UNIVERSITÉ DE STRASBOURG

*ÉCOLE DOCTORALE « des Sciences de la vie et de la santé »*

IGBMC – Inserm U 964 – CNRS UMR 7104

# THÈSE présentée par :

**Lise-Marie DONNIO**

Soutenue le : **15 décembre 2014**

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

Discipline/Spécialité : **Sciences du Vivant/Aspects Moléculaires et Cellulaires de la Biologie**

## Étude transcriptionnelle des mutations dans le Médiateur ou dans son partenaire NIPBL à l'origine de maladies génétiques

**Thèse dirigée par :**

**Dr EGLY Jean-Marc** DRE INSERM, IGBMC, Strasbourg

**RAPPORTEURS :**

**Dr JALINOT Pierre** DR CNRS, ENS Lyon
**Dr BROUSSET Pierre** PR INSERM, CHU Toulouse
**Pr DOLLFUS Hélène** PR INSERM, CHRU Strasbourg

igbmc
institut de génétique et de biologie moléculaire et cellulaire

# ACKNOWLEDGEMENTS

*Je tiens tout d'abord à exprimer mes remerciements à Pierre JALINOT, Pierre BROUSSET et Hélène DOLLFUS de m'avoir fait l'honneur d'accepter d'être membres de mon jury et ainsi participer à la dernière étape de ma thèse.*

*Je remercie tout particulièrement Jean-Marc EGLY, pour m'avoir accueilli dans son laboratoire situé dans l'un des meilleurs centres scientifiques internationaux et pour avoir été mon directeur de thèse. Merci pour sa confiance tout au long de ces quatre années.*

*Merci à l'Association de la Recherche contre le Cancer et à la Ligue contre le cancer, ainsi que leurs nombreux donateurs pour avoir financés les deux dernières années de mes études doctorales.*

*Merci à tous les membres du laboratoire présents et passés (Annabel, Anton, Amita, Alex, Alexey, Baptiste, Cathy, Carlos, Charlotte, Christophe, Federico, Frédéric, Hussein, Izarn, Jikta, Manu, Maryssa, Marc, Nicolas, Philippe, Renier, Salim, Satoru, Serena, Sergey, Ulrik, Valérie, Yakov, Zita). Je remercie particulièrement Baptiste, mon compagnon de galère sur le projet Médiateur, pour les discussions scientifiques et bavardages en tous genres, ainsi que Manu pour toutes ses corrections. Un merci général pour l'ambiance (au labo et en dehors), ainsi que pour les nombreuses et vraiment très diverses discussions que nous avons eues. Je tiens à clamer haut et fort que le box 4025 est le plus cool.*

*Je remercie tous les membres des services communs de l'IGBMC et en particulier les membres du service de culture cellulaire.*

*Je remercie l'ensemble de mes amis (Julia, Christelle, Nico, Yvan, Céline, Elodie, Cyril, William, Laurence, Julien…etc) pour les nombreuses discussions sans rapport avec la recherche et les très nombreuses soirées jeux. Du côté musical, je n'oublie pas les membres des Sixtoys et de l'orchestre Vulcania, qui m'ont permis le temps d'une répétition ou d'un concert, d'oublier le quotidien du labo.*

*Finalement, et il va sans dire, je remercie de tout mon cœur ma famille (papa, maman, mes sœurs, mon petit frère, mes neveux et nièces) sans qui rien de tout cela n'aurait été possible, même si le sujet de ma thèse restera en partie un mystère pour eux.*

