	−	1/3
Genito-urinary						
Genital anomalies	+	+	+	−	+	3/3
Others	Extra Nipples					Extra Nipbl 1/3 Thoracic kyphosis 2/3
Serum						
JUN	−	=	−	−	=	=
FOS	=	=	+	+	+	+
EGR1	−	−	=	=	−	=

ASD, Atrial septal defect; LVH, left ventricular hypertrophy.

(H3K9ac) while heterochromatin which inhibits RNA synthesis, is characterized by a different set of chromatin markers such as dimethylation of H3K9 (H3K9me2) (46,47). Here, we observed an increase of H3K9ac marks around the FOS and EGR1 promoter(s) concomitantly with a decrease of H3K9me^2 in the two cell lines, following genes activation (Fig. 2B6, B7 and 2C6, C7). In contrast, the JUN promoter remains in a heterochromatic state in R1295Hson cells compared with WT cells, as shown by the absence of both H3K9ac accumulation and H3K9me^2 loss (Fig. 2A6 and A7).

Altogether the above data clearly showed that the MED12/p.R1295H causal mutation drastically affects the expression of *JUN* by disturbing the formation of an active transactivation complex on its promoter.

MED12/p.R1295H mutation deregulates a large set of genes

We next examined the biological consequences of *IEGs* deregulation on the expression of *LRGs*. We have sequenced total RNA extract 3h after serum induction, from both R1295Hson and WT cell lines. We determine log2-fold change between non-treated and treated cells in both control and R1295H patient (Fig. 3A). With a minimum ratio of 2 between inductions, we observed a severe deregulation of a large set of gene (Fig. 3A). Globally, ChIP-seq experiments on both Pol II and MED12 further showed a significant reduction of the recruitment of these two essential proteins around the TSSs all over the genome, 3 h after serum induction (Fig. 3B). Interestingly, we noticed the identical

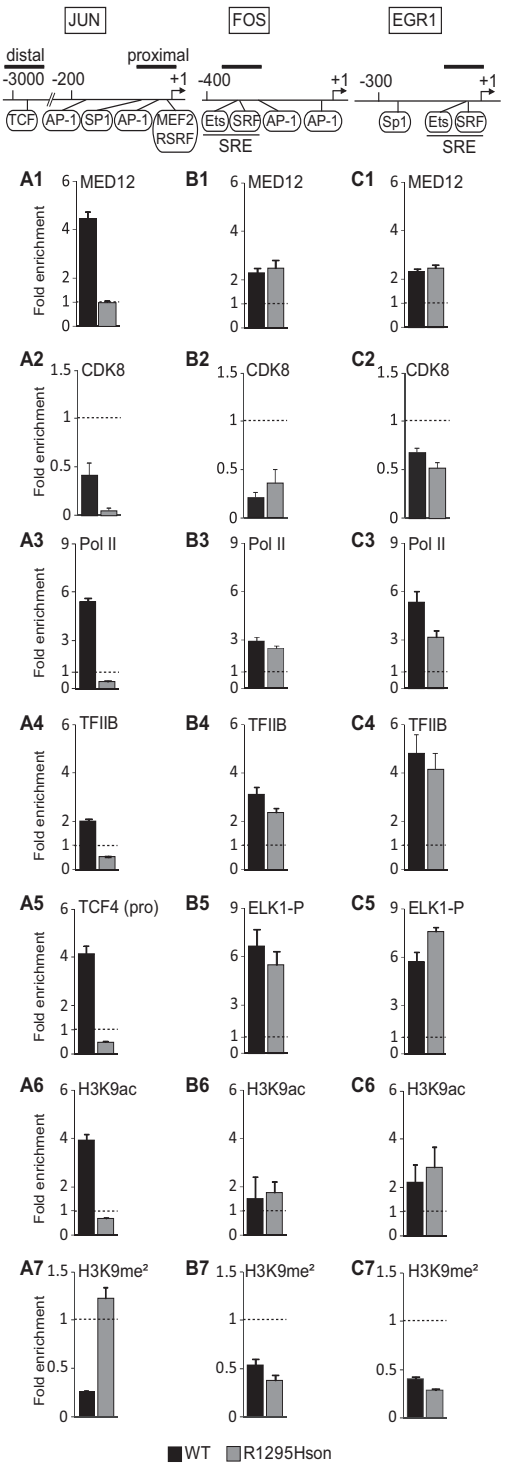


Figure 2. Promoter dynamics of *IEGs* in *MED12* mutated fibroblast. Schematic representations of *JUN*, *FOS* and *EGR1* promoters show the designed PCR amplicons. ChIP experiment monitoring the serum-dependent recruitment of MED12 **(A1, B1, C1)**, CDK8 **(A2, B2, C2)**, Pol II **(A3, B3, C3)**, TFIIB **(A4, B4, C4)**, TCF4 **(A5)**, ELK1-P **(B5, C5)**, H3K9ac **(A6, B6, C6)** and H3K9me^2 **(A7, B7, C7)** on the IEGs promoter in WT or patient cells. Values are expressed as fold enrichment (ratio of the INPUT percentage between treated and non-treated cells). Present values are representative of at least three independent experiments.

positioning of both Pol II and Mediator around TSS in WT cells. Gene expression deregulation includes significant down-regulation of 3550 genes in R1295Hson cell compared with normal cells (Fig. 3A, orange, supplementary material). By motif analysis, searching for AP1 responsive element 'TGACTC' on both strands, we measure up to 70% of down regulated genes to contain a AP1-binding site for JUN/FOS DNA binding factor (Fig. 3C). On such genes, we detected a much lower enrichment of elongating Pol II along the genes body in R1295H patient cells compared with WT cells (Fig. 3D). Interestingly, genome ontology reveals that these genes are implicated in several neurological processes (Fig. 3E). Notably, numerous genes are implicated in neural differentiation and proliferation, two pathways known to be associated with macrocephaly.

LRGs expression

AP1 is a transcription factor composed of proteins belonging to JUN and FOS families that regulates lately expressed genes through its binding to specific AP1 sites, as show on diagrams Figure 4 (below expression panel of each genes) (48). We then tested the expression of three AP1 responsive genes 3 h after serum induction. The matrix metalloproteinase-3 (MMP3, also known as Stromelysin-1), implicated in spatial learning, neural and synaptic plasticity (49,50), was down regulated in the R1295Hson cells compared with WT cells (Fig. 4A1), as observed by RNA-seq (Fig. 3A). The decreased *MMP3* (MIM:185250) expression was correlated with a decrease in Pol II recruitment (Fig. 4A2); we especially noticed a deregulation in the presence of the JUN/FOS AP1 complex at the MMP3 promoter in R1295Hson cells compared with WT cells (Fig. 4A3 and A4). Additionally, we also observed a significant defect in MED12 recruitment. The recruitment of MED6 subunit of the core Mediator was much lower in patient cells than in WT cells (Fig. 4A5 and A6). Such defective recruitment of transcription machinery paralleled the methylated status of H3K9me^2 histone at its promoter (Fig. 4A7).

Expression of Cyclin D1 (*CCND1*, MIM:168461) that connects the AP1 complex with G1 phase progression (51), was not modified (Fig. 4B1). Pol II, Mediator subunits as well as JUN and FOS were similarly present at the CCND1 promoter (Fig. 4B2 and B6), on which the methylated status of H3K9 histones was not altered in R1295Hson cells when compared with WT (Fig. 4B7).

We also investigated the expression of the RE1 silencing transcription factor (*REST*, MIM: 600571), also known as neuron restrictive silencer factor that suppresses the non-specific expression of neuronal genes in terminally differentiated non-neuronal cells (52–54). *REST* contains two alternative TSSs producing two isoforms (NM_005612.4/NM_001193508.1); one putative AP1-binding site was located near the second TSS (REST2). *REST1* expression that lacks AP1-binding site was not affected by the MED12/p.R1295H mutation (Fig. 4C1). At the *REST1* promoter in which JUN/FOS was absent (Fig. 4C5 and C6), we repeatedly detected low and similar levels of Pol II, MED12 and MED6 in both R1295Hson and WT fibroblasts (Fig. 4C2, C5 and C6). On the contrary, REST2 isoform was up regulated in the patient cell line when compared with WT cells (Fig. 4D1). We detected a higher level of recruitment of Pol II and MED12 in patient cells (Fig. 4D2 and D5). Surprisingly, we also observed a disruption of the JUN/FOS ratio recruited at the *REST2* promoter when compared with what happens in WT cells (Fig. 3D3 and D4). The recruitment of the core Mediator (as visualized by MED6) was not affected by the MED12/p.R1295H mutation (Fig. 4D6). It was difficult to decipher the relative level of H3K9me2 enrichment around both *REST1* and *REST2* promoters due to their close proximity (Fig. 4C7 and D7).

Finally, we were curious to evaluate the expression of Synapsin I (*SYN1*, MIM:313440), a REST-dependent neuronal gene involved in the regulation of neurotransmitter release as well as axonogenesis and synaptogenesis (55). One mutation in

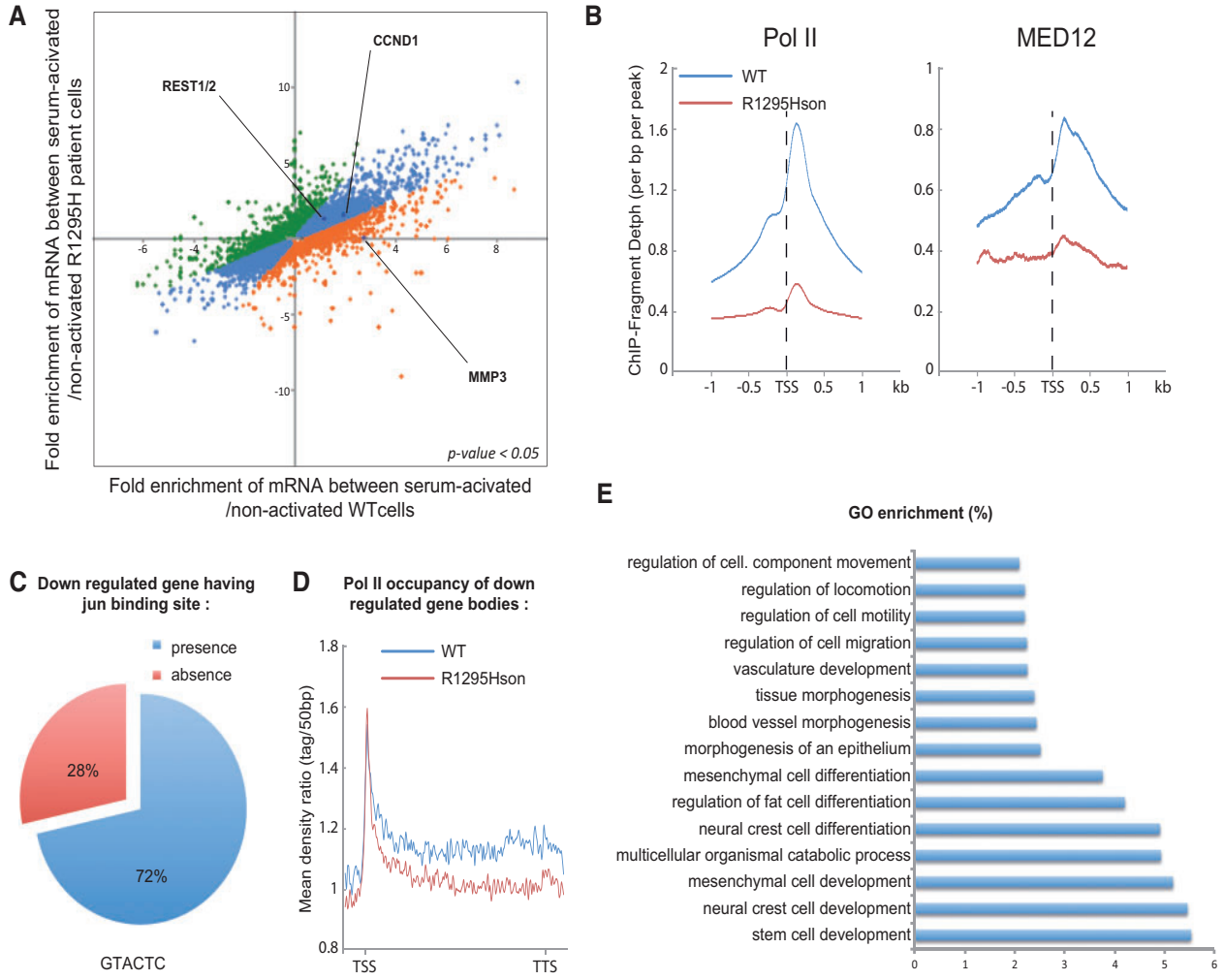


Figure 3. Deregulation of LRGs expression. **(A)** Scatter plot representing fold enrichment of mRNA in both the healthy control (x-axis) and the MED12/p.R1295H patient (y-axis). Fold enrichments have been calculated between serum starved and serum activated cells (3 h, 15%). Only fold enrichments with a P-value below 0.05 are considered. Green spots represent a minimum 2-fold increase and orange spots represent a minimum 2-fold decrease. MMP3, CCND1 and REST1/2 genes are highlighted. **(B)** ChIP-Fragment Detph (per bp per peak) of ChIPseq experiment for Pol II (left) and MED12 (right) on the 58635 TSS of the Refseq database. **(C)** Analysis of the promoter region of down-regulated genes (as defined in A). 3000bp region around TSS was searched for the presence of 'GTACTC' sequence. **(D)** Mean density (tag/50bp) in serum-treated cells for Pol II reads on the down regulated gene bodies. **(E)** Gene ontology performed on the 3550 down regulated gene on Gene Ontology Consortium for Biological processes.

this gene was found to segregate with X-linked disorders and to be associated with aggressive behaviour, macrocephaly and learning problem (56). *SYN1* expression was significantly down regulated in R1295Hson cells (Fig. 4E1). Pol II presence at *SYN1* promoter, was quite similar in both cells (Fig. 4E2) when REST is slightly more present in patient cells (Fig. 4E3). However, it seems that the lower expression of *SYN1* could also be explained by the absence of MED12 and EGR1 recruitment (Fig. 4E4 and E5). We noted that the repressive H3K9me2 marks were not observed around SYN1 promoter in both cells (Fig. 4E7).

All together our data pointed out the cascade effect that occurs on the LRGs when the JUN expression was deregulated.

Specificity of MED12 mutations on IEGs expression

To provide explanations to the large and diverse clinical spectrum covered by MED12-related patients, we enlarge the study to seven different mutations. In addition to the already identified MED12 mutations associated with Opitz-Kaveggia syndrome (p.R961W) (14), Lujan syndrome (p.N1007S) (16) and Ohdo syndrome (p.R1148H and p.S1165P) (18), we ascertain the role of three additional mutations p.R206Q, p.N898D and p.R1295H (Fig. 1B). All these residues are conserved across all MED12 orthologs, from Xenopus to human (Supplementary Material, Fig. S4). These three patients were not diagnosed a priori as having a MED12-related syndrome although they share some clinical features (Table 1). The patient with the p.R206Q substitution only shares the major symptoms such as intellectual disability, tall stature and macrocephaly. The patient with the p.N898D mutation (Supplementary Material, Fig. S2C) exhibited more clinical features such as his facial trait and 'gestalt' quite close to those of FG syndrome (57), as observed for the FG patient bearing p.G958E mutation (15). His mother presently dead was bearing the hemizygous mutation and presented some minor intellectual disability and some facial morphological features. Another patient with the p.R1295H mutation was enrolled with his brother in a cohort of families with XLID and was sharing some features associated with Lujan syndrome, despite it was difficult to arrive at a single diagnosis for both.

To further conduct our investigations, we used Epstein-Barr Virus (EBV)-immortalized lymphoblastoid cells more easily to collect. RT-qPCR showed that all the five mutations did not significantly affect mRNA expression levels of Mediator

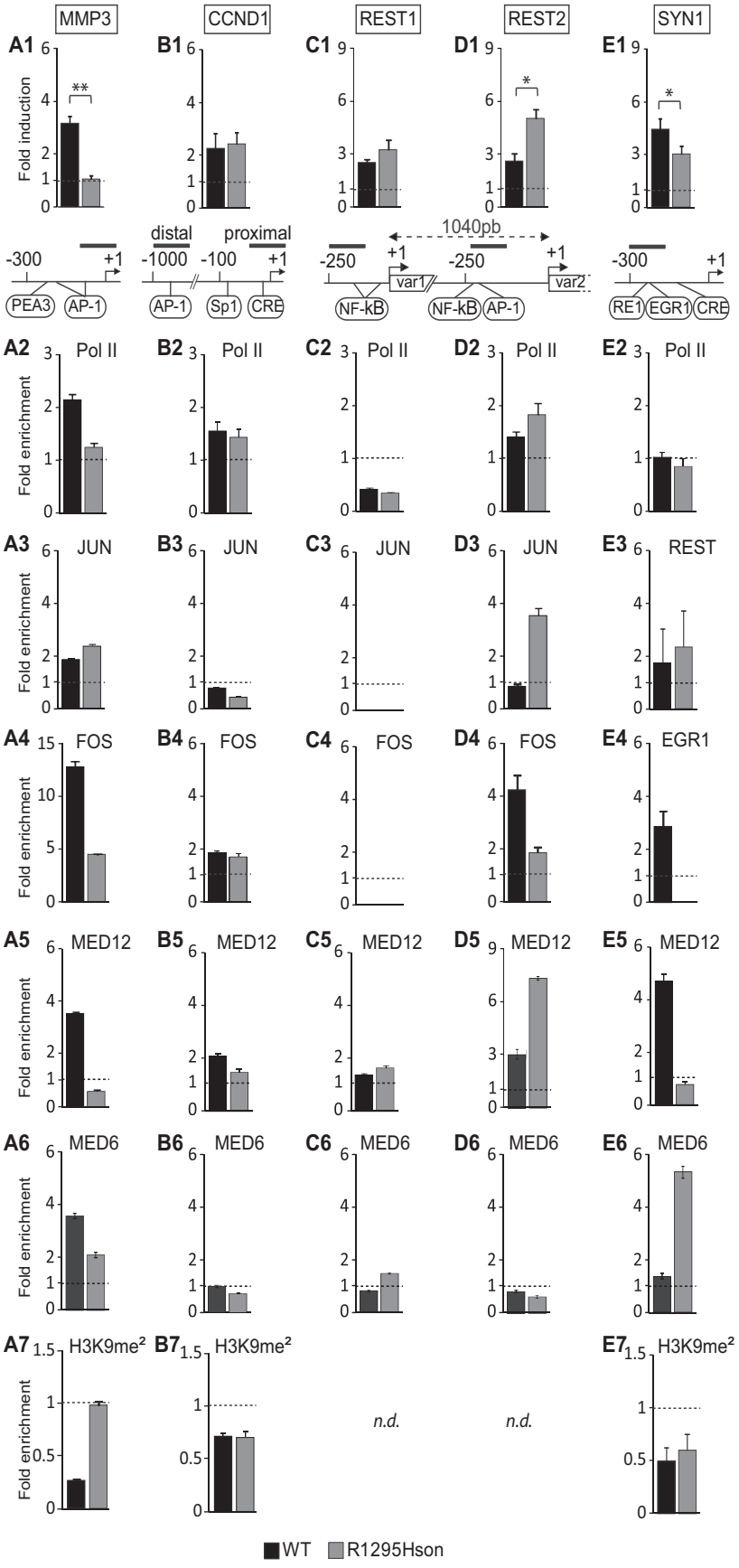


Figure 4. Responses of specific LRGs: MMP3, CCND1, REST and SYN1. Responses of *MMP3* **(A1)**, *CCND1* **(B1)**, REST1 **(C1)**, REST2 **(D1)** and *SYN1* **(E1)** expression 3 h after serum addition to serum-starved WT or MED12/p.R1295H fibroblasts. (*corresponds to $P < 0.05$; **corresponds to $P < 0.01$). Diagrams showing representations of *MMP3*, *CCND1*, *REST1*, *REST2* and *SYN1* promoters show the designed PCR amplicons. ChIP monitor the serum-dependent recruitment of Pol II **(A2, B2, C2, D2, E2)**, JUN **(A3, B3, C3, D3)**, REST **(E3)**, FOS **(A4, B4, C4, D4)**, EGR1 **(E4)**, MED12 **(A5, B5, C5, D5, E5)**, MED6 **(A6, B6, C6, D6, E6)** and H3K9me^2 **(A7, B7, E7)** on the promoter of these gene. Present values are representative of at least three independent experiments.

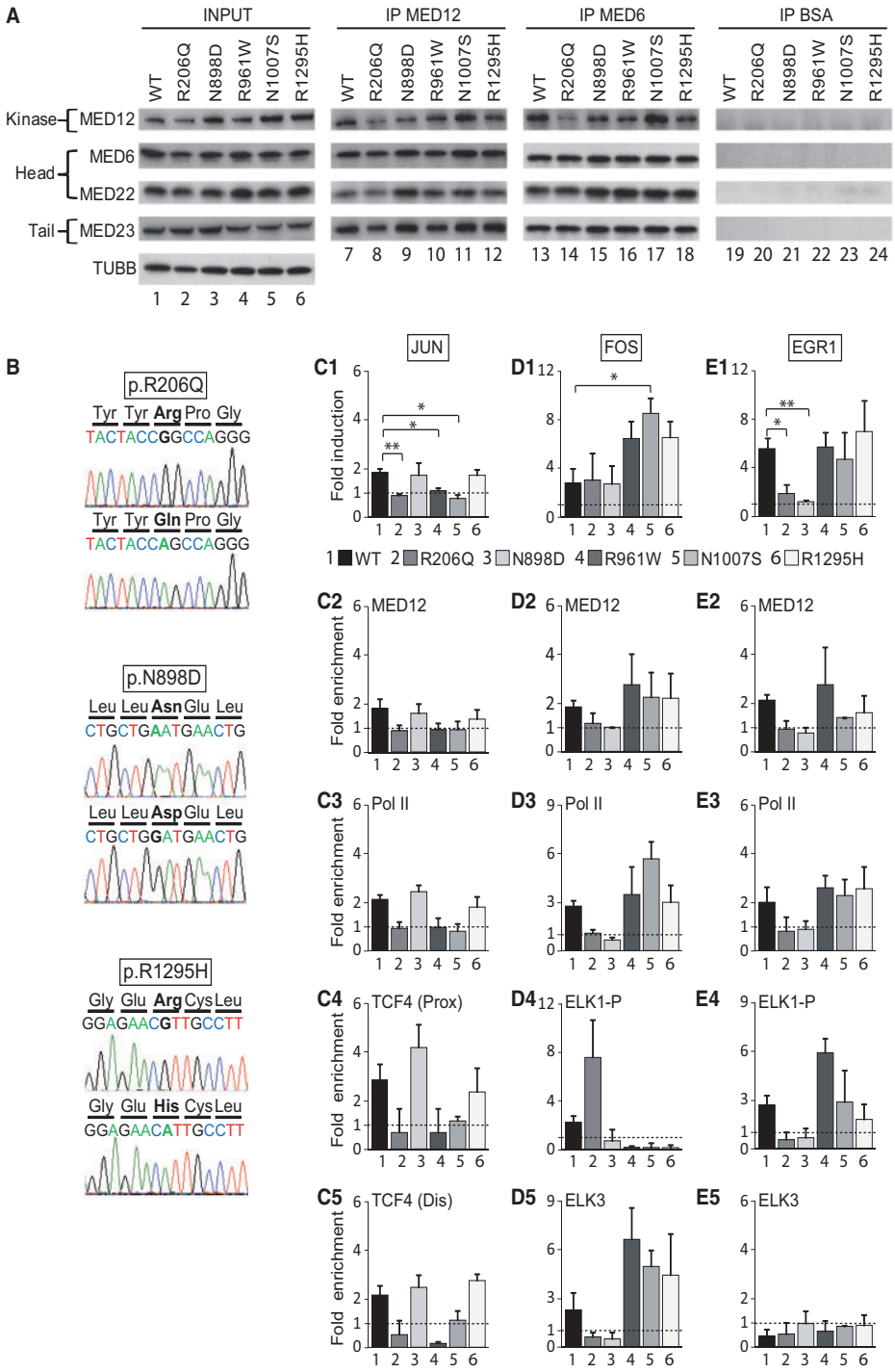


Figure 5. *MED12* mutated lymphoblats. **(A)** Analysis of Mediator composition by IP of MED12, MED6 or a control (BSA) in lymphoblastoid cells. Bound proteins were revealed by Western blot using antibodies against MED6, MED12, MED22, MED23 and TUBB. INPUT corresponds to 20% of the lysate used for IP reactions. **(B)** Electropherograms of the three new mutations (in bold) of affected individuals (bottom) and healthy control (top). **(C1, D1, E1)***JUN*, *FOS* and *EGR1* expression 30min after serum addition to serum-starved lymphoblastoid cell lines derived from healthy subject (WT) or MED12 patient. Results are presented as fold induction, meaning the ratio of treated cells relative to non-treated cells, after normalization against GAPDH. (*corresponds to $P < 0.05$; **corresponds to $P < 0.01$). ChIP monitor the serum-dependent recruitment of MED12 **(C2, D2, E2)**, Pol II **(C3, D3, E3)**, TCF4 **(C4, C5)**, ELK1-P **(D4, E4)** or ELK3 **(D5, E5)** on the IEGs promoter in WT or patient cells. Values are expressed as fold enrichment (ratio of the INPUT percentage between treated and non-treated cells). Present values are representative of at least three independent experiments.

subunits MED12, CDK8, MED17 and MED23 (Supplementary Material, Fig. S5). Immunoprecipitation (IP) experiments on whole cell extracts using antibodies against MED12 and MED6 subunits revealed, that all the MED12 variants are associated with Mediator, as observed by the presence of the head (MED6, MED22) and tail (MED23) modules at 300 mM KCl concentration (Fig. 5A). As a further control, the missense mutations were independently inserted into a plasmid expressing MED12 fused to a B10-tag and then transfected into HeLa cells. We observed that MED16 and MED23 co-precipitated with the mutated B10-MED12 in a manner similar to tagged-WT protein (Supplementary Material, Fig. S6). We next investigated the expression of *JUN*, *FOS* and *EGR1* 30 min after serum induction in patient lymphoblastoid cells. The expression of JUN is down

regulated in R206Q, R961W and N1007S cells as compared with WT, N898D and R1295H cells (Fig. 5C1, lanes 2, 4 and 5). In contrast, *FOS* is up regulated in R961W, N1007S and R1295H cells (Fig. 5D1, lanes 4–6). *EGR1* expression is down regulated in both R206Q and N898D cells compared with WT, R961W, N1007S and R1295H cells (Fig. 5E1, lanes 2 and 3). The two mutations related to Ohdo syndrome (R1148H and S1165P) also lead to an up regulation of *FOS* and a down-regulation of *EGR1* (Supplementary Material, Fig. S7), illustrating the regulatory role of MED12 in IEGs expression regulation.