*En espérant n'avoir oublié personne… merci à tous !*

# CONTENTS


FIGURES AND TABLES .................... 4

ABBREVIATIONS .................... 5

RÉSUMÉ DE LA THÈSE EN FRANÇAIS .................... 7

INTRODUCTION .................... 14

I. Transcription of class II genes .................... 15

A. RNA polymerase II .................... 15

A.1. Subunit composition .................... 15

A.2. Structure of Pol II .................... 16

A.3. CTD and its modifications .................... 17

B. Regulatory sequence in the heart of DNA .................... 17

B.1. The core promoter .................... 17

*B.1.a. The TATA-box* .................... *18*

*B.1.b. The initiator element (Inr)* .................... *18*

*B.1.c. The TFIIB recognition element (BREu and BREd)* .................... *18*

*B.1.d. The Downstream promoter element (DPE)* .................... *18*

*B.1.e. The Motif ten element (MTE)* .................... *19*

*B.1.f. The downstream core element (DCE)* .................... *19*

*B.1.g. The X core promoter element 1 (XCPE1)* .................... *19*

B.2. Regulatory response element .................... 19

C. General transcription factors .................... 20

D. The transcription cycle .................... 22

D.1. Promoter binding .................... 22

D.2. Initiation and promoter clearance .................... 22

D.3. Proximal pausing and elongation .................... 24

D.4. Termination .................... 24

D.5. Re-initiation and gene looping .................... 25

E. The activated transcription .................... 28

E.1. Immediate early response genes .................... 28

E.2. Nuclear receptors regulated genes .................... 28

*E.2.a. Nuclear receptors structure* .................... *29*

*E.2.b. Nuclear receptor classification* .................... *30*

*E.2.c. Retinoic acids receptors* .................... *30*

*E.2.d. Transcription of RAR target genes* .................... *31*

*E.2.e. Vitamin D receptors* .................... *34*

II. REVIEW on the Mediator complex .................... 35

III. Supplemental information .................... 64

*Discovery of the Mediator complex* ........ 64
*Composition and structure of the Mediator complex* ........ 64
*The Mediator complex: a general transcription factor* ........ 64
*Cohesin complex* ........ 66

MATERIALS & METHODS ........ 69

*Cells culture* ........ 70
*Isolation of mouse embryonic fibroblasts* ........ 71
*Treatments* ........ 71
*RNA extraction, reverse transcription and real-time qPCR* ........ 71
*Chromatin Immunoprecipitation (ChIP)* ........ 72
*Western blot and co-immunoprecipitation* ........ 72
*Antibodies* ........ 73
*Primers* ........ 73
*Sequences alignment* ........ 74

RESULTS ........ 76

MED12 PROJECT: Different X-linked mental retardation disorders depending on the position of MED12 mutation. ........ 77
MED17 PROJECT: Infantile Cerebral and Cerebellar Atrophy and mutation p.L371P in *MED17* ........ 102
NIPBL PROJECT: Cornelia de Lange syndrome and NIPBL mutation ........ 110
MICE PROJECT: Mice with MED12/R961W, MED17/L371P or MED23/R617Q mutation ........ 116

CONCLUSIONS ........ 120

REFERENCES ........ 122

# FIGURES AND TABLES

**Figure 1:** Organisation du complexe Médiateur humain et des maladies génétiques associées à une mutation dans une de ses sous-unités. ........ 8

**Figure 2:** Représentation schématique de la protéine MED12 et des mutations associées. ........ 9

**Table 1:** Composition of RNA polymerase II and function of the different subunits. ........ 16

**Figure 3:** Eukaryotic promoter motifs. ........ 20

**Table 2:** General transcription factors. ........ 21

**Figure 4:** Pre-initiation complex assembly. ........ 23

**Figure 5:** RNA polymerase II transcription cycle ........ 27

**Figure 6:** structural organization of Nuclear Receptors ........ 29

**Figure 7:** Classification of Nuclear Receptors ........ 30

**Figure 8:** Classical model of the activation of retinoic acid target genes ........ 32

**Figure 9:** Recapitulation of the phosphorylation cascade induced by RA. ........ 33

**Table 3:** Mediator subunits found in different organisms ........ 65

**Figure 10:** Modular and subunit organization of yeast Mediator and comparison with human Mediator ........ 66

**Figure 11:** The cohesin complex and its regulators ........ 67

**Figure 12:** MED17/L371P mutation. ........ 103

**Figure 13:** Effect of Med17/L371P mutation on *MED17* expression and protein stability. ........ 104

**Figure 14:** Predicted effect of MED17/L371P mutation. ........ 105

**Figure 15:** Effect of Med17/L371P mutation on MED composition ........ 106

**Figure 16:** Dysregulation of gene expression in MED17/L371P cells after various stimuli ........ 108

**Figure 17:** Effect of *NIPBL* mutation on its gene expression. ........ 112

**Figure 18:** Effect of *NIPBL* mutation on the expression of IEGs. ........ 112

**Figure 19:** Effect of *NIPBL* mutation on the expression of VitD response genes. ........ 113

**Figure 20:** Effect of *NIPBL* mutation on the expression of RA-target genes. ........ 114

**Figure 21:** Binding of transcriptional factors at *RAR*$\beta$ promoter in physiological condition. ........ 115

**Figure 22:** Partial protein sequence alignment among multiple species around the residues MED12/R961, MED17/L371 and MED23/R617. ........ 117