ChIP experiments next showed that the recruitment of Pol II and MED12 at their respective promoter parallels the expression level of *JUN*, *FOS* and *EGR1* (Fig. 5C2 and C3, D2 and D3 and E2 and E3). The recruitment of TCF4, at the responsive element of *JUN* promoter is impaired when Pol II and MED12 binding is defective (Fig. 5C4 and C5, lanes 2, 4 and 5). In R961W, N1007S and R1295H cells, the up-regulation of *FOS* parallel the reduced binding of ELK1-P which is compensated by an increase binding of its paralog ELK3 on the SRE element (Fig. 5D4–D5, lanes 4–6). We also observed a higher binding of ELK1-P and a lower binding of ELK3 at the *FOS* promoter in R206Q cells (Fig. 5D4 and D5, lane 2). In N898D cells, neither ELK1-P nor ELK3 were detected at SRE (Fig. 5D4 and D5, lane 3). However, in these two cell lines, *FOS* seems to be normally expressed (Fig. 5D1, lanes 2 and 3). In both R206Q and N898D cells, where *EGR1* is down regulated, neither ELK1-P nor ELK3 are recruited at SRE (Fig. 5E4 and E5, lanes 2 and 3), while in the R961W, N1007S and R1295H cells as well as in WT cells, we observed ELK1-P recruitment (lanes 1 and 4–6).

Knowing that each gene can be under the control of different stimuli, we also investigated the behaviour of *JUN* and *FOS* when cells were submitted to another cellular stress such as UV irradiation. Contrary to what was observed under serum mitogen stress, *JUN* is not down regulated but even seems to be up regulated in R206Q cells (Supplementary Material, Fig. S8A, lane 2). In UV treated R961W and R1295H cells, *FOS* seemed to be similarly overexpressed compared with what occurred in serum induced cells. However, we observed a much lower stimulation in N1007S cells and a significant increase in R206Q cells (Supplementary Material, Fig. S8B lanes 4, 6 and 5, 2, respectively). Our data show that MED12 patients might be sensitive to other cellular stress. We also investigated the effect of *MED12* mutations on the expression of the *RARβ* (MIM: 180220) gene induced by all trans retinoic acid (t-RA), that mediates cellular signalling during embryonic morphogenesis, cell growth and differentiation (58). Every t-RA treated lymphoblastoid cell lines, except N1007S cells, accumulated *RARβ* mRNA (Supplementary Material, Fig. S8C, lane 5). This demonstrate that JUN and FOS are sensitive to any kind of stress in general.

All together, the above data significantly showed that each MED12 mutations might lead to a specific pattern of IEGs expression, resulting from impairment in the formation of the transactivation complex. This will undoubtedly influence the *LRGs* in a different manner.

Discussion

In an attempt to dissect the complicated molecular mechanisms underlying the control of gene expression, we decided to deeper our knowledge on the large multi-subunit Mediator complex. In human, the central role of Mediator was highlighted by the discovery of mutations in some of its subunits leading to severe genetic disorders (8–10,15,16,18). Among MED subunits, MED12 is the part of the Kinase module that plays a crucial role in gene expression regulation (4,5,31). Our work first revealed that *MED12* knock down in human neuroblastoma cells significantly reduced the expression of *JUN*, *FOS* and *EGR1*, three genes involved in brain development and plasticity (Fig. 1D) (12,39,59,60).

We focused on several missense MED12 mutations responsible for broad and diverse clinical features. Although patients share some similar characteristics such as intellectual disabilities, several were not diagnosed a priori as having a MED12 associated syndrome, such as Lujan, Opitz-Kaveggia or Ohdo syndromes (Table 1).

As Mediator serves as a link between gene specific-binding factors and the basal transcription machinery, we investigated the activity of several genes in cells carrying MED12 mutations. We then observed that the MED12 mutations led to the abnormal and specific deregulation of *IEGs* expression. For example, MED12/p.R206Q,/p.R961W and/p.N1007S disrupt *JUN* expression (Fig. 5C1). In such cases, TCF4 was not properly recruited to its binding element (Fig. 5C5) and was then unable to contact the basal Pol II transcription machinery on proximal promoter (Fig. 5C4). Similarly, MED12/p.R961W,/pN1007S and/p.R1295H mutations up regulated *FOS* expression (Fig. 5D1), through a change in the correct configuration of the transcription machinery at the IEGs (e.g. ELK3 recruitment instead of ELK1, Fig. 5D2–D5). A third example was provided by the MED12/p.R206Q and/p.N898D that down regulated EGR1 expression due to the absence of an ELK proteins in the transactivation complex (Fig. 5E4 and E5).

We also question about the cause that prevent normal Mediator function. Although MED12 mutations were shown not to disrupt the formation of the Mediator complex in solution (Fig. 5A), it is likely that they modify in some way the intrinsic stability of the complex, leading to incomplete binding of the kinase module within the transcription complex. *In vitro* investigations predicted a role for the Kinase module of MED toward several substrates such as TFIIH, Pol II, as well as CDK8 itself in the context of activated gene expression (28). Moreover, it was also shown that the Kinase complex incorporation within MED was essential for CDK8 substrate specificity. When MED12 is required for stable incorporation of CDK8/CCNC into MED, it is also required to activate the CDK8 kinase. Here we found that the kinase module (according to ChIP experiments on MED12 and CDK8) did not properly target the *JUN* promoter (Fig. 2A1 and A2). Consequently, inaccurate expression of *IEGs* coding for DNA-binding factors, will further modify *LRGs* expression through a consecutive sequence of events. Indeed in the R1295Hson cells, the deregulation of *MMP3* and *REST2* expression might be explained by a modification in the JUN/FOS recruitment at their respective promoters (Fig. 4A3–A4). It is likely that MED12/p.R1295H associated defect in bridging DNA binding factor to Pol II machinery could also be responsible at least partially, for such deregulation. We indeed observed some changes in the Kinase(MED12)/Core-Mediator(MED6) ratio at the *MMP3* and/or *SYN1* promoter compared with WT cells (Fig. 4A5, A6 and E5, E6). It is clear that MED12 mutations lead to a modification of the correct configuration of the transcription machinery at the *IEGs* and the *LRGs* promoter, which resulted in the modification of the transcriptional programme in the patient cells. The specificity of individual mutation also depends on the cellular context as shown for the *FOS* expression in MED12/p.R1295H fibroblasts or for the expression of *RARβ* in N898D fibroblasts versus lymphoblastoids cells (Supplementary Material, Fig. S9).

Our study has shown that the MED12 mutations led to the abnormal expression of *IEGs* and resulted in changes into *LRGs* expression. Ultimately, these results suggest that the different phenotypes caused by the different mutations are essentially

the result of the above to varying degrees. The deregulation of *JUN*, regarded as a marker of neural activity and its involvement in partnership with *FOS*, partially explained the neurological phenotypes of the patients as summarized Table 1. Although the gene regulation abnormalities do not precisely explain the clinical phenotypes, we observed significant down-regulation of genes containing AP1-binding site targeted by JUN/FOS complex. Gene Ontology enrichment then shows that most of these genes are involved in development and neurological processes. Further investigations on the expression patterns of certain genes including *MMP3*, *REST2* and *SYN1* in neural cells will help to decipher the specific consequences of MED12 mutation in neurological context. Future studies should help to elucidate the genotype/phenotype relationship of the patients with MED12 mutations.

We noticed some heterozygous mother with missense mutation to share part of the symptoms (e.g. mild intellectual disability). *MED12* is located on chromosome X and one of its two alleles is inactivated in females. We have observed that most of cultivated heterozygous fibroblasts (about 80%) from the mother of the MED12/p.R1295H patient express the mutated allele. This skewed XCI phenomenon might explain why the MED12/p.R1295H and/p.N898D mothers exhibit similar but milder clinical features than their sons. Actually, the partial maternal symptoms could be explained by a stochastic effect occurring at two different times: at the early embryonic stage of the X-inactivation process and lately during the embryonic organ specification due to a random redistribution of founder cells. This random process could be systematically skewed in some cases due to preferential inactivation of the wild allele. We could not discard this hypothesis since MED12 was implicated in the activation of non-coding RNA. In this way, mutated *MED12* might have a possibility to disturb the balance of XIST and/or TSIX non-coding RNAs transcription which control the X-inactivation process (61). In fact, a defect of XIST in tiny ring X chromosome patients is already know to result in developmental and cognitive disorders (62).

A number of human diseases with *IEG* expression affecting brain development and plasticity are caused by mutations in genes encoding for MED subunits or proteins interacting with MED, such as nucleotide excision repair factors or chromatin remodellers. Mutations of MED13L, MED17, MED23 or MED25 are responsible for infantile cerebral and cerebellar atrophy, non-syndromic mental retardation and Charcot-Marie Tooth disease (8–11). Mutations in TCF4, which interact with the Mediator, result in the Pitt-Hopkins syndrome (42). Examples are also provided by xeroderma pigmentosum (XP, MIM:278730) patients that have progressive neurological degeneration as well as both trichothiodystrophy (MIM:610675) and Cockayne syndrome (MIM:610651) patients that are mentally retarded (63,64). It is likely that more and more mutations in MED and its transcription partners will be detected in patients with cognitive deficiencies as sequencing becomes more commonly used. Therefore, one could speculate that these patients should be classified on a genomic basis rather than distinct syndromes.

Material and Methods

Cell culture

Human lymphoblastoïdes cells were generated by EBV transformation of the peripheral blood lymphocytes isolated from patients. Human primary fibroblasts were isolated from patient's skin. Informed consent was obtained from all subjects before sample collection.

The cs of R1295Hmother fibroblast was done by successive dilution and confirm by Sanger sequencing of total mRNA.

CRISPR/Cas9 system was used to Knock-Down MED12 in neuroblastoma LAN-1 cell line. A targeting vector was constructed on the bases of the pPGKNEO-DTA (addgene no. 13443). 5′ arm spanning intron 2/exon 3 (1 kb, HindIII/AgeI) and 3′ arm spanning exon 5/intron 5 (960 bp, NheI/XhoI) were PCR amplified from genomic DNA of LAN-1 neuroblastoma cells and sequentially cloned into the first pPGKNEO-DTA vector using respected enzymes. Furthermore a GFP was amplified from pEGFPC1 vector and inserted via restriction digest (AgeI/NcoI) so that exon 3 and GFP will be translated in frame. A guiding RNA targeting exon 3/intron 3 junction was designed and cloned into p459 vector (addgene no. 48139). All constructs were verified by sequencing. All together the targeting vector and CRISPR/Cas9 system targeting the end of exon 3 were transfected into LAN-1 neuroblastoma cells using Amaxa nucleofection protocol. Cells were briefly selected with puromycin for first 2 days and with neomycin for next two weeks. Clones were first pre-screened for GFP expression and further verified by genomic PCR for absence of exon 4 and Med12-specific integration of targeting vector.

Med12 3 arm fw CCAAGCTAGCTAACTCCTAACACCAGGTGTACTGC
Med12 3 arm rv CCCTCGAGCAAGCTTACACAGCATGCCCTACTCTCTACC
Med12 5 arm fw TCAGATCTCGAGCTCAAGCTTGGTCAGCCTAGGAGGAGGCACTG
Med12 5 arm rvATGGTGGCGACCGGTGGGGCTAGTTGCGTGAGTGGCTTGG
px4_3ex_Med12_fwcaccgAAAAGGTAAGGTACTGTTTC
px4_3ex_Med12_rvaaacGAAACAGTACCTTACCTTTTc

All the cells were cultured in appropriate medium and maintained at 37 °C in 5% CO_2 environment.

For serum inductions, cells were incubated in red phenol-free medium without serum for 24 h before treatment by addition of serum (15% final concentration) directly into the medium. For retinoic acid inductions, cells were incubated in red phenol-free medium containing charcoal treated Fetal Calf Serum and antibiotics before treatment with 10 μM tRA (Biomol) into the same medium. For UV inductions, cells were first PBS wash then 4 millions of lymphoblastoid cells in 2 ml of PBS were spreading on a 10-cm dish. Cells were exposed to UV irradiation (20 J/m^2) and medium added back.

RNA extraction, reverse transcription and real-time qPCR

Total RNA was extracted with RNeasy mini kit (QIAgen) and reverse transcribed with poly-dT primer using Superscript II reverse transcriptase (Invitrogen). Real-time quantitative PCR was carried out on the Lightcycler 480 using SYBR Green I Master (Roche). The primer sequences are available in Supplementary Material. mRNA levels represent the ratio between values obtained from treated cells compared with untreated cells normalized against the housekeeping Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA.

Co-immunoprecipitation assays

Cells were harvest in lysis buffer (50mM Tris HCl pH7.5, 150 mM NaCl, 1mM EDTA, 1% Triton) with protease inhibitor cocktail

(PIC). 200 μg of whole cell extract were incubated with protein G magnetic bead (dynabead, invitrogen) and 2 μg of antibodies. After washes at 300 mM salt, beads were boiled in Laemmli buffer and proteins were separated by SDS-PAGE, transferred to nitrocellulose membrane and probed with antibodies.

Chromatin immunoprecipitation

After treatment, cells were crosslinked at room temperature for 15 min with 1% formaldehyde. Cells were lysed in shearing buffer (50 mM Tris-HCl pH8, 10 mM EDTA, 1% SDS) with PIC at 4 °C for 15 min. Nucleus were pelleted and resuspended in sonication buffer (50 mM Hepes pH7.6, 140 mM NaCl, 1 mM EDTA, 1% Triton X100, 0.1% Na-deoxycholate, 0.1% SDS) with PIC. Chromatin was shared using 800R sonicator (Qsonica). Samples were immunoprecipitated with antibodies at 4 °C overnight followed by addition of a pre-blocked mix of protein G- and A-Sepharose beads (GE Healthcare Life Sciences) for 3 h at 4 °C. Bound complexes were sequentially washed with sonication buffer, high salt buffer (sonication buffer at 300 mM NaCl) and TE buffer (10 mM Tris, 1mM EDTA). Protein-DNA complexes were eluted, and DNA fragments were purified using QIAquick PCR purification kit (QIAGEN). qPCR was performed as described earlier using sets of primers available in Supplementary Material. All the results are presented as 'fold recruitment' and represent the ratio of input percentage between treated and non-treated cells.

Plasmids

MED12 expression plasmid was obtained using the Gateway Invitrogen cloning method. WT MED12 was amplified from a cDNA bank of control lymphoblastoid cells and cloned into pDONR207 (Invitrogen) using standard BP reaction. The cloned sequence was then transferred by LR reaction into pSG_5 puro B10 tag vector (N-terminal fusion of the epitope B of the human estrogen receptor). This vector was constructed by inserting the attL1 and attL2 Gateway linkers (Invitrogen) into the pSG_5 vector backbone. PCR-based mutagenesis was performed using Phusion High-Fidelity DNA Polymerase (New England BioLabs, inc.) with primer bearing a point mutation for the amino changes R206Q, N898D, R961W, N1007S and R1295H.

MED12 expression plasmid was transfected using Jet PEI (Polyplus) in HeLa cells, 48 h before the experiments.

Antibodies

For Western blot, the following antibodies were used: MED6 (santacruz, sc-9434), MED12 (Béthyl, A300-774A), MED22 (santacruz, sc-393738), MED23 (BD Pharmingen, 550429), CCNC (santacruz, sc-1061) and β-tubulin (TUBB) (millipore, MAB3408).

For immunoprecipitation, the following antibodies were used: monoclonal antibodies against B10 (IGBMC), RNA Pol II (IGBMC, 1BP 7C2), ELK3 (IGBMC, 5NE 2F3A2) and polyclonal antibodies against BSA (santacruz, sc50528), MED6 (santacruz, sc-9434), MED12 (Béthyl, A300-774A), TFIIB (santacruz, sc-225), TCF4 (santacruz, sc13027), ELK1-P (santacruz, sc8406), JUN (santacruz, sc-45 or sc 822), FOS (santacruz, sc-7202), REST (Béthyl, A300-539A) and H3K9ac (cell signalling, no. 9671 or no. 9649) and $H3K9me^2$ (cell signalling, no. 9753).

Mutation screening of the MED12 gene

Genomic DNA was prepared from lymphocytes according to standard protocols. Sanger resequencing of 718 X-chromosome genes was conducted as described previously in Tarpey *et al.* (2009) for the proband in family K9338. The proband in K9467, was considered to have a phenotype similar to FG syndrome and he was sequenced for all coding exons of *MED12*. The single male in family 8935 was sequenced for all coding exons in MED12 as part of a general screen of a cohort of males with ID.

Sequences alignment

Sequences alignment was performed with BioEdit software (open source). The following sequence of Med12 has been used: *Homo sapiens* (NP_005111.2), *Mus musculus* (NP_067496.2), *Rattus norvegicus* (NP_001180221.1), *Bos Taurus* (XP_005228076.1), *Pongo abelii* (NP_001124553.2), *Pan troglodytes* (NP_001009019.1), *Danio rerio* (NP_001034550.1), *Xenopus tropicalis* (XP_002934949.2).

RNAseq

Total RNA was extracted with TRI REAGENT (Molecular Research center, inc.) and purified by phenol-chloroform extraction. Libraries was prepared with TruSeq Stranded mRNA Sample Preparation kit following guide instruction and subsequently proceed on an Illumina Hiseq 4000 as single-end 50 base reads following Illumina's instructions. Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Reads were mapped onto the hg38 assembly of the human genome. Reads count was performed with HOMER v4.8.3 (65) and differently expression was estimated with EdgeR. Genome ontology was performed on http://geneontology.org, searching for biological processes.

ChIPseq

DNA fragments analysed by ChIP-seq were prepared trough regular ChIP procedure and purified by phenol-chloroform extraction. ChIP-seq was performed on an Illumina Hiseq 4000 as single-end 50 base reads following Illumina's instructions. Image analysis and base calling were performed using RTA 1.17.20 and CASAVA 1.8.2. Reads were mapped onto the hg38 assembly of the human genome. Quantitative comparisons of the ChIP-seq data were performed using HOMER v4.8.3 (65). As reference coordinates, we used RefSeq coordinates of human genes. Tag densities were collected ±1 kilobases of the TSS or along the gene body.

Clinical reports

Patient K8935 (with mutation MED12/p.R206Q): W.D.A. is a 39½-year old African American male who has been institutionalized in facilities for those with intellectual disability since age 17 years. Psychological testing showed IQ measurements of 52, age 7 years. Vineland social quotient was 52 at age 12 years. Sequencing of the MED12 gene identified c.617G > A, p.R206Q mutation. Clinical findings are given in Table 1.

Patient K9467 (with mutation MED12/p.N898D): A single male is affected in K9467. The mother had a prior spontaneous abortion. The birth weight was 3.4 kg and the length was 50 cm. Imperforate anus and cryptochidism were present and corrected surgically. Development was globally delayed with independent walking and first clear words achieved at age three years. Special education was required. Intellectual performance was mildly impaired with a full scale IQ of 58. At age 6 years, he had a height of 113.5 cm (35th centile) and

occipital frontal circumference (OFC) of 55.3 cm (>97th centile). He has a tall and narrow forehead, downslanting palpebral fissures, small posteriorly rotated ears (4.5 cm length), maxillary dental crowding and prognathism. The thumbs were flat with angulation of the distal phalanges and the great toes were broad. Neurological examination was normal. There was a tendency to frustration and psychological lability, intolerance and a short attention span. At age of nineteen years, he has a height of 176 cm, weight of 61 kg and head circumference of 61 cm (>99th centile). Gross motor function is normal, fine motor function is slightly impaired. IQ has not been reevaluated but clinical appreciation is consistent with 6-years old IQ. Behavioural problems are prominent with marked frustration intolerance. Constipation and easy vomiting still required permanent medical management (Supplementary Material, Fig. S2C and D).

His mother presents with similar but milder dysmorphic features. She suffered from obstinate constipation. She had normal intelligence but some choleric behavioural problems. Unfortunately, she suddenly died of intracranial aneurysm rupture at the age of 47 years.

Patients K9338 (with mutation MED12/p.R1295H): K9338 has two brothers with intellectual disability born to consanguineous parents. They presented with somewhat different phenotypes. One had congenital anal stenosis that was dilated, undescended testes, an umbilical hernia and bilateral inguinal hernias He was very excitable as a child and this transitioned into tantrums and restlessness as a teen. IQ testing, age 11 years, gave a score of 58. The second male was hypotonic, slow to feed and developed slowly, and did not walk until after age 2 years. He also had undescended testes, finger contractures at the PIP joints, hammer toes and a gap between toes 1 and 2. Additional clinical findings are listed in Table 1. Sequencing of 718 genes on the X chromosome identified a c.3884G > A, p.R1295K mutation in one of the males which was subsequently found in his brother.

Patients A.J. (with mutation MED12/p.R1295H from FRANCE). A.J. was the second child of unrelated parents. His two brothers were in good health. He was born at term (40 WG) after an uneventful pregnancy. Birth weight was 3.1 kg. The neonatal course was uncomplicated but he had delayed developmental milestones. At 22 months, it was noted a global hypotony and some morphologic particularities. Ophthalmological and heart examinations, as well as abdominal ultrasound, echocardiography and CT scan findings were unremarkable. He walked at age of 24 months. Speech was slightly delayed. At age 4 years, he had an orchiopexy because his left testis was on the inguinal canal. During childhood, he suffered from repetitive strain injury. From the age of 7, he had school difficulties and was enrolled in a school mainstreaming class. He had a nasal speech and a long and thin habitus. At age 20 years, his height was 188 cm (+2DS), weight was 68.5 (+1.5DS) and OFC was 57 cm (+2DS). Arm span was 184 cm. Thoraco abdominal examination showed a mild scoliosis, pectus excavatum, two surnumemary nipples and a hypotonic abdominal wall. The facial features included a long and hypotonic face, moderate hypertelorism, bilateral ptosis with epicanthus, down slanting palpebral fissures, everted lower eyelids, arched, large and sparse eyebrows, medially flaring, malar hypoplasia, a long nose with high and narrow nasal bridge, thin upper lip, teeth malposition with absence of the upper lateral incisors and first molars. Ears were small and low-set with thick helices. Joint laxity, long digits, without a positive thumb sign and flat feet were present. Audiometry evaluation, EEG, metabolic screen in blood and urine, endocrinologic and cytogenetic screen (karyotype, comparative genomic hybridization (CGH) array) were all normal. Ophthalmologic evaluation showed a left divergent strabismus (Supplementary Material, Fig. S2A and B).

His mother had the same skeletal and morphologic features with a long and hypotonic face, high arched eyebrows, bilateral ptosis, a long filtrum, teeth malposition. She had moderate learning difficulties but none intellectual disability. Her height was 173 cm OFC was 56.5 cm (+1.5DS). Echocardiography showed a mitral valve prolapse requiring simple monitoring. She underwent surgery for left divergent strabismus and suffered from repeated ankle sprains and early arthrosis of the hip. As her son, she had a nasal speech.

Patients' consent. R1295H fibroblasts were obtained at CHU Clermont-Ferrand and provided by C.F. Informed and written consent was obtained from the patient and his mother. N898D cell lines were obtained at CHU Grenobles and provided by C.C. and P.S.J. Informed and written consent was obtained from the patient and his father. S1165P and R1148H cell lines were obtained at Donders Institute for Brain and provided by AdB. Informed consent was obtained from all families as published in Vulto-van Silfhout *et al.* (18). R206Q, R961W, N1007S and R1295H lymphoblastic cell lines were obtained through the X-linked mental retardation (XLMR) study at the Greenwood Genetic Centre and provided by CES. Informed consent was obtained from all families as published for R961W (14), N1007S (16) and presented at the 2014 annual meeting of the American Society of Human Genetics for R206Q and R1295H. XLMR research at the Greenwood Genetic Centre has been approved by the Institutional Review Board at Self Regional Hospital in Greenwood, SC.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

This study is dedicated to all the patients and families, particularly to Antoine J. We are gratefull to Annick Toutain, Laurence Faivre, Han G Brunner and Anneke T Vulto-van Silfhout for helping in collecting patient cells and clinical data. Cindy Skinner coordinated sample collection and sharing of materials. We also acknowledge the IGBMC facilities for their technical expertise. We also are grateful to Frederic Coin, Emanuel Compe, Jamel Chelly, Jean-Louis Mandel for fruitful discussion.

Conflict of Interest statement. All authors disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work.

Funding

This study was supported (to J.M.E.) by an ERC Advanced grant; l'Agence Nationale de la Recherche [no. ANR1 08MIEN-022-03]; l'Association pour la Recherche contre le Cancer; l'Association Nationale des Membres de l'Ordre National du Merite ANMONM; la ligue contre le cancer fellowships (to L.M.D.); Fondation pour la Recherche Médicale (to B.B.), young scientist grant from Inserm (to S.H.); Grant-in-Aid for Scientific Research from MEXT and JSPS Japan (to S.H.); NICHD grant (to C.E.S.) [HD 26202] and a grant from the South Carolina Department of Disabilities and Special Needs.

Web Resources

The URLs for data presented herein are as follows:
OMIM: http://www.omim.org
RefSeq: http://www.ncbi.nlm.nih.gov/RefSeq

References

1. Kornberg, R.D. (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci*, **30**, 235–239.
2. Malik, S. and Roeder, R.G. (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.*, **11**, 761–772.
3. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I. and Young, R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
4. Allen, B.L. and Taatjes, D.J. (2015) The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.*, **16**, 155–166.
5. Ansari, S.A. and Morse, R.H. (2013) Mechanisms of Mediator complex action in transcriptional activation. *Cell. Mol. Life Sci. CMLS*, **70**, 2743–2756.
6. Conaway, R.C. and Conaway, J.W. (2011) Function and regulation of the Mediator complex. *Curr. Opin. Genet. Dev.*, **21**, 225–230.
7. Lee, T.I. and Young, R.A. (2013) Transcriptional regulation and its misregulation in disease. *Cell*, **152**, 1237–1251.
8. Kaufmann, R., Straussberg, R., Mandel, H., Fattal-Valevski, A., Ben-Zeev, B., Naamati, A., Shaag, A., Zenvirt, S., Konen, O., Mimouni-Bloch, A. *et al.* (2010) Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am. J. Hum. Genet.*, **87**, 667–670.
9. Leal, A., Huehne, K., Bauer, F., Sticht, H., Berger, P., Suter, U., Morera, B., Valle, G., Lupski, J.R., Ekici, A. *et al.* (2009) Identification of the variant Ala335Val of MED25 as responsible for CMT2B2: molecular data, functional studies of the SH3 recognition motif and correlation between wild-type MED25 and PMP22 RNA levels in CMT1A animal models. *Neurogenetics*, **10**, 275–287.
10. Hashimoto, S., Boissel, S., Zarhrate, M., Rio, M., Munnich, A., Egly, J.-M. and Colleaux, L. (2011) MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. *Science*, **333**, 1161–1163.
11. Adegbola, A., Musante, L., Callewaert, B., Maciel, P., Hu, H., Isidor, B., Picker-Minh, S., Le Caignec, C., Delle Chiaie, B., Vanakker, O. *et al.* (2015) Redefining the MED13L syndrome. *Eur. J. Hum. Genet.*, **23**, 1308–1317.
12. Berk, A.J. (2012) Yin and yang of mediator function revealed by human mutants. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 19519–19520.
13. Graham, J.M. and Schwartz, C.E. (2013) MED12 related disorders. *Am. J. Med. Genet. A*, **161A**, 2734–2740.
14. Risheg, H., Graham, J.M., Clark, R.D., Rogers, R.C., Opitz, J.M., Moeschler, J.B., Peiffer, A.P., May, M., Joseph, S.M., Jones, J.R. *et al.* (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat. Genet.*, **39**, 451–453.
15. Rump, P., Niessen, R., Verbruggen, K., Brouwer, O., de Raad, M. and Hordijk, R. (2011) A novel mutation in MED12 causes FG syndrome (Opitz-Kaveggia syndrome). *Clin. Genet.*, **79**, 183–188.
16. Schwartz, C.E., Tarpey, P.S., Lubs, H.A., Verloes, A., May, M.M., Risheg, H., Friez, M.J., Futreal, P.A., Edkins, S., Teague, J. *et al.* (2007) The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J. Med. Genet.*, **44**, 472–477.
17. Isidor, B., Lefebvre, T., Le Vaillant, C., Caillaud, G., Faivre, L., Jossic, F., Joubert, M., Winer, N., Le Caignec, C., Borck, G. *et al.* (2014) Blepharophimosis, short humeri, developmental delay and hirschsprung disease: expanding the phenotypic spectrum of MED12 mutations. *Am. J. Med. Genet. A*, **164A**, 1821–1825.
18. Vulto-van Silfhout, A.T., de Vries, B.B.A., van Bon, B.W.M., Hoischen, A., Ruiterkamp-Versteeg, M., Gilissen, C., Gao, F., van Zwam, M., Harteveld, C.L., van Essen, A.J. *et al.* (2013) Mutations in MED12 cause X-linked Ohdo syndrome. *Am. J. Hum. Genet.*, **92**, 401–406.
19. Lesca, G., Moizard, M.-P., Bussy, G., Boggio, D., Hu, H., Haas, S.A., Ropers, H.-H., Kalscheuer, V.M., Des Portes, V., Labalme, A. *et al.* (2013) Clinical and neurocognitive characterization of a family with a novel MED12 gene frameshift mutation. *Am. J. Med. Genet. A*, **161A**, 3063–3071.
20. Fieremans, N., Van Esch, H., Holvoet, M., Van Goethem, G., Devriendt, K., Rosello, M., Mayo, S., Martinez, F., Jhangiani, S., Muzny, D.M. *et al.* (2016) Identification of intellectual disability genes in female patients with a Skewed X-Inactivation pattern. *Hum. Mutat.*, **37**, 804–811.
21. Prescott, T.E., Kulseth, M.A., Heimdal, K.R., Stadheim, B., Hopp, E., Gambin, T., Coban Akdemir, Z.H., Jhangiani, S.N., Muzny, D.M., Gibbs, R.A. *et al.* (2016) Two male sibs with severe micrognathia and a missense variant in MED12. *Eur. J. Med. Genet.*, **59**, 367–372.
22. Prontera, P., Ottaviani, V., Rogaia, D., Isidori, I., Mencarelli, A., Malerba, N., Cocciadiferro, D., Rolph, P., Stangoni, G., Vulto-van Silfhout, A. *et al.* (2016) A novel MED12 mutation: Evidence for a fourth phenotype. *Am. J. Med. Genet. A*, **170**, 2377–2382.
23. Tzschach, A., Grasshoff, U., Beck-Woedl, S., Dufke, C., Bauer, C., Kehrer, M., Evers, C., Moog, U., Oehl-Jaschkowitz, B., Di Donato, N. *et al.* (2015) Next-generation sequencing in X-linked intellectual disability. *Eur. J. Hum. Genet.*, **23**, 1513–1518.
24. Yamamoto, T. and Shimojima, K. (2015) A novel MED12 mutation associated with non-specific X-linked intellectual disability. *Hum. Genome Var.*, **2**, 15018.
25. Langley, K.G., Brown, J., Gerber, R.J., Fox, J., Friez, M.J., Lyons, M. and Schrier Vergano, S.A. (2015) Beyond Ohdo syndrome: a familial missense mutation broadens the MED12 spectrum. *Am. J. Med. Genet. A*, **167A**, 3180–3185.
26. Bouazzi, H., Lesca, G., Trujillo, C., Alwasiyah, M.K. and Munnich, A. (2015) Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His)]. *Clin. Case Rep.*, **3**, 604–609.
27. Liu, Y., Ranish, J.A., Aebersold, R. and Hahn, S. (2001) Yeast nuclear extract contains two major forms of RNA polymerase II mediator complexes. *J. Biol. Chem.*, **276**, 7169–7175.
28. Knuesel, M.T., Meyer, K.D., Donner, A.J., Espinosa, J.M. and Taatjes, D.J. (2009) The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of mediator. *Mol. Cell. Biol.*, **29**, 650–661.
29. Nemet, J., Jelicic, B., Rubelj, I. and Sopta, M. (2014) The two faces of Cdk8, a positive/negative regulator of transcription. *Biochimie*, **97**, 22–27.
30. Hengartner, C.J., Myer, V.E., Liao, S.M., Wilson, C.J., Koh, S.S. and Young, R.A. (1998) Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol. Cell*, **2**, 43–53.
31. Akoulitchev, S., Chuikov, S. and Reinberg, D. (2000) TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature*, **407**, 102–106.

32. Kim, S., Xu, X., Hecht, A. and Boyer, T.G. (2006) Mediator is a transducer of Wnt/beta-catenin signaling. *J. Biol. Chem.*, **281**, 14066–14075.
33. Moghal, N. and Sternberg, P.W. (2003) A component of the transcriptional mediator complex inhibits RAS-dependent vulval fate specification in C. elegans. *Development*, **130**, 57–69.
34. Rau, M.J., Fischer, S. and Neumann, C.J. (2006) Zebrafish Trap230/Med12 is required as a coactivator for Sox9-dependent neural crest, cartilage and ear development. *Dev. Biol.*, **296**, 83–93.
35. Treisman, J. (2001) Drosophila homologues of the transcriptional coactivation complex subunits TRAP240 and TRAP230 are required for identical processes in eye-antennal disc development. *Dev. Camb. Engl.*, **128**, 603–615.
36. Yoda, A., Kouike, H., Okano, H. and Sawa, H. (2005) Components of the transcriptional Mediator complex are required for asymmetric cell division in C. elegans. *Development*, **132**, 1885–1893.
37. Zhou, X., Khan, S.G., Tamura, D., Ueda, T., Boyle, J., Compe, E., Egly, J.-M., DiGiovanna, J.J. and Kraemer, K.H. (2013) Abnormal XPD-induced nuclear receptor transactivation in DNA repair disorders: trichothiodystrophy and xeroderma pigmentosum. *Eur. J. Hum. Genet.*, **21**, 831–837.
38. Callier, P., Aral, B., Hanna, N., Lambert, S., Dindy, H., Ragon, C., Payet, M., Collod-Beroud, G., Carmignac, V., Delrue, M.A. *et al.* (2013) Systematic molecular and cytogenetic screening of 100 patients with marfanoid syndromes and intellectual disability. *Clin. Genet.*, **84**, 507–521.
39. Pérez-Cadahía, B., Drobic, B. and Davie, J.R. (2011) Activation and function of immediate-early genes in the nervous system. *Biochem. Cell Biol. Biochim. Biol. Cell*, **89**, 61–73.
40. Forrest, M.P., Waite, A.J., Martin-Rendon, E. and Blake, D.J. (2013) Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PloS One*, **8**, e73169.
41. Nateri, A.S., Spencer-Dene, B. and Behrens, A. (2005) Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. *Nature*, **437**, 281–285.
42. Peippo, M. and Ignatius, J. (2012) Pitt-Hopkins Syndrome. *Mol. Syndromol.*, **2**, 171–180.
43. Besnard, A., Galan-Rodriguez, B., Vanhoutte, P. and Caboche, J. (2011) Elk-1 a transcription factor with multiple facets in the brain. *Front. Neurosci.*, **5**,
44. Buchwalter, G., Gross, C. and Wasylyk, B. (2004) Ets ternary complex transcription factors. *Gene*, **324**, 1–14.
45. Gille, H., Kortenjann, M., Thomae, O., Moomaw, C., Slaughter, C., Cobb, M.H. and Shaw, P.E. (1995) ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and transactivation. *embo J.*, **14**, 951–962.
46. Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.*, **21**, 381–395.
47. Li, B., Carey, M. and Workman, J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707–719.
48. Hess, J., Angel, P. and Schorpp-Kistner, M. (2004) AP-1 subunits: quarrel and harmony among siblings. *J. Cell Sci.*, **117**, 5965–5973.
49. Meighan, S.E., Meighan, P.C., Choudhury, P., Davis, C.J., Olson, M.L., Zornes, P.A., Wright, J.W. and Harding, J.W. (2006) Effects of extracellular matrix-degrading proteases matrix metalloproteinases 3 and 9 on spatial learning and synaptic plasticity. *J. Neurochem.*, **96**, 1227–1241.
50. Pauly, T., Ratliff, M., Pietrowski, E., Neugebauer, R., Schlicksupp, A., Kirsch, J. and Kuhse, J. (2008) Activity-dependent shedding of the NMDA receptor glycine binding site by matrix metalloproteinase 3: a PUTATIVE mechanism of postsynaptic plasticity. *PloS One*, **3**, e2681.
51. Baldin, V., Lukas, J., Marcote, M.J., Pagano, M. and Draetta, G. (1993) Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev.*, **7**, 812–821.
52. Chong, J.A., Tapia-Ramírez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuller, Y.M., Frohman, M.A., Kraner, S.D. and Mandel, G. (1995) REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell*, **80**, 949–957.
53. Roopra, A., Qazi, R., Schoenike, B., Daley, T.J. and Morrison, J.F. (2004) Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. *Mol. Cell*, **14**, 727–738.
54. Schoenherr, C.J. and Anderson, D.J. (1995) The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science*, **267**, 1360–1363.
55. Cesca, F., Baldelli, P., Valtorta, F. and Benfenati, F. (2010) The synapsins: key actors of synapse function and plasticity. *Prog. Neurobiol.*, **91**, 313–348.
56. Garcia, C.C., Blair, H.J., Seager, M., Coulthard, A., Tennant, S., Buddles, M., Curtis, A. and Goodship, J.A. (2004) Identification of a mutation in synapsin I, a synaptic vesicle protein, in a family with epilepsy. *J. Med. Genet.*, **41**, 183–186.
57. Clark, R.D., Graham, J.M., Friez, M.J., Hoo, J.J., Jones, K.L., McKeown, C., Moeschler, J.B., Raymond, F.L., Rogers, R.C., Schwartz, C.E. *et al.* (2009) FG syndrome, an X-linked multiple congenital anomaly syndrome: the clinical phenotype and an algorithm for diagnostic testing. *Genet. Med. Off. J. Am. Coll. Med. Genet.*, **11**, 769–775.
58. Al Tanoury, Z., Piskunov, A. and Rochette-Egly, C. (2013) Vitamin A and retinoid signaling: genomic and nongenomic effects. *J. Lipid Res.*, **54**, 1761–1775.
59. Flanagan, P.M., Kelleher, R.J., Sayre, M.H., Tschochner, H. and Kornberg, R.D. (1991) A mediator required for activation of RNA polymerase II transcription in vitro. *Nature*, **350**, 436–438.
60. Koleske, A.J., Buratowski, S., Nonet, M. and Young, R.A. (1992) A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. *Cell*, **69**, 883–894.
61. Masui, O., Bonnet, I., Le Baccon, P., Brito, I., Pollex, T., Murphy, N., Hupé, P., Barillot, E., Belmont, A.S. and Heard, E. (2011) Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. *Cell*, **145**, 447–458.
62. Migeon, B.R., Luo, S., Stasiowski, B.A., Jani, M., Axelman, J., Van Dyke, D.L., Weiss, L., Jacobs, P.A., Yang-Feng, T.L. and Wiley, J.E. (1993) Deficient transcription of XIST from tiny ring X chromosomes in females with severe phenotypes. *Proc. Natl. Acad. Sci. U. S. A.*, **90**, 12025–12029.
63. Brooks, P.J. (2013) Blinded by the UV light: How the focus on transcription-coupled NER has distracted from understanding the mechanisms of Cockayne syndrome neurologic disease. *DNA Repair*, **12**, 656–671.
64. Lehmann, A.R. (2003) DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie*, **85**, 1101–1111.
65. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H. and Glass, C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.

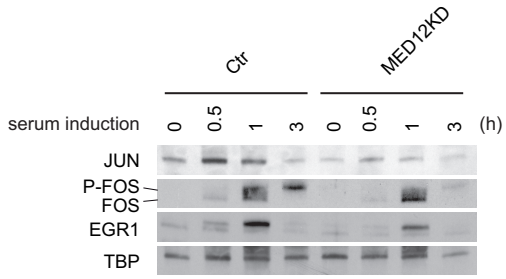


Figure S1: Effect of MED12 knocking down. WT and MED12-KD neuroblastoma whole cell extracts were resolved by SDS-PAGE and processed by Western Blot analysis using antibodies against JUN, FOS and EGR1.

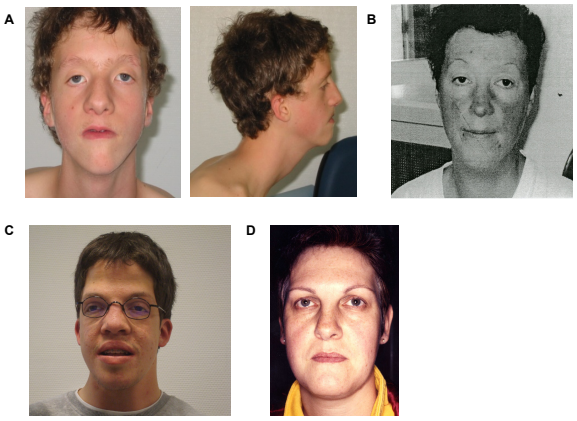


Figure S2: Photographs of patients with MED12 mutations: **(A)** Patient A.J. (p.R1295H/hemizygote) and **(B)** his mother (p.R1295H/heterozygote). **(C)** Patient T.R. in family K9467 (p.N898D/hemizygote) and **(D)** his mother (p.N898D/heterozygote).

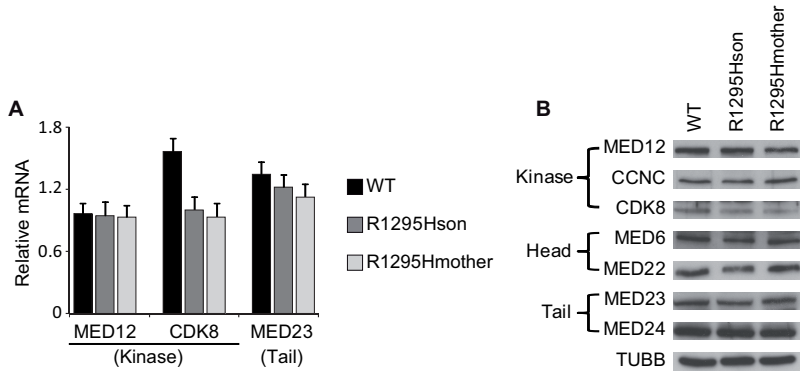


Figure S3: Effect of the MED12/R1295H mutation on Mediator gene expression in fibroblasts: **(A)** Basal expression of *MED12*, *CDK8* and *MED23* mRNA in normal (WT) and R1295H fibroblast cells. The values were normalized relatively to *GAPDH* gene expression. Error bars represent the standard deviation of at least three independent experiments. (* corresponds to p<0,05; ** corresponds to p<0,01). **(B)** Western blot of 40µg of whole cell extract revealed with antibodies targeting MED12, CCNC, CDK8, MED6, MED22, MED23, MED24 and β-tubulin (TUBB).

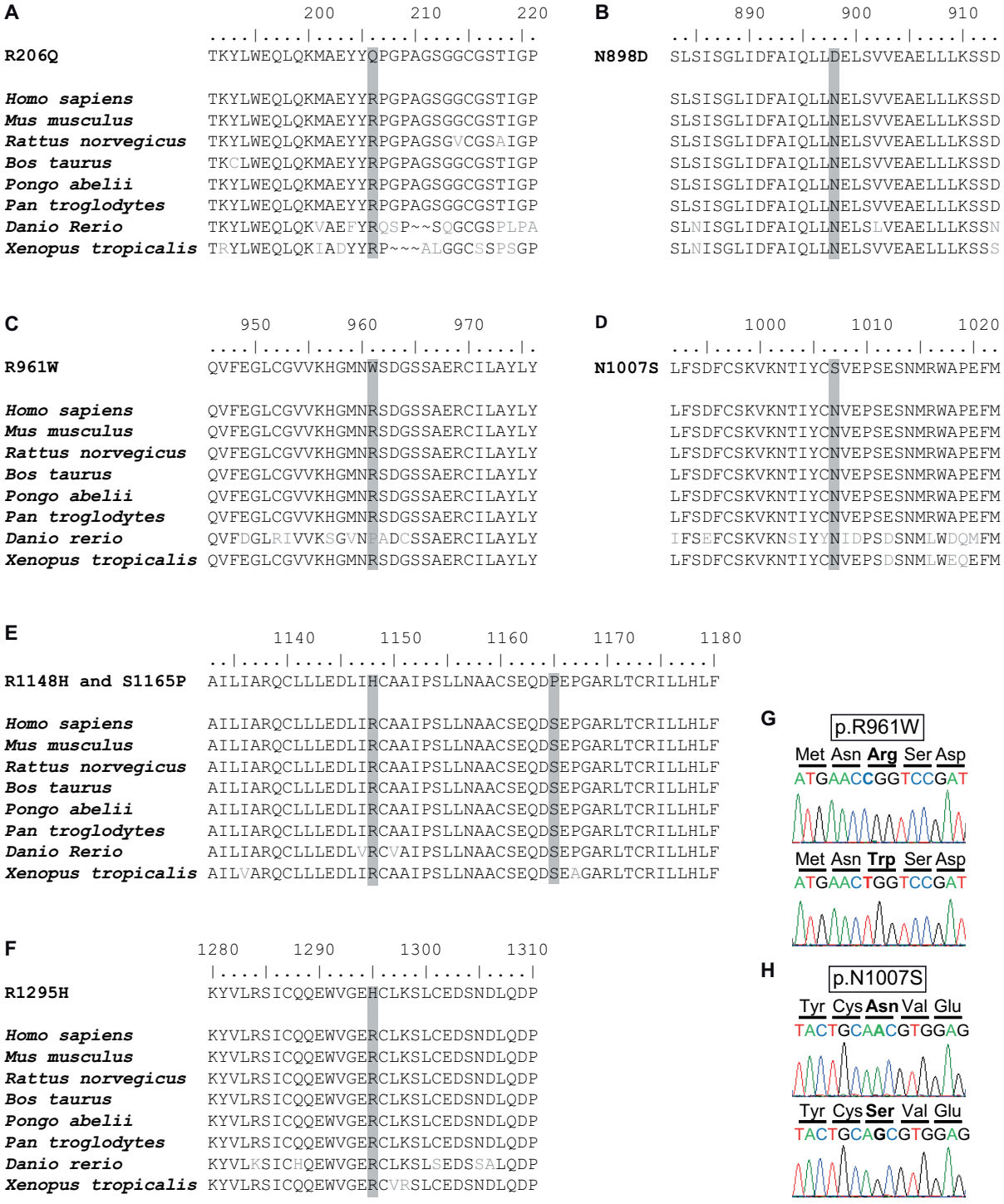


Figure S4: MED12 sequence conservation : Amino acid conservation between species around the residues R206 **(A)**, N898 **(B)**, R961 **(C)**, N1007 **(D)**, R1148 and S1165 **(E)** and R1295 **(F)** in MED12 (mutated residues are highlighted in grey). Electropherograms showing the R961W **(G)** or N1007S **(H)** mutation (in bold) in MED12 sequence of affected individuals (bottom) and healthy control (top).