**Figure 23:** embryos from a cross between Med23 heterozygous mice (+/R617Q). ........ 118

# ABBREVIATIONS

| | |
|---|---|
| **AICD** | Amyloid Precursor Protein Intracellular Domain |
| **ARC** | Activator Recruited Factor |
| **bp** | base pair |
| **BRE** | TFIIB recognition element |
| **CAK** | Cdk-Activated Kinase |
| **CDK** | Cyclin Dependent Kinase |
| **CdLS** | Cornelia de Lange Syndrome |
| **CF1** | Cleavage Factor 1 |
| **ChIP** | Chromatin Immunoprecipitation |
| **CMT** | Charcot-Marie-Tooth disease |
| **CPSF** | Cleavage and Polyadenylation Specificity Factor |
| **CRSP** | Cofactor Required for Sp1 activation |
| **CTCF** | CCCTC-binding Factor |
| **CTD** | C-Terminal Domain |
| **CS** | Cockayne Syndrome |
| **CtsF** | Cleavage stimulatory Factor |
| **DBD** | DNA binding domain |
| **DNA** | Deoxyribonucleic Acid |
| **DNMT1** | DNA (cytosine-5)-methyltransferase 1 |
| **DPE** | Downstream promoter element |
| **DRIP** | Vitamin D Receptor Interacting Protein |
| **DSIF** | DRB Sensitivity Inducing Factor |
| **EGR1** | Early Growth Response protein 1 |
| **EM** | Electron microscopy |
| **mES** | mouse Embryonic Stem cells |
| **GTF** | General Transcription Factor |
| **GR** | Glucocorticoid receptor |
| **HAT** | Histone Acetyltransferase |
| **HDAC** | Histone Deacetylase |
| **HP1** | Heterochromatin Protein 1 |
| **ID** | intellectual disability |
| **IEG** | Immediate Early Gene |
| **Inr** | Initiator |
| **LBD** | Ligand Binding Domain |
| **LCs** | Lymphoblastoids cells |
| **MAT1** | Ménage à Trois |
| **MED** | Mediator |

| | |
|---|---|
| **MEFs** | Mouse Embryonic Fibroblasts |
| **MudPIT** | Multidimensional Protein Identification Technology |
| **NAT** | Negative regulator of Activated Transcription |
| **NCoR** | Nuclear receptor co-repressor 1 |
| **ncRNA-a** | non-coding RNA-activating |
| **NELF** | Negative Elongation Factor |
| **NER** | Nucleotide Excision Repair |
| **NIPBL** | Nipped-B-like |
| **NR** | Nuclear Receptor |
| **PC2** | Positive Cofactor |
| **PIC** | Preinitiation Complex |
| **Pol II** | RNA polymerase II |
| **PTHS** | Pitt-Hopkins Syndrome |
| **PTOV** | Prostate Tumour Overexpressed protein 1 |
| **qPCR** | quantitative Polymerase Chain Reaction |
| **RAR** | Retinoic Acid Receptor |
| **RARE** | Retinoic Acid Response Element |
| **REST** | RE1-silencing transcription factor |
| **RNA** | Ribonucleic acid |
| **RPB** | RNA Polymerase B |
| **RT** | Reverse-transcription |
| **RXR** | Retinoic X Receptor |
| **SHH** | Sonic hedgehog |
| **SMC** | Structural Maintenance of Chromosomes) |
| **TBP** | TATA box binding protein |
| **TDD** | Trichothiodystrophy |
| **TF** | Transcription Factor |
| **tRA** | trans Retinoic Acid |
| **TRα** | Thyroid hormone Receptor α |
| **TRAP** | Thyroid hormone Receptor-Associated Protein |
| **UV** | Ultra-violet |
| **VDR** | Vitamin D Receptor |
| **VWA** | Von Willebrand factor type A |
| **XLID** | X-linked Intellectual Disability |
| **XP** | Xeroderma Pigmentosum |