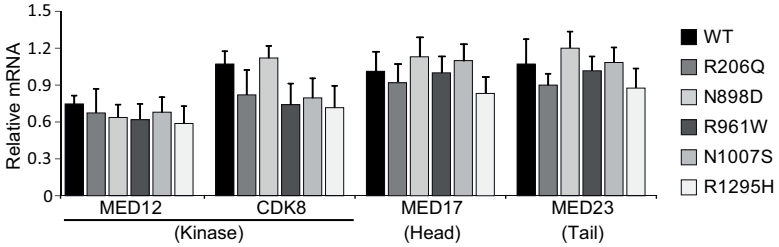


Figure S5: Expression of *MED12*, *CDK8, MED17* and *MED23* genes in normal (WT) and mutant lymphoblastoïd cells. The values were normalized relatively to *GAPDH* gene expression. Error bars represent the standard deviation of at least three independent experiments.

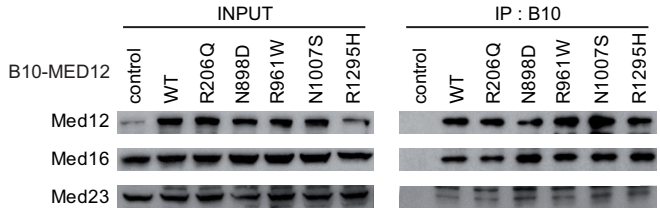


Figure S6: Effect of MED12 mutations on Mediator complex composition in HeLa cells: WT or mutated form of MED12 fused to a B10 tag were ectopically expressed in HeLa cells prior to B10 immunoprecipitation (IP) on whole-cell lysates. Immunoprecipitates were resolved by SDS-PAGE and processed by Western Blot analysis using antibodies against MED12, MED16 and MED23 as indicated. Input cooresponds to 20% of the lysate used for IP reactions.

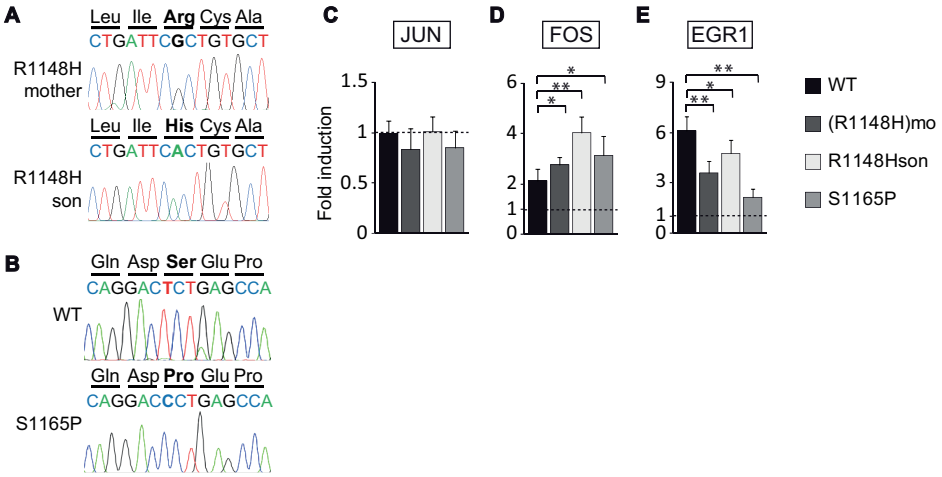


Figure S7: Response of IEGs to serum in MED12 mutated Ohdo lymphoblats. **(A)** Electropherograms of the mutations c.3443 G>A (in bold) of affected son (bottom) and mother control (top). **(B)** Electropherograms of the mutations c.3493 T>C (in bold) of affected patient (bottom) and healthy control (top). Induction of *JUN* **(C)**, *FOS* **(D)** and *EGR1* **(E)** expression 30min after serum addition to serum-starved lymphoblastoid cells from a healthy subject (WT) or MED12 patient. (* corresponds to p<0,05; ** corresponds to p<0,01). Values of mRNA expression represent at least three different experiments. Results are presented as fold induction, meaning the ratio of treated cells relative to non-treated cells, after normalization against GAPDH.

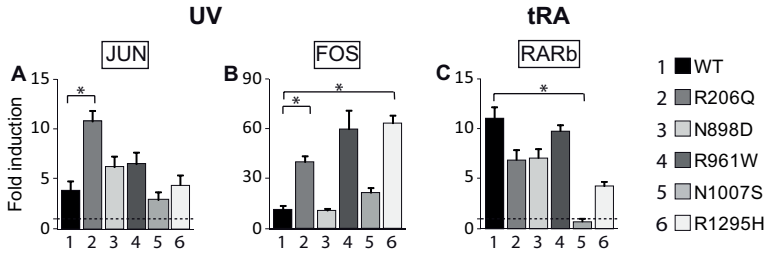


Figure S8: IEG expression after UV irradiation and tRA treatment. Relative mRNA expression of *JUN* **(A)** and *FOS* **(B)** genes after UV-irradiation (20 J/m²) and *RARb* **(C)** gene after all-trans retinoic acid treatment (10µM) in WT and patients lymphoblastoid cells (* corresponds to p<0,05; ** corresponds to p<0,01). The values from three different experiments are presented in fold induction, which means the ratio of treated cells relative to non-treated cells after normalization against GAPDH.

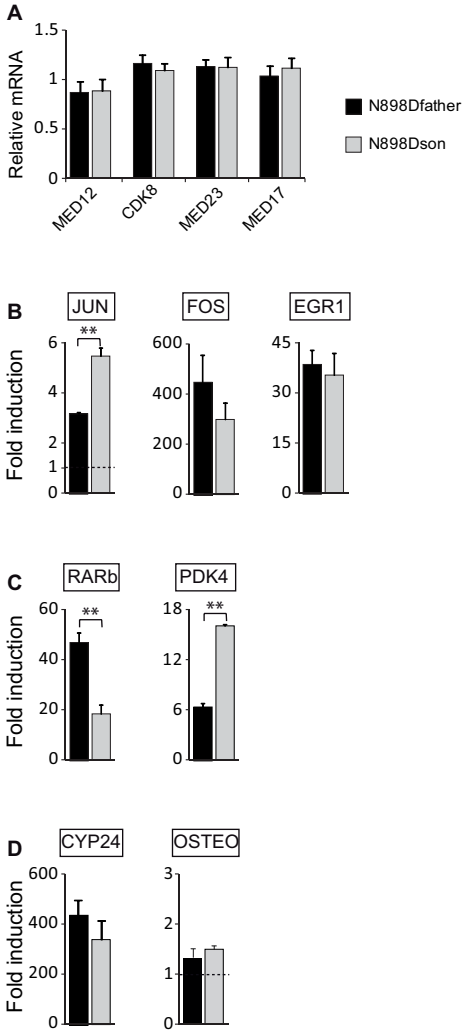


Figure S9: Effect of *MED12* N898D mutation in fibroblasts on Mediator gene expression and IEG responses: **(A)** Basal expression of *MED12*, *CDK8, MED23 and MED17* mRNA in normal (Father) and N898D (Son) fibroblast cells. **(B)** Induction of *JUN*, *FOS* and *EGR1* expression 30min after serum addition to serum-starved cells. **(C)** *RARb* and *PDK4* induction after all-trans retinoic acid treatment. **(D)** *CYP24* and *Osteopontin* induction after vitamin D treatment. (* corresponds to p<0,05; ** corresponds to p<0,01).

mRNA primers	Forward	Reverse
GAPDH	TCGACAGTCAGCCGCATCTTCTTT	ACCAAATCCGTTGACTCCGACCTT
RARb	CCAGCAAGCCTCACATGTTTCCAA	TACACGCTCTGCACCTTTAGCACT
MED12	GCAGAAGAGCATGTCCCTATT	TGGCTGTAGAGGGAGGTAAG
MED17	AGTCCAGTGAAGGGCTTCTGGAAA	CGGCTTGCTAAGCTGTCAATGGTT
MED23	AATGCGCTATGAATGCACGA	GTTTGGAAAGGGACCAGGAGA
CDK8	GGGATCTCTATGTCGGCATGTAG	AAATGACGTTTGGATGCTTAAGC
JUN	AGCGCCTGATAATCCAGTCC	CTGCTCATCTGTCACGTTCTTG
FOS	CAAGCGGAGACAGACCAACT	AGTCAGATCAAGGGAAGCCA
EGR1	AGCACCTGACCGCAGAGTCTTT	CACCAGCACCTTCTCGTTGTT
MMP3	CACTCACAGACCTGACTCGG	AGTCAGGGGGAGGTCCATAG
CCND1	GGTGCTGCGGGCCAT	CTCGCAGACCTCCAGCAT
REST1	CCCGAAACTCCAGCAACAAAG	CCTGGGTGGCCATAACTGTA
REST2	CCGGCTGCGCGAATACAG	CAGGGCCATTCCAATGTTGC
SYN1	TGCTCAGCAGTACAACGTACC	GACACTTGCGATGTCCTGGAA

ChIP primers	Forward	Reverse
JUN prox	CCAGAGAAGAATCTTCTAGG	CCCCAAGGCCTTCCCATTGG
JUN distal	CCGTCTCACTCTCTTGCTCTTC	CAACTGGACAAAATGGCTCTG
FOS pro	GAGCAGTTCCCGTCAATCC	GCATTTCGCAGTTCCTGTCT
EGR1 pro	CTGCCATATTAGGGCTTCCTGCTT	TATTTGAAGGGTCTGGAACGGCAC
MMP3 pro	TCCTGCTGCCATTTGGATGA	GCCTCCTTGTAGGTCCAACC
CCND1 AP1	AACCTTCGGTGGTCTTGTCC	AGCTGAGAAACAGTGATCTCCA
CCND1 pro	ATTCTCTGCCGGGCTTTGAT	TGCAACTTCAACAAAACTCCCC
REST1 pro	TGACCTAAGGGCAGGAGTGA	CTACCAAGCAAGGAGTGCCC
REST2 pro	GGAAGGCGCCGTTGAGT	TGAAGCGCAGAAATCGCTGT
SYN1 pro	CATTCCCCAAATTGCGCATCC	CGAAGGCACTGTCCGCGGTGC

Conclusion

In order to decipher the complex mechanisms that regulate the transcription of protein coding genes by the RNA polymerase II, we need to properly define the structures, interactions, partnership, enzymatic activities and target genes of the multiple components of the transcription machinery. For this purposes, it is useful to combine the fundamental approaches with the clinical one. Indeed, clinical data and biological samples give fruitful insights to protein specific roles that cannot be gained by molecular or biochemical approaches. In the same time, fundamental techniques bring priceless perspectives on the etiology of mutation-associated diseases, to both physicians and patients.

I - The NER factor XPC regulates transcription

Our lab has previously uncovered the fundamental roles of the DNA repair factor XPC in the formation of the Pre-Initiation Complex. XPC allows the recruitment of XPA, RPA, XPG and XPF-ERCC1 to the promoter of the NR activated gene *RARβ2*, what is necessary for its proper transactivation(Le May et al., 2010a). Indeed, they are required for DNA demethylation and active HPTM. XPG and XPF have also been found to be necessary for the formation of the promoter-terminator loop, by promoting CTCF recruitment(Le May et al., 2012).

Our current study unveiled new unexpected roles for XPC in transcription. XPC was found to be recruited at the promoter of a certain set of genes together with Pol II. Using patient cells constitutively depleted for XPC, we showed that XPC was necessary for the proper expression of these genes. XPC is strictly required for the acetylation of H3K9 and the tri-methylation of H3K4 at the promoter of XPC bound genes. It cooperates with the transcription factor E2F1 to recruit the acetyl transferase GCN5, as part of the ATAC complex. It is interesting to note that RPA is also acetylated by GCN5 in the context of nucleotide excision repair, while it is recruited subsequently to XPC(He et al., 2017; Zhao et al., 2017).

Interestingly, relations between XPC and E2F1 have recently been shown in the context of global genome NER. E2F1 notably interact with its partner HR23B. It was also found to localize around UV-induced DNA damages to promote NER factors recruitment through H3K9 acetylation, acting in a non-transcriptional process(Guo et al., 2011; Singh and Dagnino, 2016).

In the present case, E2F1 seems to be responsible for specificity, as most the regulated genes do contain an E2F1 response element. The genes that are down-regulated in absence of XPC are mainly implicate in the regulation of the chromatin structure and present oncogenic and immunologic ontology. We showed here that XPC regulated DAPK1 and it was also shown that XPC can induce interleukin-6(Schreck et al., 2016), while both are implicated in cancer progression. Among the regulated genes, we also founded the *SET1* methyl transferase. Our results suggest that the absence of XPC is also directly responsible for the decrease of H3K4me3 mark at the promoter XPC-regulated genes.

Our work greatly improves the comprehension of the mechanism of action of XPC and enlightens the overlapping pathways where XPC engage in the context of transcription and DNA repair. Therefore, these two fundamental processes of the cell appear more and more intricate.

II - The Mediator subunit MED12 in neurodevelopmental disorders

We published earlier the relation between a mutation in *MED23* and intellectual disability in a large Algerian consanguineous family(Hashimoto et al., 2011). This study unveiled the role of Mediator in the regulation of immediate early genes and the implication of Mediator in brain development and functioning.

By studying several *MED12* mutations, we showed that MED12 is strongly implicated in the regulation of IEG, especially *JUN*(Donnio et al., 2017). What is more, the position of the mutation on *MED12* differentially impacted the *JUN* expression. Considering the roles of IEG in the cognitive defects(Berk, 2012; Pérez-Cadahía et al., 2011), it may be possible that the position of the mutation differentially influence the neurological development, thus explaining the phenotypic differences among patients. The mediator serves as a bridge between the transcription machinery and the specific transcription factors. Mutations that affect one of brick of the bridge undoubtedly perturbed the normal transcription program.

Since the first case, almost twenty *MED12* mutations have been published with closely related symptoms. The affected patients are considered among one of

the three definite syndromes (Ohdo syndrome, FG syndrome and Lujan syndrome) or as non-syndromic ID. They share some phenotypic trait like intellectual disability, neurodevelopmental delay, speech difficulty and congenital digestive issues, defining a “MED12 phenotype”. Therefore, syndrome boundaries are becoming vague for some patients.

Other Mediator subunits have been found to be mutated and to give rise to several form of intellectual disability associated with various phenotypes/syndromes. MED13L can easily be compared with MED12 phenotype. This proximity is not surprising since both MED12 and MED13L belongs to the same Mediator module. Therefore, some authors have proposed the concept of a MED12/M13L clinical spectrum(Caro-Llopis et al., 2016).

In a larger perspective, considering the increasing number of patients with a mutation in one of the Mediator subunit and the relatively close neurodevelopmental issues, the concept of “Mediatorophaties” is emerging(Caro-Llopis et al., 2016).

Altogether, these results brought a new comprehension on the mechanism of transcription and demonstrated the substantial benefits that arise from the combination of clinical and fundamental research.

Résumé en français

I - INTRODUCTION

La transcription est un mécanisme biochimique qui permet de copier l'information génétique présent sur le génome. Elle correspond à la synthèse d'un ARN à partir d'une matrice d'ADN et environ 75% du génome peut potentiellement être transcrit. La transcription se déroule dans le noyau et résulte d'une cascade d'évènements temporellement et spatialement orchestrés. La régulation de ce mécanisme est essentielle au développement des cellules, à leurs divisions, à leur différentiations et tout simplement indispensable à leurs vies.

Les gènes dont la séquence code pour des protéines sont transcrits par l'ARN polymérase II (Pol II), donnant naissance à des ARN messagers, qui seront ensuite traduit en protéines. L'initiation de la transcription par la Pol II fait intervenir de nombreux complexes protéiques tels que les facteurs généraux de transcription, le complexe Médiateur, des co-activateurs, des facteurs de remodelage de la chromatine ainsi que la Pol II elle même. Ils sont assemblés au niveau de la région promotrice du gène pour former le complexe d'initiation préalable à la transcription (Figure 1). Cette mise en place nécessite un remodelage de la chromatine permettant l'accès à la séquence du gène.

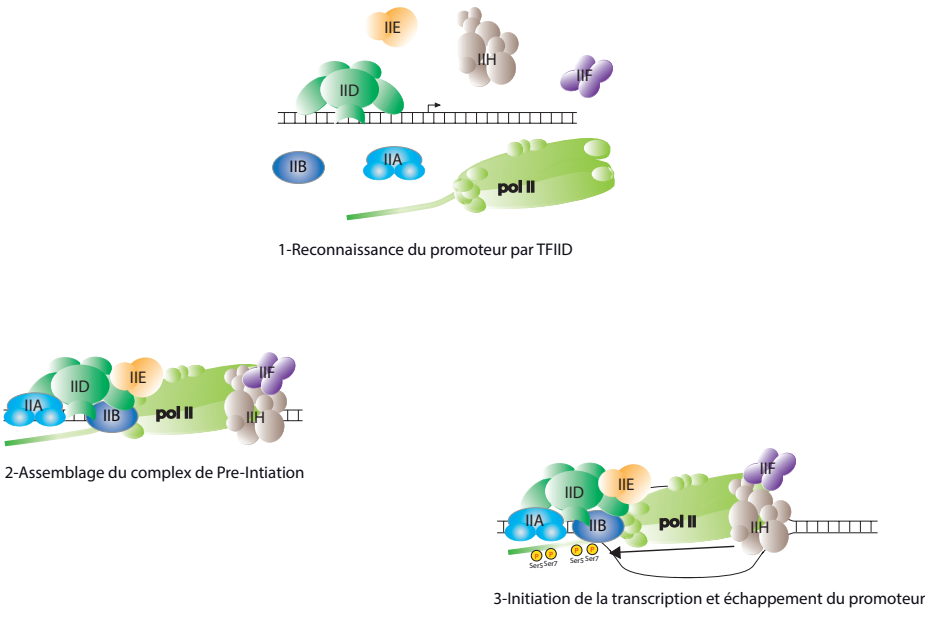


Figure 10: Initiation de la transcription

Tout événement qui perturbe le positionnement ou la fonction d'une telle machinerie peut conduire à une altération de l'expression des gènes et potentiellement être à l'origine de pathologies telles que des cancers ou des maladies génétiques. La thématique de notre équipe est de comprendre les mécanismes régissant l'initiation de la transcription, pour mieux appréhender leurs dérégulations dans un contexte pathologique. En ce qui me concerne, je me suis focalisé sur deux composants de la machinerie de transcription, la sous-unité MED12 du complexe Médiateur et le facteur de réparation XPC.

En effet, des études menées au sein du laboratoire ont permis de mettre en évidence une implication des protéines XPC, CSB, TFIIH, XPA, XPG, XPF-ERCC1 dans la régulation de l'expression de certains gènes. Ces différentes protéines ont été caractérisées à l'origine comme faisant partie de la machinerie de réparation de l'ADN par excision de nucléotide (NER)(Compe and Egly, 2012; Le May et al., 2010b). La protéine XPC est connue pour être responsable de la reconnaissance de lésion de l'ADN et ainsi permettre leurs éliminations par la voie NER.

Il a été montré que ces facteurs NER sont recrutés de manière séquentielle en aval de la machinerie transcriptionnelle au niveau du promoteur suite à l'activation des gènes. La présence de ces facteurs est également corrélée à plusieurs changements au niveau de la chromatine telles la modification post-traductionnelles des histones, l'induction de coupures de l'ADN ou la déméthylation de cytosines. L'ensemble de ces modifications se révèle nécessaire à la formation d'une boucle de chromatine entre le promoteur et le terminateur du gène activé et permet son expression optimale(Le May et al., 2010a, 2012). Des mutations au niveau des gènes codant pour différents facteurs NER ont été associées à des maladies génétiques humaines telles que le Xeroderma Pigmentosum (XP), la Trichothiodystrophie (TTD) et le syndrome de Cockayne (CS) (Table 2), pathologies dont les symptômes ne peuvent pas être expliqués uniquement par des défauts de réparation de l'ADN (Table 5), mais pourrait également être dû à des défauts transcriptionnelle.

XPC semble notamment régir la modification des histones par acétylation et méthylation au niveau de certains promoteurs. La première partie de mon projet

consistait donc à identifier les enzymes impliquées dans ces processus et à définir le mécanisme d'action de ces dernières en collaboration avec XPC.

Table 5: Facteurs NER et maladies associées

Protéine	Activité	Maladies associées
XPA	Reconnaissance des dommages	XP
XPB	3'-5' ATP dpdt hélicase	XP; XP/CS; TTD
XPC	Reconnaissance des dommages, remodelage de la chromatine	XP
XPD	5'-3' ATP dpdt hélicage; 5'-3' ATP dpdt translocase	XP; XP/CS; TTD; XP/TTD
XPE	Reconnaissance des dommages	XP
XPF	Endonucléase	XP
XPG	Endonucléase	XP; XPC/CS
XPV	ADN polymérase η	XP
CSA	Ubiquitin ligase	CS
CSB	ADN dpdt ATPase of SWI/SNF family	CS
p8/TTDA	Impliquer dans la NER	TTD

Également recruté au niveau du promoteur lors de la mise en place de la machinerie transcriptionnelle, le Médiateur (MED) est un complexe multi-protéique conservé au cours de l'évolution et constitué d'environ 30 sous-unités formant 4 modules distincts ('tête', 'milieu', 'queue' et 'kinase'). Son principal rôle est d'intégrer les différents signaux transmis par les facteurs de transcription fixés en amont du promoteur ou sur des séquences spécifiques afin de délivrer un message coordonné à la machinerie de transcription. De plus, certaines sous-unités du MED semblent être dédiées à la régulation spécifique de certains programmes d'expression génique(Malik and Roeder, 2010).

Au cours de la dernière décennie, de nombreux travaux ont montré l'association de certaines affections et de mutations dans les gènes codant pour les sous-unités du MED (Figure 2). Ces affections sont principalement caractérisées par des malformations congénitales, des retards mentaux et des cancers. Notre laboratoire a par exemple montré que la mutation (p.R617Q) de MED23 était à l'origine d'une déficience intellectuelle non syndromique, via la dérégulation de plusieurs gènes de réponse immédiate (IEGs)(Hashimoto et al., 2011). La seconde partie de mon projet portait sur l'étude de mutations dans le gène *MED12*, et leurs

conséquences sur la mise en place de la machinerie transcriptionnelle chez des patients présentant de larges troubles du développement.

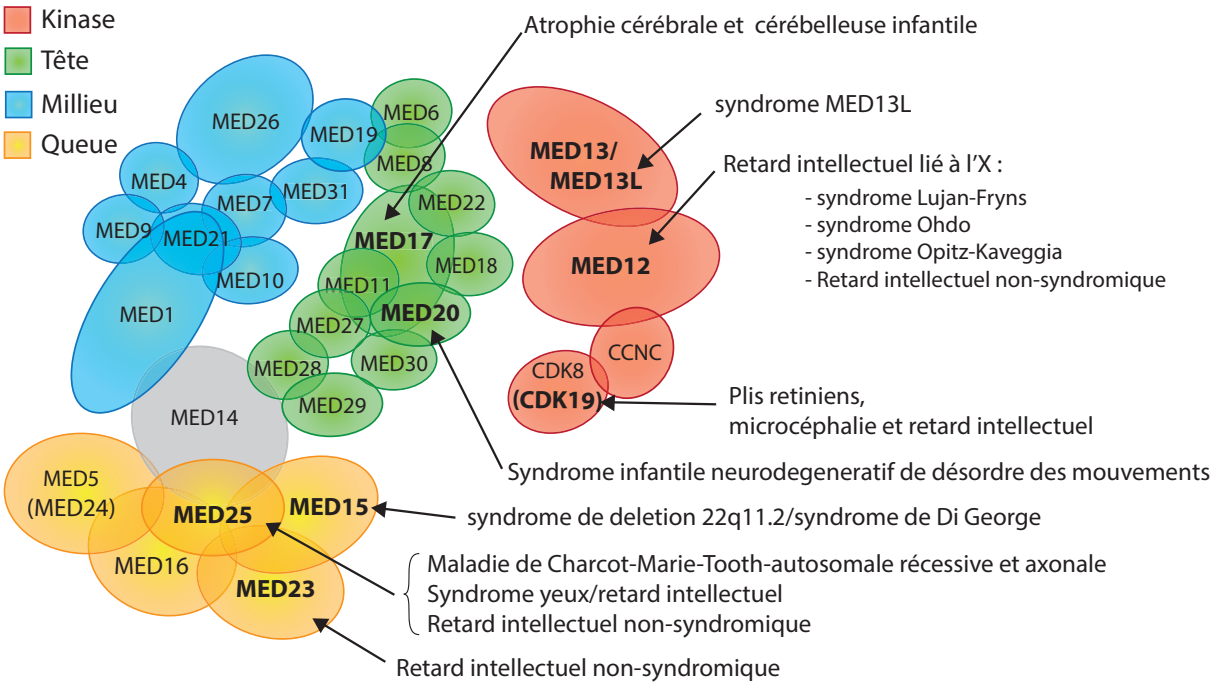


Figure 11: complexe Médiateur et maladies associées

II - RÉSULTAS

1) Le facteur de réparation de l'ADN XPC

En analysant des cellules de patients n'exprimant pas XPC ou dans lesquelles son expression a été restaurée, nous avons pu mettre en évidence son activité dans la régulation de l'expression de certains gènes. Dans le cadre de stimulations à l'acide transrétinoïque, la présence de la protéine XPC au niveau du promoteur de *RARβ2* est nécessaire à la mise en place d'un environnement chromatinien propice à la transcription, via notamment deux modifications post-traductionnelle de l'histone H3, l'acétylation de la lysine K9 (H3K9ac) et la triméthylation de la lysine K4 (H3K4me3). De plus, ce phénomène dépendant de la présence de XPC, semble indépendant des autres facteurs NER. Après immuno-précipitation, nous avons constaté que la protéine GCN5, une Histone Acétyl Transférase (HAT) co-précipitait avec XPC. Nous avons caractérisé cette interaction XPC/GCN5 par des expériences

in vitro et *in vivo*. Puis nous avons noté par Immuno-Précipitation de Chromatine (ChIP) que la présence de la protéine XPC était nécessaire au recrutement de GCN5 et à l'apparition de la marque H3K9ac au niveau du promoteur de *RARβ2*. Nous avons également reproduit ces résultats dans d'autres modèles cellulaires par l'utilisation de shRNA permettant de bloquer l'expression de XPC.

Afin d'évaluer l'ampleur de ce phénomène au niveau du génome, nous avons combiné les techniques de ChIPseq et de RNAseq. Nous avons ainsi pu définir un groupe d'environ 300 gènes dont l'expression est dérégulée en l'absence de XPC. Dans les cellules restaurées, leur promoteur est ciblé par la protéine XPC. Ils sont principalement impliqués dans la régulation de la structure de la chromatine et dans le contrôle de l'expression génique. Dans les cellules de patient n'exprimant pas XPC, on note l'absence de GCN5 et une diminution drastique des marques H3K9ac et H3K4me3 au niveau du Site d'Initiation de la Transcription (TSS) de ces gènes. On constate également ces résultats individuellement, sur différents gènes choisis aléatoirement parmi les 300 gènes.

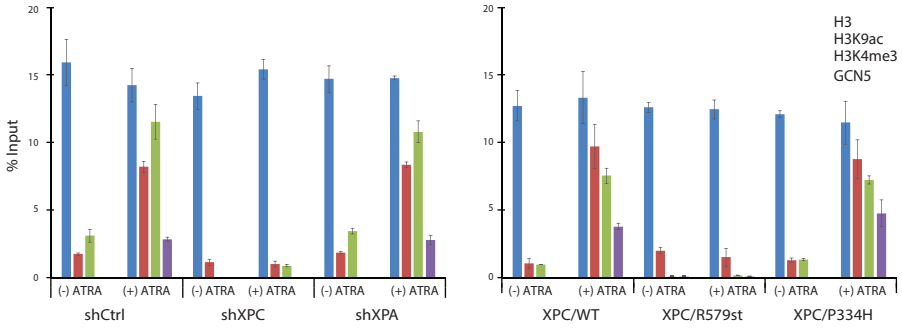


Figure 12: XPC contrôle l'acétylation de H3K9 et la tri-méthylation de H3K4

L'analyse de la séquence de ces promoteurs a montré un enrichissement de l'élément de fixation du facteur de transcription E2F1. Sur le promoteur de ces différents gènes, l'arrivée de la protéine XPC semble directement corrélée à la présence de ce facteur. Nous avons d'abord montré que la protéine XPC interagit à la fois avec E2F1 et avec GCN5, *in vitro* et *in vivo*. Ensuite, l'utilisation de siRNA ciblant E2F1 nous a permis de montrer l'importance de celui-ci dans la fixation de la protéine XPC sur les promoteurs. En effet, XPC et E2F1 s'avèrent chacun nécessaire au recrutement de l'autre, et tous les deux nécessaire à l'arrivée de GCN5 et donc à l'acétylation adéquate de H3K9.

Des expériences d'acétylation *in vitro* ont montrées que ni E2F1 ni XPC n'avait d'influence sur l'activité de l'enzyme GCN5. Les deux protéines semblent donc uniquement nécessaires à son recrutement au niveau du promoteur. Par contre, la présence du facteur général de transcription TFIIH augmente drastiquement l'activité de GCN5. De plus, l'absence des protéines XPC ou E2F1 semble perturber le recrutement de TFIIH au niveau des promoteurs.

En plus d'en dévoiler le mécanisme moléculaire, ces résultats permettent d'expliquer les défauts de transcription observés chez certains patients affectés par une mutation du gène *XPC*.

2) La sous-unité du Médiateur MED12

En parallèle, je me suis intéressé au rôle de la sous-unité MED12, dont le gène est porté par le chromosome X. Nous avons tout d'abord concentré notre travail sur l'étude de la mutation c.G3884A faux-sens conduisant à la substitution p.R1295H présente chez un patient hémizygote souffrant de retard mental, ainsi que chez sa mère hétérozygote. Dans des fibroblastes de peau isolés et mis en culture, j'ai dans un premier temps constaté un défaut d'expression de certains IEGs, et notamment de *JUN*, facteur de transcription membre de complexe AP-1, dans le cadre de stimulation au sérum. Nous avons ensuite mis en évidence par la technique de ChIP que ce défaut d'expression était dû à un problème de recrutement de la machinerie de transcription sur le promoteur de *JUN*.

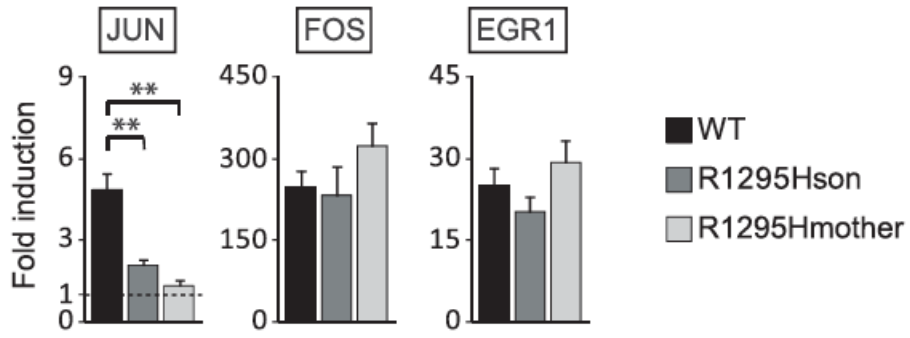


Figure 13: expression des gènes JUN, FOS et EGR1

La dérégulation de *Jun* induit une modification de l'expression des gènes de réponse tardive (LRGs), de manière directe et indirecte, ce que nous avons quantifié par RNAseq. Les gènes sous-exprimés, dont une majorité contient un élément de

fixation pour AP-1, sont entre autre impliqués dans la régulation du développement des crêtes neuronales. On note notamment une dérégulation de l'expression de la métallo-protéase MMP3, impliquée dans la plasticité et la migration neuronale. En parallèle, on observe également la surexpression de certains gènes dont *REST2*, codant pour un facteur de transcription connu pour inhiber l'expression des gènes neuronaux. L'augmentation de REST2 s'explique par une hausse du recrutement de la Pol II et un changement de recrutement conjoint des facteurs Jun et Fos sur son promoteur. En conséquence, *SYN1* (synapsin1), gène neuronal dont le promoteur est ciblé par la protéine REST2, voit son expression inhibée.

Dans un second temps, nous avons élargi l'analyse à d'autres types de stimulation et d'autres mutations, responsables des substitutions p.R206Q, p.N898D, p.R961W, p.N1007S, p.R1148H, p.S1165S et p.R1295H. Ces dernières sont à l'origine de maladies comme le syndrome de Lujan, le syndrome d'Opitz-Kaveggia, le syndrome d'Odho ou de retard mental lié à l'X.

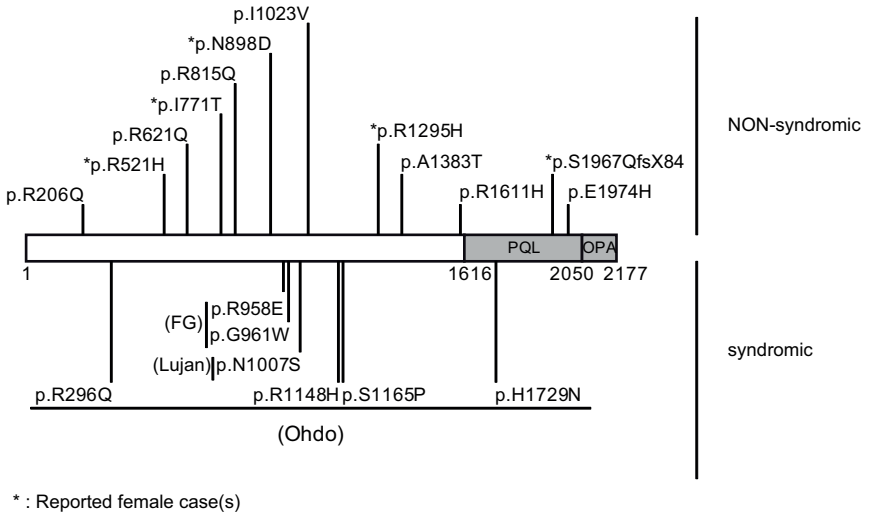


Figure 14: Mutations identifiées dans le gène de la protéine MED12

Nous avons soumis des lymphocytes périphériques transformés à différents stress (sérum, acide transrétinoïque, UV) et avons analysé l'expression de plusieurs gènes. Parmi différents IEGs, nous avons confirmé la dérégulation des gènes *FOS*, *JUN* et *EGR1* codant pour des facteurs de transcription, dans les cellules de patients. De plus, en fonction de la position de la mutation sur le gène *MED12* et du gène cible considéré, l'altération de son expression varie. L'étude du promoteur de ces gènes nous a permis de montrer que ces défauts de transcription étaient

également corrélés à une modification du recrutement de la machinerie transcriptionnelle.

De manière intéressante, les différentes mutations de MED12 n'affectent pas nécessairement l'expression des même gènes de réponse immédiate, ni n'entrainent les même altérations de ceux-ci. Ces différences que nous notons entre lignées cellulaires pourraient être une piste intéressante pour permettre d'expliquer la variabilité des symptômes observés chez les patients.

III - CONCLUSION

Ces résultats nous éclairent sur la position et la fonction de la protéine XPC et de sous-unité MED12 du complexe Médiateur au moment de l'initiation de la transcription. Outre l'aspect fondamental, l'étude des mécanismes moléculaires qui sous-tendent ces maladies permet d'en décrypter l'étiologie.

Bibliography

Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protić, M., Hübscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859–868.

Adegbola, A., Musante, L., Callewaert, B., Maciel, P., Hu, H., Isidor, B., Picker-Minh, S., Le Caignec, C., Delle Chiaie, B., Vanakker, O., et al. (2015). Redefining the MED13L syndrome. Eur. J. Hum. Genet. EJHG *23*, 1308–1317.

Akhtar, W., and Veenstra, G.J.C. (2011). TBP-related factors: a paradigm of diversity in transcription initiation. Cell Biosci. *1*, 23.

Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000). TFIIH is negatively regulated by cdk8-containing mediator complexes. Nature *407*, 102–106.

Alekseev, S., Nagy, Z., Sandoz, J., Weiss, A., Egly, J.-M., Le May, N., and Coin, F. (2017). Transcription without XPB Establishes a Unified Helicase-Independent Mechanism of Promoter Opening in Eukaryotic Gene Expression. Mol. Cell *65*, 504–514.e4.

Allfrey, V.G. (1966). Structural modifications of histones and their possible role in the regulation of ribonucleic acid synthesis. Proc. Can. Cancer Conf. *6*, 313–335.

Anish, R., Hossain, M.B., Jacobson, R.H., and Takada, S. (2009). Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements. PloS One *4*, e5103.

Ansari, S.A., He, Q., and Morse, R.H. (2009). Mediator complex association with constitutively transcribed genes in yeast. Proc. Natl. Acad. Sci. U. S. A. *106*, 16734–16739.

Araki, K., Nakajima, Y., Eto, K., and Ikeda, M.-A. (2003). Distinct recruitment of E2F family members to specific E2F-binding sites mediates activation and repression of the E2F1 promoter. Oncogene *22*, 7632–7641.

Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J.,

Ohkuma, Y., and Hanaoka, F. (2001). Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. J. Biol. Chem. *276*, 18665–18672.

Araújo, S.J., Tirode, F., Coin, F., Pospiech, H., Syväoja, J.E., Stucki, M., Hübscher, U., Egly, J.M., and Wood, R.D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. Genes Dev. *14*, 349–359.

Arents, G., and Moudrianakis, E.N. (1995). The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc. Natl. Acad. Sci. U. S. A. *92*, 11170–11174.

Armache, K.-J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *100*, 6964–6968.

Asadollahi, R., Oneda, B., Sheth, F., Azzarello-Burri, S., Baldinger, R., Joset, P., Latal, B., Knirsch, W., Desai, S., Baumer, A., et al. (2013). Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. Eur. J. Hum. Genet. EJHG *21*, 1100–1104.

Asadollahi, R., Zweier, M., Gogoll, L., Schiffmann, R., Sticht, H., Steindl, K., and Rauch, A. (2017). Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. Eur. J. Med. Genet. *60*, 451–464.

Asturias, F.J., Jiang, Y.W., Myers, L.C., Gustafsson, C.M., and Kornberg, R.D. (1999). Conserved structures of mediator and RNA polymerase II holoenzyme. Science *283*, 985–987.

Baek, H.J., Kang, Y.K., and Roeder, R.G. (2006). Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. J. Biol. Chem. *281*, 15172–15181.

Balamotis, M.A., Pennella, M.A., Stevens, J.L., Wasylyk, B., Belmont, A.S., and Berk, A.J. (2009). Complexity in transcription control at the activation domain-

mediator interface. Sci. Signal. *2*, ra20.

Bandara, L.R., Lam, E.W., Sørensen, T.S., Zamanian, M., Girling, R., and La Thangue, N.B. (1994). DP-1: a cell cycle-regulated and phosphorylated component of transcription factor DRTF1/E2F which is functionally important for recognition by pRb and the adenovirus E4 orf 6/7 protein. EMBO J. *13*, 3104–3114.

Barreto, G., Schäfer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Döderlein, G., Maltry, N., Wu, W., Lyko, F., et al. (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature *445*, 671–675.

Bártfai, R., Hoeijmakers, W.A.M., Salcedo-Amaya, A.M., Smits, A.H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T.-W., Françoijs, K.-J., and Stunnenberg, H.G. (2010). H2A.Z demarcates intergenic regions of the plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. PLoS Pathog. *6*, e1001223.

Basel-Vanagaite, L., Smirin-Yosef, P., Essakow, J.L., Tzur, S., Lagovsky, I., Maya, I., Pasmanik-Chor, M., Yeheskel, A., Konen, O., Orenstein, N., et al. (2015). Homozygous MED25 mutation implicated in eye-intellectual disability syndrome. Hum. Genet. *134*, 577–587.

Baumli, S., Hoeppner, S., and Cramer, P. (2005). A conserved mediator hinge revealed in the structure of the MED7.MED21 (Med7.Srb7) heterodimer. J. Biol. Chem. *280*, 18171–18178.

Bedford, M.T., and Clarke, S.G. (2009). Protein arginine methylation in mammals: who, what, and why. Mol. Cell *33*, 1–13.

Beerens, N., Hoeijmakers, J.H.J., Kanaar, R., Vermeulen, W., and Wyman, C. (2005). The CSB protein actively wraps DNA. J. Biol. Chem. *280*, 4722–4729.

Berger, J., Senti, K.-A., Senti, G., Newsome, T.P., Asling, B., Dickson, B.J., and Suzuki, T. (2008). Systematic identification of genes that regulate neuronal wiring in the Drosophila visual system. PLoS Genet. *4*, e1000085.

Berk, A.J. (2012). Yin and yang of mediator function revealed by human mutants. Proc. Natl. Acad. Sci. U. S. A. *109*, 19519–19520.

Bernardes de Jesus, B.M., Bjørås, M., Coin, F., and Egly, J.M. (2008). Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. Mol. Cell. Biol. *28*, 7225–7235.

Bernecky, C., and Taatjes, D.J. (2012). Activator–Mediator Binding Stabilizes RNA Polymerase II Orientation within the Human Mediator–RNA Polymerase II–TFIIF Assembly. J. Mol. Biol. *417*, 387–394.

Bernecky, C., Grob, P., Ebmeier, C.C., Nogales, E., and Taatjes, D.J. (2011). Molecular architecture of the human Mediator-RNA polymerase II-TFIIF assembly. PLoS Biol. *9*, e1000603.

Bernecky, C., Plitzko, J.M., and Cramer, P. (2017). Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp. Nat. Struct. Mol. Biol.

Berti, L., Mittler, G., Przemeck, G.K., Stelzer, G., Günzler, B., Amati, F., Conti, E., Dallapiccola, B., Hrabé de Angelis, M., Novelli, G., et al. (2001). Isolation and characterization of a novel gene from the DiGeorge chromosomal region that encodes for a mediator subunit. Genomics *74*, 320–332.

Beyer, K.S., Klauck, S.M., Benner, A., Poustka, F., and Poustka, A. (2002). Association studies of theHOPA dodecamer duplication variant in different subtypes of autism. Am. J. Med. Genet. *114*, 110–115.

Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451–454.

Black, J.C., Choi, J.E., Lombardo, S.R., and Carey, M. (2006). A mechanism for coordinating chromatin modification and preinitiation complex assembly. Mol. Cell *23*, 809–818.

van den Boom, V., Citterio, E., Hoogstraten, D., Zotter, A., Egly, J.-M., van Cappellen, W.A., Hoeijmakers, J.H.J., Houtsmuller, A.B., and Vermeulen, W. (2004).

DNA damage stabilizes interaction of CSB with the transcription elongation machinery. J. Cell Biol. *166*, 27–36.

Bouazzi, H., Lesca, G., Trujillo, C., Alwasiyah, M.K., and Munnich, A. (2015). Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His)]. Clin. Case Rep. *3*, 604–609.

Boube, M., Faucher, C., Joulia, L., Cribbs, D.L., and Bourbon, H.M. (2000). Drosophila homologs of transcriptional mediator complex subunits are required for adult cell and segment identity specification. Genes Dev. *14*, 2906–2917.

Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.-C., Ahmad, Y., Neel, H., et al. (2010). HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. Mol. Cell *39*, 912–924.

Bourbon, H.-M., Aguilera, A., Ansari, A.Z., Asturias, F.J., Berk, A.J., Bjorklund, S., Blackwell, T.K., Borggrefe, T., Carey, M., Carlson, M., et al. (2004). A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. Mol. Cell *14*, 553–557.

Boutry-Kryza, N., Labalme, A., Till, M., Schluth-Bolard, C., Langue, J., Turleau, C., Edery, P., and Sanlaville, D. (2012). An 800 kb deletion at 17q23.2 including the MED13 (THRAP1) gene, revealed by aCGH in a patient with a SMC 17p. Am. J. Med. Genet. A. *158A*, 400–405.

Brooks, P.J. (2013). Blinded by the UV light: How the focus on transcription-coupled NER has distracted from understanding the mechanisms of Cockayne syndrome neurologic disease. DNA Repair *12*, 656–671.

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell *56*, 549–561.

Burke, T.W., and Kadonaga, J.T. (1996). Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev. *10*, 711–724.

Burke, T.W., and Kadonaga, J.T. (1997). The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes Dev. *11*, 3020–3031.

Burke, J.R., Liban, T.J., Restrepo, T., Lee, H.-W., and Rubin, S.M. (2014). Multiple mechanisms for E2F binding inhibition by phosphorylation of the retinoblastoma protein C-terminal domain. J. Mol. Biol. *426*, 245–255.

Cafiero, C., Marangi, G., Orteschi, D., Ali, M., Asaro, A., Ponzi, E., Moncada, A., Ricciardi, S., Murdolo, M., Mancano, G., et al. (2015). Novel de novo heterozygous loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. Eur. J. Hum. Genet. EJHG *23*, 1499–1504.

Cai, G., Imasaki, T., Takagi, Y., and Asturias, F.J. (2009). Mediator structural conservation and implications for the regulation mechanism. Struct. Lond. Engl. 1993 *17*, 559–567.

Callier, P., Aral, B., Hanna, N., Lambert, S., Dindy, H., Ragon, C., Payet, M., Collod-Beroud, G., Carmignac, V., Delrue, M.A., et al. (2013). Systematic molecular and cytogenetic screening of 100 patients with marfanoid syndromes and intellectual disability. Clin. Genet. *84*, 507–521.

Cantin, G.T., Stevens, J.L., and Berk, A.J. (2003). Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. Proc. Natl. Acad. Sci. U. S. A. *100*, 12003–12008.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A.M., Taylor, M.S., Engström, P.G., Frith, M.C., et al. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nat. Genet. *38*, 626–635.

Caro-Llopis, A., Rosello, M., Orellana, C., Oltra, S., Monfort, S., Mayo, S., and Martinez, F. (2016). De novo mutations in genes of mediator complex causing syndromic intellectual disability: mediatorpathy or transcriptomopathy? Pediatr. Res. *80*, 809–815.

Cattoglio, C., Zhang, E.T., Grubisic, I., Chiba, K., Fong, Y.W., and Tjian, R.