# RÉSUMÉ DE LA THÈSE EN FRANÇAIS

**INTRODUCTION**

La transcription est une des étapes clés de l'expression des gènes d'une cellule. Elle nécessite le recrutement au niveau du promoteur d'un gène activé de l'ARN polymérase II (Pol II), des facteurs généraux de transcription (TFIIA, -B, -D, -E, -F, -H), du Médiateur, de co-activateurs, de protéines de remodelage de la chromatine et de facteurs de réparation de l'ADN (Compe and Egly, 2012). Une meilleure compréhension des mécanismes de régulation de l'expression des gènes passe donc par l'étude des différents complexes participant à la transcription. Notre intérêt s'est porté sur le Médiateur, complexe qui interagit avec le facteur général de transcription et de réparation TFIIH, facteur étudié en détail dans notre laboratoire.

Le Médiateur (MED) est un complexe multi-protéique conservé au cours de l'évolution et constitué de 25 sous-unités chez la levure et de 30 ou plus chez les organismes supérieurs. Les sous-unités sont organisées en trois modules principaux (la Tête, le Milieu et la Queue) et un module dissociable, le module Kinase (Figure 1). La principale fonction du Médiateur est de transmettre à la machinerie basale de transcription les différents signaux fournis par les facteurs fixés sur des séquences d'ADN spécifique (Poss et al., 2013). Cependant, le Médiateur ne régule pas seulement l'initiation de la transcription mais aussi l'élongation, la terminaison, ainsi que le remodelage de la chromatine (Zhu et al., 2011; Conaway and Conaway, 2013; Mukundan and Ansari, 2013; Whyte et al., 2013).

Ces dernières années, de nombreux travaux ont montré que des mutations dans certaines sous-unités du MED ou dans ses partenaires sont à l'origine de diverses pathologies telles que des malformations congénitales, des troubles neurodéveloppementaux ou parfois des cancers (Spaeth et al., 2011).