(2015). Functional and mechanistic studies of XPC DNA-repair complex as transcriptional coactivator in embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *112*, E2317–E2326.

Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J. *18*, 4835–4845.

Chatzinikolaou, G., Apostolou, Z., Aid-Pavlidis, T., Ioannidou, A., Karakasilioti, I., Papadopoulos, G.L., Aivaliotis, M., Tsekrekou, M., Strouboulis, J., Kosteas, T., et al. (2017). ERCC1-XPF cooperates with CTCF and cohesin to facilitate the developmental silencing of imprinted genes. Nat. Cell Biol. *19*, 421–432.

Chen, P., and Li, G. (2010). Dynamics of the higher-order structure of chromatin. Protein Cell *1*, 967–971.

Chen, C.-P., Lin, M.-H., Chen, Y.-Y., Chern, S.-R., Chen, Y.-N., Wu, P.-S., Pan, C.-W., Lee, M.-S., and Wang, W. (2015). Prenatal diagnosis and array comparative genomic hybridization characterization of interstitial deletions of 8q23.3-q24.11 and 8q24.13 associated with Langer-Giedion syndrome, Cornelia de Lange syndrome and haploinsufficiency of TRPS1, RAD21 and EXT1. Taiwan. J. Obstet. Gynecol. *54*, 592–596.

Chen, J., Larochelle, S., Li, X., and Suter, B. (2003). Xpd/Ercc2 regulates CAK activity and mitotic progression. Nature *424*, 228–232.

Chen, L., Wei, T., Si, X., Wang, Q., Li, Y., Leng, Y., Deng, A., Chen, J., Wang, G., Zhu, S., et al. (2013). Lysine acetyltransferase GCN5 potentiates the growth of non-small cell lung cancer via promotion of E2F1, cyclin D1, and cyclin E1 expression. J. Biol. Chem. *288*, 14510–14521.

Chen, W., Rogatsky, I., and Garabedian, M.J. (2006). MED14 and MED1 differentially regulate target-specific gene activation by the glucocorticoid receptor. Mol. Endocrinol. Baltim. Md *20*, 560–572.

Cheng, B., Li, T., Rahl, P.B., Adamson, T.E., Loudas, N.B., Guo, J., Varzavand, K., Cooper, J.J., Hu, X., Gnatt, A., et al. (2012). Functional association of

Gdown1 with RNA polymerase II poised on human genes. Mol. Cell *45*, 38–50.

Choi, C.H., Hiromura, M., and Usheva, A. (2003). Transcription factor IIB acetylates itself to regulate transcription. Nature *424*, 965–969.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834–840.

Clagett-Dame, M., and DeLuca, H.F. (2002). The role of vitamin A in mammalian reproduction and embryonic development. Annu. Rev. Nutr. *22*, 347–381.

Clark, R.D., Graham, J.M., Friez, M.J., Hoo, J.J., Jones, K.L., McKeown, C., Moeschler, J.B., Raymond, F.L., Rogers, R.C., Schwartz, C.E., et al. (2009). FG syndrome, an X-linked multiple congenital anomaly syndrome: The clinical phenotype and an algorithm for diagnostic testing. Genet. Med. *11*, 769–775.

Cleaver, J.E. (1968). Defective repair replication of DNA in xeroderma pigmentosum. Nature *218*, 652–656.

Cleaver, J.E. (2005). Cancer in xeroderma pigmentosum and related disorders of DNA repair. Nat. Rev. Cancer *5*, 564–573.

Codina-Solà, M., Rodríguez-Santiago, B., Homs, A., Santoyo, J., Rigau, M., Aznar-Laín, G., Del Campo, M., Gener, B., Gabau, E., Botella, M.P., et al. (2015). Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. Mol. Autism *6*, 21.

Coin, F., Marinoni, J.C., Rodolfo, C., Fribourg, S., Pedrini, A.M., and Egly, J.M. (1998). Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. Nat. Genet. *20*, 184–188.

Coin, F., Oksenych, V., Mocquet, V., Groh, S., Blattner, C., and Egly, J.M. (2008). Nucleotide excision repair driven by the dissociation of CAK from TFIIH. Mol. Cell *31*, 9–20.

Compe, E., and Egly, J.-M. (2012). TFIIH: when transcription met DNA repair. Nat. Rev. Mol. Cell Biol. *13*, 343–354.

Compe, E., and Egly, J.-M. (2016). Nucleotide Excision Repair and Transcriptional Regulation: TFIIH and Beyond. Annu. Rev. Biochem. *85*, 265–290.

Conaway, R.C., and Conaway, J.W. (2013). The Mediator complex and transcription elongation. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1829*, 69–75.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C., and Chambon, P. (1980). Promoter sequences of eukaryotic protein-coding genes. Science *209*, 1406–1414.

Daniels, D.L., Ford, M., Schwinn, M.K., Benink, H., Galbraith, M.D., Amunugama, R., Jones, R., Allen, D., Okazaki, N., Yamakawa, H., et al. (2013). Mutual exclusivity of MED12/MED12L, MED13/13L, and CDK8/19 paralogs revealed within the CDK-Mediator kinase module. J Proteomics Bioinform S *2*, 2.

Daury, L., Chailleux, C., Bonvallet, J., and Trouche, D. (2006). Histone H3.3 deposition at E2F-regulated genes is linked to transcription. EMBO Rep. *7*, 66–71.

Davis, J.A., Takagi, Y., Kornberg, R.D., and Asturias, F.A. (2002). Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. Mol. Cell *10*, 409–415.

Davis, M.A., Larimore, E.A., Fissel, B.M., Swanger, J., Taatjes, D.J., and Clurman, B.E. (2013). The SCF-Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with Mediator. Genes Dev. *27*, 151–156.

DeGregori, J., and Johnson, D.G. (2006). Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. Curr. Mol. Med. *6*, 739–748.

DeGregori, J., Leone, G., Miron, A., Jakoi, L., and Nevins, J.R. (1997). Distinct roles for E2F proteins in cell growth control and apoptosis. Proc. Natl. Acad. Sci. U.

S. A. *94*, 7245–7250.

DeLisi, L.E., Smith, A.B., Razi, K., Stewart, J., Wang, Z., Sandhu, H.K., and Philibert, R.A. (2000). Investigation of a candidate gene for schizophrenia on Xq13 previously associated with mental retardation and hypothyroidism. Am. J. Med. Genet. *96*, 398–403.

De Luca, A., Conti, E., Grifone, N., Amati, F., Spalletta, G., Caltagirone, C., Bonaviri, G., Pasini, A., Gennarelli, M., Stefano, B., et al. (2003). Association study between CAG trinucleotide repeats in the PCQAP gene (PC2 glutamine/Q-rich-associated protein) and schizophrenia. Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet. *116B*, 32–35.

Deng, W., and Roberts, S.G.E. (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. Genes Dev. *19*, 2418–2423.

Devault, A., Martinez, A.M., Fesquet, D., Labbé, J.C., Morin, N., Tassan, J.P., Nigg, E.A., Cavadore, J.C., and Dorée, M. (1995). MAT1 ('menage à trois') a new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and Xenopus CAK. EMBO J. *14*, 5027–5036.

Dikstein, R., Ruppert, S., and Tjian, R. (1996). TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. Cell *84*, 781–790.

Ding, N., Zhou, H., Esteve, P.-O., Chin, H.G., Kim, S., Xu, X., Joseph, S.M., Friez, M.J., Schwartz, C.E., Pradhan, S., et al. (2008). Mediator Links Epigenetic Silencing of Neuronal Gene Expression with X-Linked Mental Retardation. Mol. Cell *31*, 347–359.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101–108.

Donnio, L.-M., Bidon, B., Hashimoto, S., May, M., Epanchintsev, A., Ryan, C., Allen, W., Hackett, A., Gecz, J., Skinner, C., et al. (2017). MED12-related XLID disorders are dose-dependent of immediate early genes (IEGs) expression. Hum. Mol. Genet. *26*, 2062–2075.

Dotson, M.R., Yuan, C.X., Roeder, R.G., Myers, L.C., Gustafsson, C.M., Jiang, Y.W., Li, Y., Kornberg, R.D., and Asturias, F.J. (2000). Structural organization of yeast and mammalian mediator complexes. Proc. Natl. Acad. Sci. U. S. A. *97*, 14307–14310.

Drané, P., Compe, E., Catez, P., Chymkowitch, P., and Egly, J.-M. (2004). Selective regulation of vitamin D receptor-responsive genes by TFIIH. Mol. Cell *16*, 187–197.

Dubaele, S., Proietti De Santis, L., Bienstock, R.J., Keriel, A., Stefanini, M., Van Houten, B., and Egly, J.-M. (2003). Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. Mol. Cell *11*, 1635–1646.

Ebmeier, C.C., and Taatjes, D.J. (2010). Activator-Mediator binding regulates Mediator-cofactor interactions. Proc. Natl. Acad. Sci. U. S. A. *107*, 11283–11288.

Elmlund, H., Baraznenok, V., Lindahl, M., Samuelsen, C.O., Koeck, P.J.B., Holmberg, S., Hebert, H., and Gustafsson, C.M. (2006). The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *103*, 15788–15793.

Escriva, H., Bertrand, S., and Laudet, V. (2004). The evolution of the nuclear receptor superfamily. Essays Biochem. *40*, 11–26.

Esnault, C., Ghavi-Helm, Y., Brun, S., Soutourina, J., Van Berkum, N., Boschiero, C., Holstege, F., and Werner, M. (2008). Mediator-Dependent Recruitment of TFIIH Modules in Preinitiation Complex. Mol. Cell *31*, 337–346.

Eyboulet, F., Cibot, C., Eychenne, T., Neil, H., Alibert, O., Werner, M., and Soutourina, J. (2013). Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment. Genes Dev. *27*, 2549–2562.

Feaver, W.J., Svejstrup, J.Q., Bardwell, L., Bardwell, A.J., Buratowski, S., Gulyas, K.D., Donahue, T.F., Friedberg, E.C., and Kornberg, R.D. (1993). Dual roles of a multiprotein complex from S. cerevisiae in transcription and DNA repair. Cell *75*, 1379–1387.

Fieremans, N., Van Esch, H., Holvoet, M., Van Goethem, G., Devriendt, K., Rosello, M., Mayo, S., Martinez, F., Jhangiani, S., Muzny, D.M., et al. (2016). Identification of Intellectual Disability Genes in Female Patients with a Skewed X-Inactivation Pattern. Hum. Mutat. *37*, 804–811.

Figueiredo, T., Melo, U.S., Pessoa, A.L.S., Nobrega, P.R., Kitajima, J.P., Correa, I., Zatz, M., Kok, F., and Santos, S. (2015). Homozygous missense mutation in MED25 segregates with syndromic intellectual disability in a large consanguineous family. J. Med. Genet. *52*, 123–127.

Fishburn, J., Tomko, E., Galburt, E., and Hahn, S. (2015). Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation. Proc. Natl. Acad. Sci. U. S. A. *112*, 3961–3966.

Fitch, M.E., Nakajima, S., Yasui, A., and Ford, J.M. (2003). In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. J. Biol. Chem. *278*, 46906–46910.

Flanagan, P.M., Kelleher, R.J., Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. Nature *350*, 436–438.

Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z.F., and Reinberg, D. (1991). The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. Proc. Natl. Acad. Sci. U. S. A. *88*, 9999–10003.

Fondell, J.D., Ge, H., and Roeder, R.G. (1996). Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. Proc. Natl. Acad. Sci. U. S. A. *93*, 8329–8333.

Fong, Y.W., Inouye, C., Yamaguchi, T., Cattoglio, C., Grubisic, I., and Tjian, R. (2011). A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. Cell *147*, 120–131.

Fong, Y.W., Cattoglio, C., and Tjian, R. (2013). The intertwined roles of transcription and repair proteins. Mol. Cell *52*, 291–302.

Forget, D., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J., and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. Proc. Natl. Acad. Sci. U. S. A. *94*, 7150–7155.

Fousteri, M., Vermeulen, W., van Zeeland, A.A., and Mullenders, L.H.F. (2006). Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. Mol. Cell *23*, 471–482.

Fowler, T., Sen, R., and Roy, A.L. (2011). Regulation of primary response genes. Mol. Cell *44*, 348–360.

Friedberg, E.C., Aguilera, A., Gellert, M., Hanawalt, P.C., Hays, J.B., Lehmann, A.R., Lindahl, T., Lowndes, N., Sarasin, A., and Wood, R.D. (2006). DNA repair: from molecular mechanism to human disease. DNA Repair *5*, 986–996.

Fryns, J.P., and Buttiens, M. (1987). X-linked mental retardation with marfanoid habitus. Am. J. Med. Genet. *28*, 267–274.

Gershenzon, N.I., and Ioshikhes, I.P. (2005). Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinforma. Oxf. Engl. *21*, 1295–1300.

Ghare, S.S., Joshi-Barve, S., Moghe, A., Patil, M., Barker, D.F., Gobejishvili, L., Brock, G.N., Cave, M., McClain, C.J., and Barve, S.S. (2014). Coordinated histone H3 methylation and acetylation regulate physiologic and pathologic fas ligand gene expression in human CD4+ T cells. J. Immunol. Baltim. Md 1950 *193*, 412–421.

Giglia-Mari, G., Coin, F., Ranish, J.A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N.G.J., Raams, A., Argentini, M., van der Spek, P.J., et al. (2004). A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. Nat. Genet. *36*, 714–719.

Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W.M., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature

511, 344–347.

Girling, R., Partridge, J.F., Bandara, L.R., Burden, N., Totty, N.F., Hsuan, J.J., and La Thangue, N.B. (1993). A new component of the transcription factor DRTF1/E2F. Nature *365*, 468.

Gong, F., Fahy, D., and Smerdon, M.J. (2006). Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. Nat. Struct. Mol. Biol. *13*, 902–907.

Gozukara, E.M., Khan, S.G., Metin, A., Emmert, S., Busch, D.B., Shahlavi, T., Coleman, D.M., Miller, M., Chinsomboon, N., Stefanini, M., et al. (2001). A stop codon in xeroderma pigmentosum group C families in Turkey and Italy: molecular genetic evidence for a common ancestor. J. Invest. Dermatol. *117*, 197–204.

Greenberg, M.E., and Ziff, E.B. (1984). Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. Nature *311*, 433–438.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat. Rev. Genet. *13*, 343–357.

Groisman, R., Kuraoka, I., Chevallier, O., Gaye, N., Magnaldo, T., Tanaka, K., Kisselev, A.F., Harel-Bellan, A., and Nakatani, Y. (2006). CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. Genes Dev. *20*, 1429–1434.

Gu, B., Eick, D., and Bensaude, O. (2013). CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. Nucleic Acids Res. *41*, 1591–1603.

Guermah, M., Malik, S., and Roeder, R.G. (1998). Involvement of TFIID and USA components in transcriptional activation of the human immunodeficiency virus promoter by NF-kappaB and Sp1. Mol. Cell. Biol. *18*, 3234–3244.

Guermah, M., Tao, Y., and Roeder, R.G. (2001). Positive and Negative TAFII Functions That Suggest a Dynamic TFIID Structure and Elicit Synergy with TRAPs in Activator-Induced Transcription. Mol. Cell. Biol. *21*, 6882–6894.

Guo, R., Chen, J., Zhu, F., Biswas, A.K., Berton, T.R., Mitchell, D.L., and Johnson, D.G. (2010). E2F1 localizes to sites of UV-induced DNA damage to enhance nucleotide excision repair. J. Biol. Chem. *285*, 19308–19315.

Guo, R., Chen, J., Mitchell, D.L., and Johnson, D.G. (2011). GCN5 and E2F1 stimulate nucleotide excision repair by promoting H3K9 acetylation at sites of damage. Nucleic Acids Res. *39*, 1390–1397.

Ha, I., Roberts, S., Maldonado, E., Sun, X., Kim, L.U., Green, M., and Reinberg, D. (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev. *7*, 1021–1032.

van Haelst, M.M., Monroe, G.R., Duran, K., van Binsbergen, E., Breur, J.M., Giltay, J.C., and van Haaften, G. (2015). Further confirmation of the MED13L haploinsufficiency syndrome. Eur. J. Hum. Genet. EJHG *23*, 135–138.

Hamdan, F.F., Srour, M., Capo-Chichi, J.-M., Daoud, H., Nassif, C., Patry, L., Massicotte, C., Ambalavanan, A., Spiegelman, D., Diallo, O., et al. (2014). De novo mutations in moderate or severe intellectual disability. PLoS Genet. *10*, e1004772.

Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. Nat. Rev. Mol. Cell Biol. *9*, 958–970.

Hantsche, M., and Cramer, P. (2017). Conserved RNA polymerase II initiation complex structure. Curr. Opin. Struct. Biol. *47*, 17–22.

Harlen, K.M., Trotta, K.L., Smith, E.E., Mosaheb, M.M., Fuchs, S.M., and Churchman, L.S. (2016). Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. Cell Rep. *15*, 2147–2158.

Hashimoto, S., Boissel, S., Zarhrate, M., Rio, M., Munnich, A., Egly, J.-M., and Colleaux, L. (2011). MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. Science *333*, 1161–1163.

He, H., Wang, J., and Liu, T. (2017). UV-Induced RPA1 Acetylation Promotes Nucleotide Excision Repair. Cell Rep. *20*, 2010–2025.

He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. Nature *495*, 481–486.

Healy, S., Khan, P., and Davie, J.R. (2013). Immediate early response genes and cell transformation. Pharmacol. Ther. *137*, 64–77.

Herschman, H.R. (1991). Primary response genes induced by growth factors and tumor promoters. Annu. Rev. Biochem. *60*, 281–319.

Hirabayashi, S., Saitsu, H., and Matsumoto, N. (2016). Distinct but milder phenotypes with choreiform movements in siblings with compound heterozygous mutations in the transcription preinitiation mediator complex subunit 17 (MED17). Brain Dev. *38*, 118–123.

Ho, J.J., Cattoglio, C., McSwiggen, D.T., Tjian, R., and Fong, Y.W. (2017). Regulation of DNA demethylation by the XPC DNA repair complex in somatic and pluripotent stem cells. Genes Dev. *31*, 830–844.

Hoeppner, S., Baumli, S., and Cramer, P. (2005). Structure of the mediator subunit cyclin C and its implications for CDK8 function. J. Mol. Biol. *350*, 833–842.

Hong, S.-K., Haldin, C.E., Lawson, N.D., Weinstein, B.M., Dawid, I.B., and Hukriede, N.A. (2005). The zebrafish kohtalo/trap230 gene is required for the development of the brain, neural crest, and pronephric kidney. Proc. Natl. Acad. Sci. *102*, 18473–18478.

Hsieh, T.-H.S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108–119.

Hsu, L.M. (2002). Promoter clearance and escape in prokaryotes. Biochim. Biophys. Acta *1577*, 191–207.

Hu, J., Adar, S., Selby, C.P., Lieb, J.D., and Sancar, A. (2015). Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. Genes Dev. *29*, 948–960.

Hu, X., Malik, S., Negroiu, C.C., Hubbard, K., Velalar, C.N., Hampton, B.,

Grosu, D., Catalano, J., Roeder, R.G., and Gnatt, A. (2006). A Mediator-responsive form of metazoan RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *103*, 9506–9511.

Huang, Y., Li, W., Yao, X., Lin, Q., Yin, J., Liang, Y., Heiner, M., Tian, B., Hui, J., and Wang, G. (2012). Mediator Complex Regulates Alternative mRNA Processing via the MED23 Subunit. Mol. Cell *45*, 459–469.

Imasaki, T., Calero, G., Cai, G., Tsai, K.-L., Yamada, K., Cardelli, F., Erdjument-Bromage, H., Tempst, P., Berger, I., Kornberg, G.L., et al. (2011). Architecture of the Mediator head module. Nature *475*, 240–243.

Ingram, L., Munro, S., Coutts, A.S., and La Thangue, N.B. (2011). E2F-1 regulation by an unusual DNA damage-responsive DP partner subunit. Cell Death Differ. *18*, 122–132.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.-H., Narzisi, G., Leotta, A., et al. (2012). De novo gene disruptions in children on the autistic spectrum. Neuron *74*, 285–299.

Isidor, B., Lefebvre, T., Le Vaillant, C., Caillaud, G., Faivre, L., Jossic, F., Joubert, M., Winer, N., Le Caignec, C., Borck, G., et al. (2014). Blepharophimosis, short humeri, developmental delay and hirschsprung disease: expanding the phenotypic spectrum of MED12 mutations. Am. J. Med. Genet. A. *164A*, 1821–1825.

Ito, M., Yuan, C.-X., Malik, S., Gu, W., Fondell, J.D., Yamamura, S., Fu, Z.-Y., Zhang, X., Qin, J., and Roeder, R.G. (1999). Identity between TRAP and SMCC Complexes Indicates Novel Pathways for the Function of Nuclear Receptors and Diverse Mammalian Activators. Mol. Cell *3*, 361–370.

Ito, S., Kuraoka, I., Chymkowitch, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.-M., and Tanaka, K. (2007). XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for Cockayne syndrome in XP-G/CS patients. Mol. Cell *26*, 231–243.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S.T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells.

Mol. Cell. Biol. *14*, 116–127.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions. Nat. Genet. *41*, 941–945.

Jin, Q., Yu, L.-R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.-E., Wang, C., Brindle, P.K., Dent, S.Y.R., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. EMBO J. *30*, 249–262.

Jishage, M., Malik, S., Wagner, U., Uberheide, B., Ishihama, Y., Hu, X., Chait, B.T., Gnatt, A., Ren, B., and Roeder, R.G. (2012). Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. Mol. Cell *45*, 51–63.

Johnson, K.M., and Carey, M. (2003). Assembly of a Mediator/TFIID/TFIIA Complex Bypasses the Need for an Activator. Curr. Biol. *13*, 772–777.

Johnson, K.M., Wang, J., Smallwood, A., Arayata, C., and Carey, M. (2002). TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. Genes Dev. *16*, 1852–1863.

Johnson, R.E., Kondratick, C.M., Prakash, S., and Prakash, L. (1999). hRAD30 mutations in the variant form of xeroderma pigmentosum. Science *285*, 263–265.

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. *16*, 167–177.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kakumu, E., Nakanishi, S., Shiratori, H.M., Kato, A., Kobayashi, W., Machida, S., Yasuda, T., Adachi, N., Saito, N., Ikura, T., et al. (2017). Xeroderma

pigmentosum group C protein interacts with histones: regulation by acetylated states of histone H3. Genes Cells Devoted Mol. Cell. Mech. *22*, 310–327.

Kamileri, I., Karakasilioti, I., Sideri, A., Kosteas, T., Tatarakis, A., Talianidis, I., and Garinis, G.A. (2012). Defective transcription initiation causes postnatal growth failure in a mouse model of nucleotide excision repair (NER) progeria. Proc. Natl. Acad. Sci. U. S. A. *109*, 2995–3000.

Kang, K.A., and Hyun, J.W. (2017). Oxidative Stress, Nrf2, and Epigenetic Modification Contribute to Anticancer Drug Resistance. Toxicol. Res. *33*, 1–5.

Kaufmann, J., and Smale, S.T. (1994). Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes Dev. *8*, 821–829.

Kaufmann, W.E., and Moser, H.W. (2000). Dendritic anomalies in disorders associated with mental retardation. Cereb. Cortex N. Y. N 1991 *10*, 981–991.

Kaufmann, R., Straussberg, R., Mandel, H., Fattal-Valevski, A., Ben-Zeev, B., Naamati, A., Shaag, A., Zenvirt, S., Konen, O., Mimouni-Bloch, A., et al. (2010). Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. Am. J. Hum. Genet. *87*, 667–670.

Kaya, N., Colak, D., Albakheet, A., Al-Owain, M., Abu-Dheim, N., Al-Younes, B., Al-Zahrani, J., Mukaddes, N.M., Dervent, A., Al-Dosari, N., et al. (2012). A novel X-linked disorder with developmental delay and autistic features. Ann. Neurol. *71*, 498–508.

Kelleher, R.J., Flanagan, P.M., and Kornberg, R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell *61*, 1209–1215.