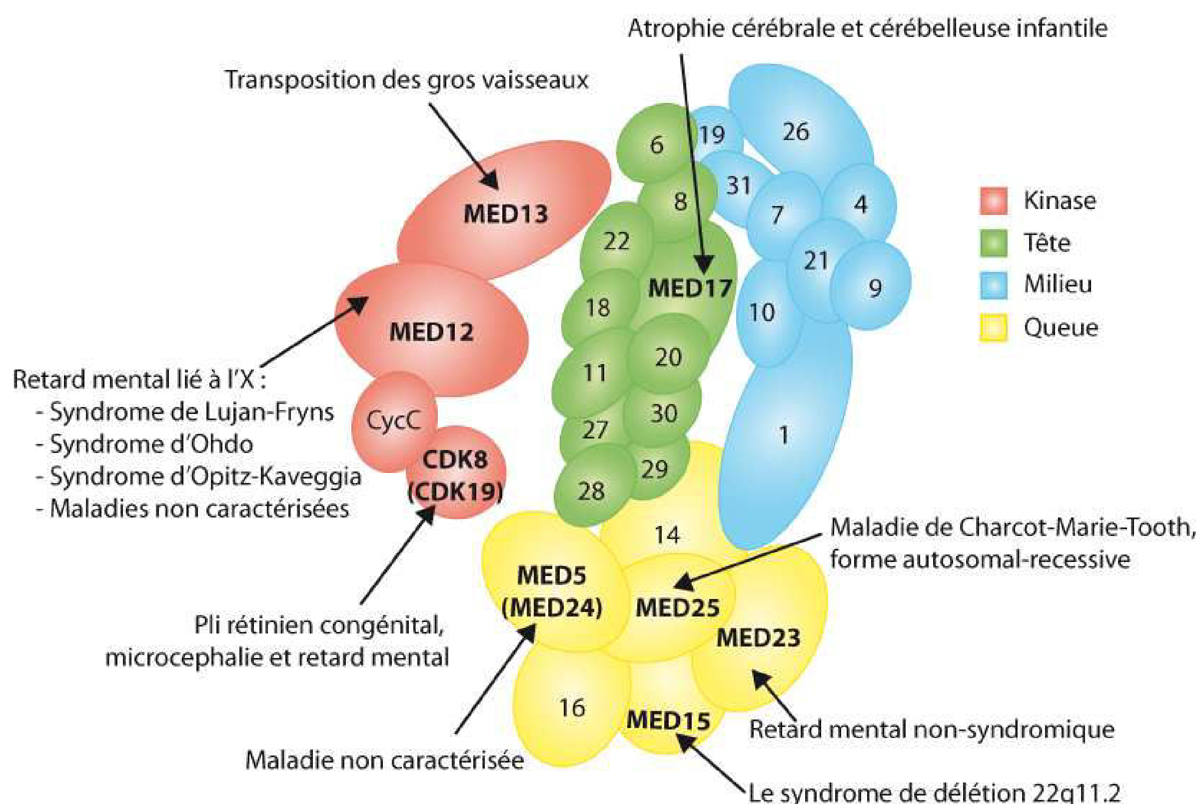



**Figure 1 : Organisation du complexe Médiateur humain et des maladies génétiques associées à une mutation dans une de ses sous-unités.**
*Cette figure sert d'illustration et est basée sur les différentes données d'interactions et de structures connues (Tsai et al., 2014). Dans ce plan 2D, une partie de la localisation, des interactions et de la taille des sous-unités n'a pas pu être respectée. Les maladies génétiques associées à une mutation dans certaines sous-unités du Médiateur sont également indiquées.*

Afin de mieux comprendre l'étiologie des ces maladies, nous nous sommes intéressés aux sous-unités MED12 et MED17 du Médiateur. Des mutations dans le gène *MED12,* localisé sur le chromosome X, sont à l'origine de divers syndromes caractérisés par une déficience intellectuelle : le syndrome d'Opitz-Kaveggia (p.G958E, p.R961W) (Risheg et al., 2007; Rump et al., 2011), le syndrome de Lujan-Fryns (p.N1007S) (Schwartz et al., 2007), le syndrome d'Ohdo (p.R1148H, p.S1165P, p.H1729N) (Vulto-van Silfhout et al., 2013), une profonde déficience intellectuelle non syndromique (p.S1967Qfsx84) (Lesca et al., 2013) et des pathologies en cours d'identification clinique (p.R206Q, p.N898D et p.R1295H) (Figure 2).De même, une mutation homozygote faux-sens dans MED17 (p.L371P) a été découverte chez des patients présentant une atrophie cérébrale et cérébelleuse avec un sévère défaut de myélinisation (Kaufmann et al., 2010). Nous nous sommes également intéressés à un des partenaires du Médiateur, le facteur NIPBL, du complexe cohésine. Ce facteur contribue à la régulation de l'expression des gènes en facilitant, entre autres, la formation de boucles d'ADN entre l'enhancers et le promoteur d'un gène (Kagey et al., 2010; Muto et al., 2014). Par ailleurs, le gène *NIPBL* est

retrouvé muté chez environ 50% des patients avec un syndrome Cornelia de Lange (CdLS), maladie qui se caractérise par une malformation du visage, associée à un retard de croissance et à une déficience intellectuelle (Liu and Krantz, 2009). Bien que certains symptômes soient communs, chaque syndrome a ses propres caractéristiques cliniques.

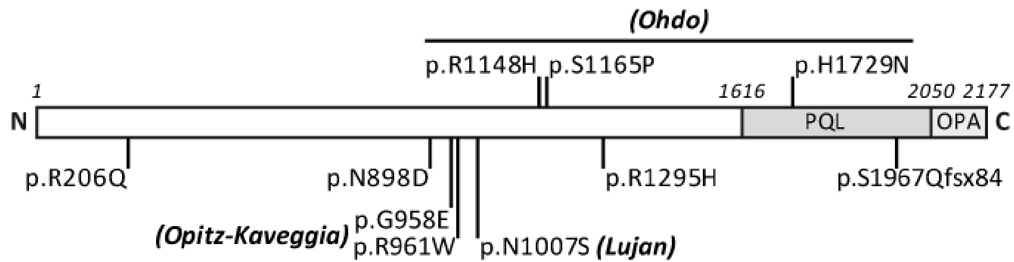



**Figure 2 : Représentation schématique de la protéine MED12 et des mutations associées.**
*Les différents domaines de MED12 sont indiqués : le domaine PQL (riche en Proline, Glutamine et Leucine) impliqué dans les interactions protéiques et le domaine OPA (riche en Glutamine). Les modifications d'acides aminés conduisant à diverses maladies avec une déficience intellectuelle sont présentées.*

Ainsi, outre une meilleure compréhension de l'étiologie de ces maladies, le but de ma thèse est de mieux appréhender le rôle des diverses sous-unités du MED et de ses partenaires lors de la transcription des gènes.

## RESULTATS

### *Maladies associées à une mutation dans MED12 ou MED17*

Nous avons pu obtenir des cellules lymphoblastoïdes provenant de patients mâles portant chacun une mutation différente (R206Q, N898D, R961W, N1007S, R1148H, S1165P ou R1295H) dans le gène *MED12*, ainsi que d'un patient portant la mutation homozygote dans le gène *MED17* (L371P). Nous avons également reçu des fibroblastes (cellules plus aisées à étudier) provenant d'un patient mâle et de sa mère portant la mutation MED12/R1295H, respectivement à l'état hémi- et hétérozygote. Lors de la vérification de la présence de la mutation, le séquençage par la méthode de Sanger a révélé que 80% des cellules de la mère expriment l'allèle *MED12* muté.

Dans un premier temps, j'ai évalué l'influence des mutations dans MED12 ou MED17 (i) sur le taux d'expression d'ARN messager (ARNm) des sous-unités du Médiateur par reverse-

transcription quantitative PCR (RT-qPCR) des ARN extraits des cellules de patients ; (ii) sur la stabilité de la protéine mutée par western blot (WB) en utilisant des extraits cellulaires de patients et (iii) sur l'architecture du complexe par co-immunoprécipitation en utilisant un anticorps dirigé contre une des sous-unités du complexe et révélée par WB. La mutation n'affecte ni le taux d'expression d'ARNm, ni la synthèse de la protéine mutée, ni celle des autres sous-unités du Médiateur. De même, elle ne modifie pas la composition du complexe Médiateur.

Dans un second temps, j'ai étudié l'effet des mutations sur l'activation de la transcription de certains gènes par RT-qPCR. Certaines caractéristiques cliniques des patients, telles que les malformations congénitales, peuvent être causées par une dérégulation hormonale de différentes voies de développement. Par conséquence, je me suis d'abord intéressée à l'expression des gènes connus pour être sous le contrôle de l'acide rétinoïque (tRA) (gènes *RAR*$\beta$, *PDK4* et *TMG2*) et de la vitamine D (vitD) (gènes *CYP24* et *OSTÉOPONTINE*). J'ai également analysé l'expression des gènes à réponse précoce (IEGs), tel que *JUN*, *FOS* et *EGR1*, dont l'expression peut être induite par irradiation des cellules aux UV ou par l'ajout de sérum dans des cultures cellulaires préalablement privées en sérum. Notre laboratoire a montré une dérégulation des IEGs, induits en réponse au sérum, dans les cellules d'un patient atteint d'une déficience intellectuelle non syndromique, maladie associée à une mutation dans MED23 (R617Q) (Hashimoto et al., 2011). Or, le symptôme commun à tous les patients portant une mutation dans MED12 ou MED17 est une déficience intellectuelle, ce qui oriente mes travaux dans cette même direction.

Après traitement des cellules à tRA, à la VitD, au sérum ou UV, j'ai observé une altération du niveau d'expression des gènes étudiés, altération différente en fonction de la localisation de la mutation et de la nature de l'activation (UV ou sérum).