Kikuchi, Y., Umemura, H., Nishitani, S., Iida, S., Fukasawa, R., Hayashi, H., Hirose, Y., Tanaka, A., Sugasawa, K., and Ohkuma, Y. (2015). Human mediator MED17 subunit plays essential roles in gene regulation by associating with the transcription and DNA repair machineries. Genes Cells Devoted Mol. Cell. Mech. *20*, 191–202.

Killeen, M., Coulombe, B., and Greenblatt, J. (1992). Recombinant TBP, transcription factor IIB, and RAP30 are sufficient for promoter recognition by mammalian RNA polymerase II. J. Biol. Chem. *267*, 9463–9466.

Kim, Y.J., Björklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell *77*, 599–608.

Kimura, A., Matsubara, K., and Horikoshi, M. (2005). A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. J. Biochem. (Tokyo) *138*, 647–662.

Kissil, J.L., Feinstein, E., Cohen, O., Jones, P.A., Tsai, Y.C., Knowles, M.A., Eydmann, M.E., and Kimchi, A. (1997). DAP-kinase loss of expression in various carcinoma and B-cell lymphoma cell lines: possible implications for role as tumor suppressor gene. Oncogene *15*, 403–407.

Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. Nature *368*, 466–469.

Koleske, A.J., and Young, R.A. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. Trends Biochem. Sci. *20*, 113–116.

Koleske, A.J., Buratowski, S., Nonet, M., and Young, R.A. (1992). A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. Cell *69*, 883–894.

Koschubs, T., Seizl, M., Larivière, L., Kurth, F., Baumli, S., Martin, D.E., and Cramer, P. (2009). Identification, structure, and functional requirement of the Mediator submodule Med7N/31. EMBO J. *28*, 69–80.

Kouzarides, T. (2007). SnapShot: Histone-modifying enzymes. Cell *131*, 822.

Kovesdi, I., Reichel, R., and Nevins, J.R. (1986). Identification of a cellular transcription factor involved in E1A trans-activation. Cell *45*, 219–228.

Krebs, A.R., Karmodiya, K., Lindahl-Allen, M., Struhl, K., and Tora, L. (2011). SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of

genes and ATAC defines a class of p300-independent enhancers. Mol. Cell *44*, 410–423.

Krivega, I., and Dean, A. (2012). Enhancer and promoter interactions-long distance calls. Curr. Opin. Genet. Dev. *22*, 79–85.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nat. Rev. Mol. Cell Biol. *12*, 283–294.

Kuschal, C., Botta, E., Orioli, D., Digiovanna, J.J., Seneca, S., Keymolen, K., Tamura, D., Heller, E., Khan, S.G., Caligiuri, G., et al. (2016). GTF2E2 Mutations Destabilize the General Transcription Factor Complex TFIIE in Individuals with DNA Repair-Proficient Trichothiodystrophy. Am. J. Hum. Genet. *98*, 627–642.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev. *12*, 34–44.

Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature *494*, 497–501.

Lainé, J.-P., Singh, B.N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. Genes Dev. *23*, 2604–2609.

Lang, S.E., McMahon, S.B., Cole, M.D., and Hearing, P. (2001). E2F transcriptional activation requires TRRAP and GCN5 cofactors. J. Biol. Chem. *276*, 32627–32634.

Langley, K.G., Brown, J., Gerber, R.J., Fox, J., Friez, M.J., Lyons, M., and Schrier Vergano, S.A. (2015). Beyond Ohdo syndrome: A familial missense mutation broadens the MED12 spectrum. Am. J. Med. Genet. A. *167A*, 3180–3185.

Larivière, L., Geiger, S., Hoeppner, S., Röther, S., Strässer, K., and Cramer, P. (2006). Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. Nat. Struct. Mol. Biol. *13*, 895–901.

Larivière, L., Plaschka, C., Seizl, M., Wenzeck, L., Kurth, F., and Cramer, P. (2012). Structure of the Mediator head module. Nature *492*, 448–451.

Larivière, L., Plaschka, C., Seizl, M., Petrotchenko, E.V., Wenzeck, L., Borchers, C.H., and Cramer, P. (2013). Model of the Mediator middle module based on protein cross-linking. Nucleic Acids Res. *41*, 9266–9273.

Larochelle, S., Amat, R., Glover-Cutter, K., Sansó, M., Zhang, C., Allen, J.J., Shokat, K.M., Bentley, D.L., and Fisher, R.P. (2012). Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. Nat. Struct. Mol. Biol. *19*, 1108–1115.

Laugel, V. (2013). Cockayne syndrome: the expanding clinical and mutational spectrum. Mech. Ageing Dev. *134*, 161–170.

Lavigne, A.C., Mengus, G., Gangloff, Y.G., Wurtz, J.M., and Davidson, I. (1999). Human TAF(II)55 interacts with the vitamin D(3) and thyroid hormone receptors and with derivatives of the retinoid X receptor that have altered transactivation properties. Mol. Cell. Biol. *19*, 5486–5494.

Leal, A., Huehne, K., Bauer, F., Sticht, H., Berger, P., Suter, U., Morera, B., Del Valle, G., Lupski, J.R., Ekici, A., et al. (2009). Identification of the variant Ala335Val of MED25 as responsible for CMT2B2: molecular data, functional studies of the SH3 recognition motif and correlation between wild-type MED25 and PMP22 RNA levels in CMT1A animal models. Neurogenetics *10*, 275–287.

Lee, D.-H., Gershenzon, N., Gupta, M., Ioshikhes, I.P., Reinberg, D., and Lewis, B.A. (2005). Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. Mol. Cell. Biol. *25*, 9674–9686.

Lee, D.K., DeJong, J., Hashimoto, S., Horikoshi, M., and Roeder, R.G. (1992). TFIIA induces conformational changes in TFIID via interactions with the basic repeat. Mol. Cell. Biol. *12*, 5189–5196.

Lee, H.-K., Park, U.-H., Kim, E.-J., and Um, S.-J. (2007). MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. EMBO J. *26*, 3545–3557.

Lee, K.-B., Wang, D., Lippard, S.J., and Sharp, P.A. (2002). Transcription-coupled and DNA damage-dependent ubiquitination of RNA polymerase II in vitro. Proc. Natl. Acad. Sci. U. S. A. *99*, 4239–4244.

Le May, N., Mota-Fernandes, D., Vélez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010a). NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. Mol. Cell *38*, 54–66.

Le May, N., Egly, J.-M., and Coin, F. (2010b). True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. J. Nucleic Acids *2010*.

Le May, N., Fradin, D., Iltis, I., Bougnères, P., and Egly, J.-M. (2012). XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. Mol. Cell *47*, 622–632.

Lesca, G., Moizard, M.-P., Bussy, G., Boggio, D., Hu, H., Haas, S.A., Ropers, H.-H., Kalscheuer, V.M., Des Portes, V., Labalme, A., et al. (2013). Clinical and neurocognitive characterization of a family with a novel *MED12* gene frameshift mutation. Am. J. Med. Genet. A. *161*, 3063–3071.

Lewis, B.A., Kim, T.K., and Orkin, S.H. (2000). A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. Proc. Natl. Acad. Sci. U. S. A. *97*, 7172–7177.

Li, L., Bales, E.S., Peterson, C.A., and Legerski, R.J. (1993). Characterization of molecular defects in xeroderma pigmentosum group C. Nat. Genet. *5*, 413–417.

Li, X., Urwyler, O., and Suter, B. (2010). Drosophila Xpd regulates Cdk7 localization, mitotic kinase activity, spindle dynamics, and chromosome segregation. PLoS Genet. *6*, e1000876.

Li, Y., Flanagan, P.M., Tschochner, H., and Kornberg, R.D. (1994). RNA polymerase II initiation factor interactions and transcription start site selection. Science *263*, 805–807.

Lim, C.Y., Santoso, B., Boulay, T., Dong, E., Ohler, U., and Kadonaga, J.T. (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev. *18*, 1606–1617.

Linder, T., Zhu, X., Baraznenok, V., and Gustafsson, C.M. (2006). The classical srb4-138 mutant allele causes dissociation of yeast Mediator. Biochem. Biophys. Res. Commun. *349*, 948–953.

Liu, J., Grogan, L., Nau, M.M., Allegra, C.J., Chu, E., and Wright, J.J. (2001). Physical interaction between p53 and primary response gene Egr-1. Int. J. Oncol. *18*, 863–870.

Loncle, N., Boube, M., Joulia, L., Boschiero, C., Werner, M., Cribbs, D.L., and Bourbon, H.-M. (2007). Distinct roles for Mediator Cdk8 module subunits in Drosophila development. EMBO J. *26*, 1045–1054.

Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc. Natl. Acad. Sci. U. S. A. *88*, 10004–10008.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Luijsterburg, M.S., Lindh, M., Acs, K., Vrouwe, M.G., Pines, A., van Attikum, H., Mullenders, L.H., and Dantuma, N.P. (2012). DDB2 promotes chromatin decondensation at UV-induced DNA damage. J. Cell Biol. *197*, 267–281.

Lujan, J.E., Carlin, M.E., and Lubs, H.A. (1984). A form of X-linked mental retardation with marfanoid habitus. Am. J. Med. Genet. *17*, 311–322.

Lyons, M.J., Graham, J.M., Neri, G., Hunter, A.G.W., Clark, R.D., Rogers, R.C., Moscarda, M., Boccuto, L., Simensen, R., Dodd, J., et al. (2009). Clinical experience in the evaluation of 30 patients with a prior diagnosis of FG syndrome. J. Med. Genet. *46*, 9–13.

Maden, M. (2007). Retinoic acid in the development, regeneration and

maintenance of the nervous system. Nat. Rev. Neurosci. *8*, 755–765.

Mahony, S., Mazzoni, E.O., McCuine, S., Young, R.A., Wichterle, H., and Gifford, D.K. (2011). Ligand-dependent dynamics of retinoic acid receptor binding during early neurogenesis. Genome Biol. *12*, R2.

Maldonado, E., Ha, I., Cortes, P., Weis, L., and Reinberg, D. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: role of transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. Mol. Cell. Biol. *10*, 6335–6347.

Malik, S., and Roeder, R.G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. Nat. Rev. Genet. *11*, 761–772.

Malik, S., Gu, W., Wu, W., Qin, J., and Roeder, R.G. (2000). The USA-derived transcriptional coactivator PC2 is a submodule of TRAP/SMCC and acts synergistically with other PCs. Mol. Cell *5*, 753–760.

Malik, S., Barrero, M.J., and Jones, T. (2007). Identification of a regulator of transcription elongation as an accessory factor for the human Mediator coactivator. Proc. Natl. Acad. Sci. U. S. A. *104*, 6182–6187.

Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., et al. (1995). The nuclear receptor superfamily: the second decade. Cell *83*, 835–839.

Martínez, F., Caro-Llopis, A., Roselló, M., Oltra, S., Mayo, S., Monfort, S., and Orellana, C. (2017). High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. J. Med. Genet. *54*, 87–92.

Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P.J., and Bootsma, D. (1994). Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. EMBO J. *13*, 1831–1843.

Matsui, T., Segall, J., Weil, P.A., and Roeder, R.G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. J. Biol. Chem. *255*, 11992–11996.

Maxon, M.E., Goodrich, J.A., and Tjian, R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. Genes Dev. *8*, 515–524.

Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., and Cramer, P. (2012). CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science *336*, 1723–1725.

McKune, K., Moore, P.A., Hull, M.W., and Woychik, N.A. (1995). Six human RNA polymerase subunits functionally substitute for their yeast counterparts. Mol. Cell. Biol. *15*, 6895–6900.

Meyer, K.D., Lin, S.-C., Bernecky, C., Gao, Y., and Taatjes, D.J. (2010). p53 activates transcription by directing structural shifts in Mediator. Nat. Struct. Mol. Biol. *17*, 753–760.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., et al. (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell *87*, 1261–1270.

Mocquet, V., Kropachev, K., Kolbanovskiy, M., Kolbanovskiy, A., Tapias, A., Cai, Y., Broyde, S., Geacintov, N.E., and Egly, J.-M. (2007). The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. EMBO J. *26*, 2923–2932.

Moghal, N. (2003). A component of the transcriptional mediator complex inhibits RAS-dependent vulval fate specification in C. elegans. Development *130*, 57–69.

Morgan, J.I., and Curran, T. (1991). Stimulus-transcription coupling in the nervous system: involvement of the inducible proto-oncogenes fos and jun. Annu.

Rev. Neurosci. *14*, 421–451.

Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L.H.F., and Fousteri, M.I. (2007). Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. Mol. Cell *27*, 311–323.

Mukhopadhyay, A., Kramer, J.M., Merkx, G., Lugtenberg, D., Smeets, D.F., Oortveld, M.A.W., Blokland, E.A.W., Agrawal, J., Schenck, A., van Bokhoven, H., et al. (2010). CDK19 is disrupted in a female patient with bilateral congenital retinal folds, microcephaly and mild mental retardation. Hum. Genet. *128*, 281–291.

Mukundan, B., and Ansari, A. (2011). Novel Role for Mediator Complex Subunit Srb5/Med18 in Termination of Transcription. J. Biol. Chem. *286*, 37053–37057.

Mukundan, B., and Ansari, A. (2013). Srb5/Med18-mediated Termination of Transcription Is Dependent on Gene Looping. J. Biol. Chem. *288*, 11384–11394.

Mullegama, S.V., Jensik, P., Li, C., Dorrani, N., UCLA Clinical Genomics Center, Kantarci, S., Blumberg, B., Grody, W.W., and Strom, S.P. (2017). Coupling clinical exome sequencing with functional characterization studies to diagnose a patient with familial Mediterranean fever and MED13L haploinsufficiency syndromes. Clin. Case Rep. *5*, 833–840.

Muncke, N., Jung, C., Rüdiger, H., Ulmer, H., Roeth, R., Hubert, A., Goldmuntz, E., Driscoll, D., Goodship, J., Schön, K., et al. (2003). Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). Circulation *108*, 2843–2850.

Näär, A.M., Beaurang, P.A., Zhou, S., Abraham, S., Solomon, W., and Tjian, R. (1999). Composite co-activator ARC mediates chromatin-directed transcriptional activation. Nature *398*, 828–832.

Nagy, Z., Riss, A., Fujiyama, S., Krebs, A., Orpinell, M., Jansen, P., Cohen, A., Stunnenberg, H.G., Kato, S., and Tora, L. (2010). The metazoan ATAC and SAGA

coactivator HAT complexes regulate different sets of inducible target genes. Cell. Mol. Life Sci. CMLS *67*, 611–628.

Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., Hosseini, M., Behjati, F., Haas, S., Jamali, P., et al. (2011). Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature *478*, 57–63.

Narayanan, D.L., and Phadke, S.R. (2017). A novel variant in MED12 gene: Further delineation of phenotype. Am. J. Med. Genet. A. *173*, 2257–2260.

Narita, T., Narita, K., Takedachi, A., Saijo, M., and Tanaka, K. (2015). Regulation of Transcription Elongation by the XPG-TFIIH Complex Is Implicated in Cockayne Syndrome. Mol. Cell. Biol. *35*, 3178–3188.

Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugasawa, K., and Hanaoka, F. (2005). Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. Mol. Cell. Biol. *25*, 5664–5674.

Nogi, Y., and Fukasawa, T. (1980). A novel mutation that affects utilization of galactose in Saccharomyces cerevisiae. Curr. Genet. *2*, 115–120.

Nozawa, K., Schneider, T.R., and Cramer, P. (2017). Core Mediator structure at 3.4 Å extends model of transcription initiation complex. Nature *545*, 248–251.

Ohdo, S., Madokoro, H., Sonoda, T., and Hayakawa, K. (1986). Mental retardation associated with congenital heart disease, blepharophimosis, blepharoptosis, and hypoplastic teeth. J. Med. Genet. *23*, 242–244.

Ohkuma, Y., and Roeder, R.G. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. Nature *368*, 160–163.

Ohler, U., Liao, G., Niemann, H., and Rubin, G.M. (2002). Computational analysis of core promoters in the Drosophila genome. Genome Biol. *3*, RESEARCH0087.

Oksenych, V., Bernardes de Jesus, B., Zhovmer, A., Egly, J.-M., and Coin, F.

(2009). Molecular insights into the recruitment of TFIIH to sites of DNA damage. EMBO J. *28*, 2971–2980.

Okuda, M., Tanaka, A., Arai, Y., Satoh, M., Okamura, H., Nagadoi, A., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2004). A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. J. Biol. Chem. *279*, 51395–51403.

Opitz, J.M., and Kaveggia, E.G. (1974). Studies of malformation syndromes of man 33: the FG syndrome. An X-linked recessive syndrome of multiple congenital anomalies and mental retardation. Z. Für Kinderheilkd. *117*, 1–18.

Ormondroyd, E., de la Luna, S., and La Thangue, N.B. (1995). A new member of the DP family, DP-3, with distinct protein products suggests a regulatory role for alternative splicing in the cell cycle transcription factor DRTF1/E2F. Oncogene *11*, 1437–1446.

Parry, T.J., Theisen, J.W.M., Hsu, J.-Y., Wang, Y.-L., Corcoran, D.L., Eustice, M., Ohler, U., and Kadonaga, J.T. (2010). The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. Genes Dev. *24*, 2013–2018.

Patil, S.J., Somashekar, P.H., Shukla, A., Siddaiah, S., Bhat, V., Girisha, K.M., and Rao, P.N. (2017). Clinical Variability in Familial X-Linked Ohdo Syndrome-Maat-Kievit-Brunner Type with MED12 Mutation. J. Pediatr. Genet. *6*, 198–204.

Peippo, M., and Ignatius, J. (2011). Pitt-Hopkins Syndrome. Mol. Syndromol.

Pérez-Cadahía, B., Drobic, B., and Davie, J.R. (2011). Activation and function of immediate-early genes in the nervous system. Biochem. Cell Biol. Biochim. Biol. Cell. *89*, 61–73.

Perissi, V., Aggarwal, A., Glass, C.K., Rose, D.W., and Rosenfeld, M.G. (2004). A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. Cell *116*, 511–526.

Perissi, V., Jepsen, K., Glass, C.K., and Rosenfeld, M.G. (2010). Deconstructing repression: evolving models of co-repressor action. Nat. Rev. Genet. *11*, 109–123.

Pham, A.D., and Sauer, F. (2000). Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila. Science *289*, 2357–2360.

Philibert, R.A., Bohle, P., Secrest, D., Deaderick, J., Sandhu, H., Crowe, R., and Black, D.W. (2007). The association of the HOPA12bp polymorphism with schizophrenia in the NIMH genetics initiative for schizophrenia sample. Am. J. Med. Genet. B Neuropsychiatr. Genet. *144B*, 743–747.

Pikaard, C.S., Haag, J.R., Ream, T., and Wierzbicki, A.T. (2008). Roles of RNA polymerase IV in gene silencing. Trends Plant Sci. *13*, 390–397.

Plaschka, C., Larivière, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E.V., Borchers, C.H., Baumeister, W., Herzog, F., et al. (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. Nature *518*, 376–380.

Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. Nature *533*, 353–358.

Prescott, T.E., Kulseth, M.A., Heimdal, K.R., Stadheim, B., Hopp, E., Gambin, T., Coban Akdemir, Z.H., Jhangiani, S.N., Muzny, D.M., Gibbs, R.A., et al. (2016). Two male sibs with severe micrognathia and a missense variant in MED12. Eur. J. Med. Genet. *59*, 367–372.

Prontera, P., Ottaviani, V., Rogaia, D., Isidori, I., Mencarelli, A., Malerba, N., Cocciadiferro, D., Rolph, P., Stangoni, G., Vulto-van Silfhout, A., et al. (2016). A novel MED12 mutation: Evidence for a fourth phenotype. Am. J. Med. Genet. A. *170*, 2377–2382.

Rachez, C., Suldan, Z., Ward, J., Chang, C.P., Burakov, D., Erdjument-Bromage, H., Tempst, P., and Freedman, L.P. (1998). A novel protein complex that

interacts with the vitamin D3 receptor in a ligand-dependent manner and enhances VDR transactivation in a cell-free system. Genes Dev. *12*, 1787–1800.

Rau, M.J., Fischer, S., and Neumann, C.J. (2006). Zebrafish Trap230/Med12 is required as a coactivator for Sox9-dependent neural crest, cartilage and ear development. Dev. Biol. *296*, 83–93.

Raveh, T., and Kimchi, A. (2001). DAP kinase-a proapoptotic gene that functions as a tumor suppressor. Exp. Cell Res. *264*, 185–192.

Redin, C., Gérard, B., Lauer, J., Herenger, Y., Muller, J., Quartier, A., Masurel-Paulet, A., Willems, M., Lesca, G., El-Chehadeh, S., et al. (2014). Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. J. Med. Genet. *51*, 724–736.

Reinberg, D., and Roeder, R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. J. Biol. Chem. *262*, 3310–3321.

Reinberg, D., Horikoshi, M., and Roeder, R.G. (1987). Factors involved in specific transcription in mammalian RNA polymerase II. Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. J. Biol. Chem. *262*, 3322–3330.

Riedl, T., Hanaoka, F., and Egly, J.-M. (2003). The comings and goings of nucleotide excision repair factors on damaged DNA. EMBO J. *22*, 5293–5303.

Risheg, H., Graham, J.M., Clark, R.D., Rogers, R.C., Opitz, J.M., Moeschler, J.B., Peiffer, A.P., May, M., Joseph, S.M., Jones, J.R., et al. (2007). A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. Nat. Genet. *39*, 451–453.

Riss, A., Scheer, E., Joint, M., Trowitzsch, S., Berger, I., and Tora, L. (2015). Subunits of ADA-two-A-containing (ATAC) or Spt-Ada-Gcn5-acetyltrasferase (SAGA) Coactivator Complexes Enhance the Acetyltransferase Activity of GCN5. J. Biol. Chem. *290*, 28997–29009.

Robinson, P.J., Trnka, M.J., Pellarin, R., Greenberg, C.H., Bushnell, D.A., Davis, R., Burlingame, A.L., Sali, A., and Kornberg, R.D. (2015). Molecular architecture of the yeast Mediator complex. eLife *4*.

Robinson, P.J., Trnka, M.J., Bushnell, D.A., Davis, R.E., Mattei, P.-J., Burlingame, A.L., and Kornberg, R.D. (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. Cell *166*, 1411–1422.e16.

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the “30-nm” chromatin fiber: evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. U. S. A. *103*, 6506–6511.

Robinson, P.J.J., Bushnell, D.A., Trnka, M.J., Burlingame, A.L., and Kornberg, R.D. (2012). Structure of the mediator head module bound to the carboxy-terminal domain of RNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. *109*, 17931–17935.

Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.M., and Chambon, P. (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. Cell *90*, 97–107.

Rodríguez-Navarro, S. (2009). Insights into SAGA function during gene expression. EMBO Rep. *10*, 843–850.

Roeder, R.G., and Rutter, W.J. (1970). Specific nucleolar and nucleoplasmic RNA polymerases. Proc. Natl. Acad. Sci. U. S. A. *65*, 675–682.

Rump, P., Niessen, R., Verbruggen, K., Brouwer, O., de Raad, M., and Hordijk, R. (2011). A novel mutation in MED12 causes FG syndrome (Opitz-Kaveggia syndrome). Clin. Genet. *79*, 183–188.

Ryu, S., Zhou, S., Ladurner, A.G., and Tjian, R. (1999). The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. Nature *397*, 446–450.

Sakai, A., Shimizu, Y., and Hishinuma, F. (1988). Isolation and characterization of mutants which show an oversecretion phenotype in Saccharomyces cerevisiae. Genetics *119*, 499–506.

Sakai, M., Tujimura-Hayakawa, T., Yagi, T., Yano, H., Mitsushima, M., Unoki-Kubota, H., Kaburagi, Y., Inoue, H., Kido, Y., Kasuga, M., et al. (2016). The GCN5-CITED2-PKA signalling module controls hepatic glucose metabolism through a cAMP-induced substrate switch. Nat. Commun. *7*, 13147.

Sakurai, H., and Fukasawa, T. (1997). Yeast Gal11 and Transcription Factor IIE Function through a Common Pathway in Transcriptional Regulation. J. Biol. Chem. *272*, 32663–32669.

Sakurai, H., and Fukasawa, T. (2000). Functional Connections between Mediator Components and General Transcription Factors of Saccharomyces cerevisiae. J. Biol. Chem. *275*, 37251–37256.

Samuels, M., Fire, A., and Sharp, P.A. (1982). Separation and characterization of factors mediating accurate transcription by RNA polymerase II. J. Biol. Chem. *257*, 14419–14427.

Sandrock, B., and Egly, J.M. (2001). A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIH. J. Biol. Chem. *276*, 35328–35333.

Sato, S., Tomomori-Sato, C., Parmely, T.J., Florens, L., Zybailov, B., Swanson, S.K., Banks, C.A.S., Jin, J., Cai, Y., Washburn, M.P., et al. (2004). A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. Mol. Cell *14*, 685–691.

Sato, S., Tomomori-Sato, C., Tsai, K.-L., Yu, X., Sardiu, M., Saraf, A., Washburn, M.P., Florens, L., Asturias, F.J., Conaway, R.C., et al. (2016). Role for the MED21-MED7 Hinge in Assembly of the Mediator-RNA Polymerase II Holoenzyme. J. Biol. Chem. *291*, 26886–26898.

Sawadogo, M., and Roeder, R.G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell *43*, 165–175.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H., Chambon, P., and Egly, J.M. (1993). DNA repair helicase: a

component of BTF2 (TFIIH) basic transcription factor. Science *260*, 58–63.

Schiano, C., Casamassimi, A., Vietri, M.T., Rienzo, M., and Napoli, C. (2014a). The roles of mediator complex in cardiovascular diseases. Biochim. Biophys. Acta *1839*, 444–451.

Schiano, C., Casamassimi, A., Rienzo, M., de Nigris, F., Sommese, L., and Napoli, C. (2014b). Involvement of Mediator complex in malignancy. Biochim. Biophys. Acta *1845*, 66–83.

Schmitz, K.-M., Schmitt, N., Hoffmann-Rohrer, U., Schäfer, A., Grummt, I., and Mayer, C. (2009). TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. Mol. Cell *33*, 344–353.

Schreck, I., Grico, N., Hansjosten, I., Marquardt, C., Bormann, S., Seidel, A., Kvietkova, D.L., Pieniazek, D., Segerbäck, D., Diabaté, S., et al. (2016). The nucleotide excision repair protein XPC is essential for bulky DNA adducts to promote interleukin-6 expression via the activation of p38-SAPK. Oncogene *35*, 908–918.

Schroeder, S.C., Schwer, B., Shuman, S., and Bentley, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev. *14*, 2435–2440.

Schulze, A., Zerfass, K., Spitkovsky, D., Middendorp, S., Bergès, J., Helin, K., Jansen-Dürr, P., and Henglein, B. (1995). Cell cycle regulation of the cyclin A gene promoter is mediated by a variant E2F site. Proc. Natl. Acad. Sci. U. S. A. *92*, 11264–11268.

Schwartz, C.E., Tarpey, P.S., Lubs, H.A., Verloes, A., May, M.M., Risheg, H., Friez, M.J., Futreal, P.A., Edkins, S., Teague, J., et al. (2007). The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. J. Med. Genet. *44*, 472–477.

Schwertman, P., Vermeulen, W., and Marteijn, J.A. (2013). UVSSA and USP7, a new couple in transcription-coupled DNA repair. Chromosoma *122*, 275–284.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1993). Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. Nature *363*, 371–374.

Shannon, S.R., Moise, A.R., and Trainor, P.A. (2017). New insights and changing paradigms in the regulation of vitamin A metabolism in development. Wiley Interdiscip. Rev. Dev. Biol. *6*.

Shao, H., Revach, M., Moshonov, S., Tzuman, Y., Gazit, K., Albeck, S., Unger, T., and Dikstein, R. (2005). Core promoter binding by histone-like TAF complexes. Mol. Cell. Biol. *25*, 206–219.

Shin, C.H., Chung, W.-S., Hong, S.-K., Ober, E.A., Verkade, H., Field, H.A., Huisken, J., and Y. R. Stainier, D. (2008). Multiple roles for Med12 in vertebrate endoderm development. Dev. Biol. *317*, 467–479.

Singh, N., and Han, M. (1995). sur-2, a novel gene, functions late in the let-60 ras-mediated signaling pathway during Caenorhabditis elegans vulval induction. Genes Dev. *9*, 2251–2265.

Singh, R.K., and Dagnino, L. (2016). E2F1 interactions with hHR23A inhibit its degradation and promote DNA repair. Oncotarget *7*, 26275–26292.

Singh, A., Compe, E., Le May, N., and Egly, J.-M. (2015). TFIIH subunit alterations causing xeroderma pigmentosum and trichothiodystrophy specifically disturb several steps during transcription. Am. J. Hum. Genet. *96*, 194–207.

Smale, S.T., and Baltimore, D. (1989). The “initiator” as a transcription control element. Cell *57*, 103–113.

Søgaard, T.M.M., and Svejstrup, J.Q. (2007). Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. J. Biol. Chem. *282*, 14113–14120.

Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C., and Werner, M. (2011). Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo. Science *331*, 1451–1454.

Spedale, G., Timmers, H.T.M., and Pijnappel, W.W.M.P. (2012). ATAC-king the complexity of SAGA during evolution. Genes Dev. *26*, 527–541.

Spotswood, H.T., and Turner, B.M. (2002). An increasingly complex code. J. Clin. Invest. *110*, 577–582.

Stadelmann, B., Heyer, W.D., Thuriaux, P., Kohli, J., Smith, C., Agris, P.F., Kuo, K.C., and Gehrke, C. (1986). Antisuppressor mutations and sulfur-carrying nucleosides in transfer RNAs of Schizosaccharomyces pombe. J. Biol. Chem. *261*, 16351–16355.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490–495.

Stevens, J.L. (2002). Transcription Control by E1A and MAP Kinase Pathway via Sur2 Mediator Subunit. Science *296*, 755–758.

Sugasawa, K. (2010). Regulation of damage recognition in mammalian global genomic nucleotide excision repair. Mutat. Res. *685*, 29–37.

Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. Genes Dev. *15*, 507–521.

Sugasawa, K., Shimizu, Y., Iwai, S., and Hanaoka, F. (2002). A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. DNA Repair *1*, 95–107.

Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., Mori, T., Iwai, S., Tanaka, K., Tanaka, K., et al. (2005). UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. Cell *121*, 387–400.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat. Struct. Biol. *7*, 1121–1124.

Syed, S.H., Goutte-Gattat, D., Becker, N., Meyer, S., Shukla, M.S., Hayes, J.J., Everaers, R., Angelov, D., Bednar, J., and Dimitrov, S. (2010). Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. Proc. Natl. Acad. Sci. U. S. A. *107*, 9620–9625.

Taatjes, D.J., Schneider-Poetsch, T., and Tjian, R. (2004). Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. Nat. Struct. Mol. Biol. *11*, 664–671.

Takagi, Y., Calero, G., Komori, H., Brown, J.A., Ehrensberger, A.H., Hudmon, A., Asturias, F., and Kornberg, R.D. (2006). Head Module Control of Mediator Interactions. Mol. Cell *23*, 355–364.

Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A.S., Kong, S.E., Szutorisz, H., Swanson, S.K., Martin-Brown, S., Washburn, M.P., et al. (2011). Human mediator subunit MED26 functions as a docking site for transcription elongation factors. Cell *146*, 92–104.

Takedachi, A., Saijo, M., and Tanaka, K. (2010). DDB2 complex-mediated ubiquitylation around DNA damage is oppositely regulated by XPC and Ku and contributes to the recruitment of XPA. Mol. Cell. Biol. *30*, 2708–2723.

Talbert, P.B., and Henikoff, S. (2010). Histone variants--ancient wrap artists of the epigenome. Nat. Rev. Mol. Cell Biol. *11*, 264–275.

Talbert, P.B., Ahmad, K., Almouzni, G., Ausió, J., Berger, F., Bhalla, P.L., Bonner, W.M., Cande, W.Z., Chadwick, B.P., Chan, S.W.L., et al. (2012). A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin *5*, 7.

Tanaka, A., Watanabe, T., Iida, Y., Hanaoka, F., and Ohkuma, Y. (2009). Central forkhead domain of human TFIIE beta plays a primary role in binding double-stranded DNA at transcription initiation. Genes Cells Devoted Mol. Cell. Mech. *14*, 395–405.

Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. Proc. Natl. Acad. Sci. U. S. A. *105*, 5160–5165.

Taubert, S., Gorrini, C., Frank, S.R., Parisi, T., Fuchs, M., Chan, H.-M., Livingston, D.M., and Amati, B. (2004). E2F-dependent histone acetylation and recruitment of the Tip60 acetyltransferase complex to chromatin in late G1. Mol. Cell. Biol. *24*, 4546–4556.

Thaler, F., and Mercurio, C. (2014). Towards selective inhibition of histone deacetylase isoforms: what has been achieved, where we are and what will be next. ChemMedChem *9*, 523–526.

Theisen, J.W.M., Lim, C.Y., and Kadonaga, J.T. (2010). Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. Mol. Cell. Biol. *30*, 3471–3479.

Thompson, C.M., and Young, R.A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. Proc. Natl. Acad. Sci. U. S. A. *92*, 4587–4590.

Thompson, C.M., Koleske, A.J., Chao, D.M., and Young, R.A. (1993). A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. Cell *73*, 1361–1375.

Tokusumi, Y., Ma, Y., Song, X., Jacobson, R.H., and Takada, S. (2007). The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. Mol. Cell. Biol. *27*, 1844–1858.

Traboulsi, H., Davoli, S., Catez, P., Egly, J.-M., and Compe, E. (2014). Dynamic partnership between TFIIH, PGC-1α and SIRT1 is impaired in trichothiodystrophy. PLoS Genet. *10*, e1004732.

Trehan, A., Brady, J.M., Maduro, V., Bone, W.P., Huang, Y., Golas, G.A., Kane, M.S., Lee, P.R., Thurm, A., Gropman, A.L., et al. (2015). MED23-associated intellectual disability in a non-consanguineous family. Am. J. Med. Genet. A. *167*, 1374–1380.

Tremeau-Bravard, A., Riedl, T., Egly, J.-M., and Dahmus, M.E. (2004). Fate of RNA polymerase II stalled at a cisplatin lesion. J. Biol. Chem. *279*, 7751–7759.

Tsai, K.-L., Tomomori-Sato, C., Sato, S., Conaway, R.C., Conaway, J.W., and Asturias, F.J. (2014). Subunit Architecture and Functional Modular Rearrangements of the Transcriptional Mediator Complex. Cell *157*, 1430–1444.

Tsai, K.-L., Yu, X., Gopalan, S., Chao, T.-C., Zhang, Y., Florens, L., Washburn, M.P., Murakami, K., Conaway, R.C., Conaway, J.W., et al. (2017). Mediator structure and rearrangements required for holoenzyme formation. Nature *544*, 196–201.

Tullai, J.W., Schaffer, M.E., Mullenbrock, S., Sholder, G., Kasif, S., and Cooper, G.M. (2007). Immediate-Early and Delayed Primary Response Genes are Distinct in Function and Genomic Architecture. J. Biol. Chem. *282*, 23981–23995.

Turner, B.M. (2005). Reading signals on the nucleosome with a new nomenclature for modified histones. Nat. Struct. Mol. Biol. *12*, 110–112.

Tzschach, A., Grasshoff, U., Beck-Woedl, S., Dufke, C., Bauer, C., Kehrer, M., Evers, C., Moog, U., Oehl-Jaschkowitz, B., Di Donato, N., et al. (2015). Next-generation sequencing in X-linked intellectual disability. Eur. J. Hum. Genet. EJHG *23*, 1513–1518.

Utami, K.H., Winata, C.L., Hillmer, A.M., Aksoy, I., Long, H.T., Liany, H., Chew, E.G.Y., Mathavan, S., Tay, S.K.H., Korzh, V., et al. (2014). Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. Hum. Mutat. *35*, 1311–1320.

Vandel, L., and Kouzarides, T. (1999). Residues phosphorylated by TFIIH are required for E2F-1 degradation during S-phase. EMBO J. *18*, 4280–4291.

van Essen, D., Engist, B., Natoli, G., and Saccani, S. (2009). Two Modes of Transcriptional Activation at Native Promoters by NF-κB p65. PLoS Biol. *7*, e73.

Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. Nat. Rev. Mol. Cell Biol. *16*, 178–189.

Verloes, A., Bremond-Gignac, D., Isidor, B., David, A., Baumann, C., Leroy, M.-A., Stevens, R., Gillerot, Y., Héron, D., Héron, B., et al. (2006). Blepharophimosis-

mental retardation (BMR) syndromes: A proposed clinical classification of the so-called Ohdo syndrome, and delineation of two new BMR syndromes, one X-linked and one autosomal recessive. Am. J. Med. Genet. A. *140A*, 1285–1296.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J. *26*, 2041–2051.

Vodopiutz, J., Schmook, M.T., Konstantopoulou, V., Plecko, B., Greber-Platzer, S., Creus, M., Seidl, R., and Janecke, A.R. (2015). MED20 mutation associated with infantile basal ganglia degeneration and brain atrophy. Eur. J. Pediatr. *174*, 113–118.

Volker, M., Moné, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H., van Driel, R., van Zeeland, A.A., and Mullenders, L.H. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. Mol. Cell *8*, 213–224.

Vulto-van Silfhout, A.T., de Vries, B.B.A., van Bon, B.W.M., Hoischen, A., Ruiterkamp-Versteeg, M., Gilissen, C., Gao, F., van Zwam, M., Harteveld, C.L., van Essen, A.J., et al. (2013). Mutations in MED12 cause X-linked Ohdo syndrome. Am. J. Hum. Genet. *92*, 401–406.

Wang, G., Balamotis, M.A., Stevens, J.L., Yamaguchi, Y., Handa, H., and Berk, A.J. (2005). Mediator Requirement for Both Recruitment and Postrecruitment Steps in Transcription Initiation. Mol. Cell *17*, 683–694.

Wang, W., Yao, X., Huang, Y., Hu, X., Liu, R., Hou, D., Chen, R., and Wang, G. (2013). Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. Transcription *4*, 39–51.

Wang, X., Yang, N., Uno, E., Roeder, R.G., and Guo, S. (2006). A subunit of the mediator complex regulates vertebrate neuronal development. Proc. Natl. Acad. Sci. *103*, 17284–17289.

Watanabe, T., Hayashi, K., Tanaka, A., Furumoto, T., Hanaoka, F., and Ohkuma, Y. (2003). The carboxy terminus of the small subunit of TFIIE regulates the

transition from transcription initiation to elongation by RNA polymerase II. Mol. Cell. Biol. *23*, 2914–2926.

Watt, F., and Molloy, P.L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev. *2*, 1136–1143.

Weber, A.R., Krawczyk, C., Robertson, A.B., Kuśnierczyk, A., Vågbø, C.B., Schuermann, D., Klungland, A., and Schär, P. (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nat. Commun. *7*, 10806.

Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. Cell *18*, 469–484.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science *303*, 1014–1016.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

Widom, J., and Klug, A. (1985). Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. Cell *43*, 207–213.

Willemsen, J., Wicht, O., Wolanski, J.C., Baur, N., Bastian, S., Haas, D.A., Matula, P., Knapp, B., Meyniel-Schicklin, L., Wang, C., et al. (2017). Phosphorylation-Dependent Feedback Inhibition of RIG-I by DAPK1 Identified by Kinome-wide siRNA Screening. Mol. Cell *65*, 403–415.e8.

Wind, M., and Reines, D. (2000). Transcription elongation factor SII. BioEssays News Rev. Mol. Cell. Dev. Biol. *22*, 327–336.

Winkles, J.A. (1998). Serum- and polypeptide growth factor-inducible gene expression in mouse fibroblasts. Prog. Nucleic Acid Res. Mol. Biol. *58*, 41–78.

Wirbelauer, C., Bell, O., and Schübeler, D. (2005). Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. Genes Dev. *19*, 1761–1766.

Wong, K.H., Jin, Y., and Struhl, K. (2014). TFIIH phosphorylation of the Pol II CTD stimulates mediator dissociation from the preinitiation complex and promoter escape. Mol. Cell *54*, 601–612.

Woychik, N.A., and Young, R.A. (1990). RNA polymerase II: subunit structure and function. Trends Biochem. Sci. *15*, 347–351.

Wu, Y., Ma, S., Xia, Y., Lu, Y., Xiao, S., Cao, Y., Zhuang, S., Tan, X., Fu, Q., Xie, L., et al. (2017). Loss of GCN5 leads to increased neuronal apoptosis by upregulating E2F1- and Egr-1-dependent BH3-only protein Bim. Cell Death Dis. *8*, e2570.

Xu, X., Zhou, H., and Boyer, T.G. (2011). Mediator is a transducer of amyloid-precursor-protein-dependent nuclear signalling. EMBO Rep. *12*, 216–222.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Mol. Cell *21*, 227–237.

Yamamoto, T., and Shimojima, K. (2015). A novel MED12 mutation associated with non-specific X-linked intellectual disability. Hum. Genome Var. *2*, 15018.

Yamamoto, T., Shimojima, K., Ondo, Y., Shimakawa, S., and Okamoto, N. (2017). MED13L haploinsufficiency syndrome: A de novo frameshift and recurrent intragenic deletions due to parental mosaicism. Am. J. Med. Genet. A. *173*, 1264–1269.

Yan, Q., Moreland, R.J., Conaway, J.W., and Conaway, R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. J. Biol. Chem. *274*, 35668–35675.

Yao, G., Lee, T.J., Mori, S., Nevins, J.R., and You, L. (2008). A bistable Rb-

E2F switch underlies the restriction point. Nat. Cell Biol. *10*, 476–482.

Yao, P.J., Petralia, R.S., and Mattson, M.P. (2016). Sonic Hedgehog Signaling and Hippocampal Neuroplasticity. Trends Neurosci. *39*, 840–850.

Yarden, G., Elfakess, R., Gazit, K., and Dikstein, R. (2009). Characterization of sINR, a strict version of the Initiator core promoter element. Nucleic Acids Res. *37*, 4234–4246.

Yin, J., and Wang, G. (2014). The Mediator complex: a master coordinator of transcription and cell lineage development. Dev. Camb. Engl. *141*, 977–987.

Yin, Y.-W., Jin, H.-J., Zhao, W., Gao, B., Fang, J., Wei, J., Zhang, D.D., Zhang, J., and Fang, D. (2015). The Histone Acetyltransferase GCN5 Expression Is Elevated and Regulated by c-Myc and E2F1 Transcription Factors in Human Colon Cancer. Gene Expr. *16*, 187–196.

Yu, Y., Song, C., Zhang, Q., DiMaggio, P.A., Garcia, B.A., York, A., Carey, M.F., and Grunstein, M. (2012). Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA. Mol. Cell *46*, 7–17.

Zehavi, Y., Sloutskin, A., Kuznetsov, O., and Juven-Gershon, T. (2014). The core promoter composition establishes a new dimension in developmental gene networks. Nucl. Austin Tex *5*, 298–303.

Zhang, C., and Burton, Z.F. (2004). Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. J. Mol. Biol. *342*, 1085–1099.

Zhang, F., Lin, M., Abidi, P., Thiel, G., and Liu, J. (2003). Specific interaction of Egr1 and c/EBPbeta leads to the transcriptional activation of the human low density lipoprotein receptor gene. J. Biol. Chem. *278*, 44246–44254.

Zhang, H.S., Gavin, M., Dahiya, A., Postigo, A.A., Ma, D., Luo, R.X., Harbour, J.W., and Dean, D.C. (2000). Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. Cell *101*, 79–89.

Zhao, M., Geng, R., Guo, X., Yuan, R., Zhou, X., Zhong, Y., Huo, Y., Zhou, M., Shen, Q., Li, Y., et al. (2017). PCAF/GCN5-Mediated Acetylation of RPA1 Promotes Nucleotide Excision Repair. Cell Rep. *20*, 1997–2009.

Zhou, H., Kim, S., Ishii, S., and Boyer, T.G. (2006). Mediator Modulates Gli3-Dependent Sonic Hedgehog Signaling. Mol. Cell. Biol. *26*, 8667–8682.

Zhou, H., Spaeth, J.M., Kim, N.H., Xu, X., Friez, M.J., Schwartz, C.E., and Boyer, T.G. (2012). MED12 mutations link intellectual disability syndromes with dysregulated GLI3-dependent Sonic Hedgehog signaling. Proc. Natl. Acad. Sci. *109*, 19763–19768.

Zhou, M., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J., Price, D.H., and Brady, J.N. (2000). Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. Mol. Cell. Biol. *20*, 5077–5086.

Zhu, X., Zhang, Y., Bjornsdottir, G., Liu, Z., Quan, A., Costanzo, M., Davila Lopez, M., Westholm, J.O., Ronne, H., Boone, C., et al. (2011). Histone modifications influence mediator interactions with chromatin. Nucleic Acids Res. *39*, 8342–8354.

Zile, M.H. (2001). Function of vitamin A in vertebrate embryonic development. J. Nutr. *131*, 705–708.

Baptiste BIDON

Médiateur et facteurs NER lors de l'initiation de la transcription

Résumé

La synthèse d'ARN messagers résulte d'une cascade d'évènements temporellement et spatialement orchestrée. Au moment de l'initiation de la transcription, divers facteurs tels que les facteurs généraux de transcription, le complexe Médiateur, des co-activateurs, des facteurs de remodelage de la chromatine ainsi que l'ARN polymérase II sont recrutés au niveau de la région promotrice du gène. Certains facteurs de la voie NER de réparation de l'ADN sont également recrutés. En utilisant des cellules de patients porteurs de mutations dans les gènes *MED12* (sous-unité du Médiateur) ou *XPC* (facteur initiant la voie NER), nous avons pu étudier le rôle de ces protéines dans la transcription. Les patients MED12 sont notamment caractérisés par une lourde déficience intellectuelle et des malformations congénitales. Nous avons montré que MED12 est impliqué dans le contrôle de certains gènes de réponse immédiate comme JUN, qui contribue notamment au développent et à la plasticité cérébrale. L'expression de ce dernier est affectée par les mutations de MED12, mais différemment en fonction de la position de la mutation, apportant une possible indication sur l'origine des variations phénotypiques observées chez les patients. En parallèle, les patients XPC se caractérisent par une forte photosensibilité. Nous avons montré que la protéine XPC, en collaboration avec le facteur E2F1, est impliquée dans le recrutement de l'histone acetyl-transférase GCN5 au niveau du promoteur d'un certain nombre de gènes. Cette dernière permet notamment l'a modification de l'environnement chromatinien, en coopération avec le facteur général de transcription TFIIH et participe ainsi à l'initiation de la transcription. En plus d'approfondir la compréhension des mécanismes régissant la transcription, ces résultats ont permis de mieux comprendre l'étiologie des maladies associées aux mutations.

Abstract

The synthesis of messenger RNA is a highly regulated process. During transcription initiation, a large number of proteins are recruited to gene promoter, including the RNA polymerase II, general transcription factors, co-activators, chromatin remodellers and the Mediator complex. Some DNA repair factors from the NER pathway are also recruited. Using cells derived from patients bearing mutations in either *MED12* gene or *XPC* gene, we studied the roles of such proteins in transcription. MED12 patients are mostly characterised by intellectual disability and developmental delay. We showed that MED12 is implicated in the transcription regulation of immediate early genes like JUN, known for its role in neurological development and neuronal plasticity. JUN expression is markedly altered by *MED12* mutations. We also showed that the position of the mutation influences this alteration, bringing possible explanation for inter-patients symptom variability. Meanwhile, XPC patients are mostly characterized by photosensitivity. We showed that XPC protein, which engages one of the NER pathways, is implicated in chromatin post-translational modification. Together with E2F1, it helps the recruitment of GCN5 acetyl-transferase to promoter of a certain set of genes. On the promoter, GCN5 notably cooperates with TFIIH to modify the chromatin environment during transcription initiation. In addition to help the comprehension of the transcription mechanisms, these results bring knew insight into the aetiology of mutations associated diseases.