Finalement, l'effet de la mutation sur le recrutement des facteurs impliqués dans la formation du complexe de transcription au niveau du promoteur des gènes dérégulés a été analysé par des expériences d'immunoprécipitation de la chromatine (ChIP) avec des anticorps dirigés contre ces facteurs. La modification du niveau d'expression des gènes observée dans les cellules de patients est la conséquence d'un défaut de recrutement de certains composants de la machinerie de transcription, ainsi que d'une perturbation des différentes modifications post-

traductionnelles des histones nécessaires à l'activation de la transcription (Li et al., 2007; Bannister and Kouzarides, 2011).

Ces résultats suggèrent que chaque mutation dans le gène *MED12* ou *MED17* provoque un dysfonctionnement spécifique de certains processus de régulation de l'expression des gènes, conduisant donc au développement de différentes pathologies.

***Syndrome Cornelia de Lange associé à une mutation dans NIPBL :***

J'ai également travaillé sur des fibroblastes d'un patient avec le syndrome Cornelia de Lange (CdLS), hétérozygote pour la mutation c.6516-6517insA dans le gène *NIPBL.* Cette mutation a pour conséquence l'expression d'une protéine tronquée. J'ai observé une réduction d'environ 30-40% du niveau d'expression du gène *NIPBL* dans les cellules mutées comparé aux cellules contrôles, alors que l'expression de ses interactants (sous-unités du complexe Médiateur et cohésines) est inchangée. Par contre, le niveau d'expression de la protéine NIPBL n'a pas pu être analysé, et ce dû à l'absence d'anticorps contre cette protéine. Les résultats préliminaires de ChIP ont révélés un plus faible recrutement de NIPBL, ainsi que de l'ARN polymérase II, au niveau du promoteur du gène *NIPBL* dans les cellules mutées comparées aux cellules contrôles. Ces résultats suggèrent une autorégulation de NIPBL.

La mutation n'a pas d'effet sur la transcription des gènes *CYP24* et *OSTÉOPONTINE* induits par la VitD ou des gènes *JUN*, *FOS* et *EGR1* induits par l'ajout de sérum. Par contre, après traitement des cellules à tRA, j'ai observé 50 fois plus d'induction du gène *RAR$\beta$* dans les cellules mutées par rapport aux cellules contrôles. Néanmoins, avant traitement, ce gène est beaucoup plus faiblement exprimé dans les cellules mutées que dans les cellules contrôles. Les résultats préliminaires de ChIP montrent un défaut de recrutement du complexe de transcription sur le promoteur du gène *RAR$\beta$* dans les conditions physiologiques. L'analyse de l'expression d'autres gènes cibles de tRA (*PDK4*, *NRIP1*, *TMG2, SMAD3* et *RAR$\alpha$*) ne montre aucune dérégulation aussi importante que celle de *RAR$\beta$* dans les cellules CdLS comparé aux cellules contrôles.

Ces résultats démontrent un rôle de NIPBL dans la régulation transcriptionnelle de certains gènes, et plus particulièrement dans l'expression de *RAR$\beta$*, basale ou induite.

***Modèles murins des maladies associées à une mutation dans le Médiateur :***

Afin d'acquérir une meilleure compréhension des mécanismes physiopathologiques impliqués dans les maladies associées à une mutation dans une des sous-unités du Médiateur, nous avons essayé de générer, en collaboration avec l'Institut Clinique de la Souris et dans le cadre du projet GENCODYS, des souris, portant soit la mutation MED12/p.R961W, MED17/p.L371P ou MED23/p.R617Q à l'état homozygote. Ces mutations provoquent respectivement chez l'homme le syndrome d'Opitz-Kaveggia (Risheg et al., 2007), une atrophie cérébrale et cérébelleuse infantile (Kaufmann et al., 2010) et une déficience intellectuelle non syndromique (Hashimoto et al., 2011).

Pour Med12, nous n'avons pas pu obtenir des souris chimères, c'est-à-dire des souris provenant d'embryons auxquels on a injecté au stade blastocytes des cellules souches embryonnaires portant la mutation. Nous avons donc essayé de générer des souris arborant un knock-out conditionnel (cKO) pour le gène Med12 et ce fût encore une fois un échec.

Les souris homozygotes Med23/R617Q meurent au cours du développement embryonnaires, tandis que les souris homozygotes Med17/L371P ne survivent que jusqu'à 7-8 semaines après la naissance. Un suivi du développement de ces dernières (taille, masse et température corporelle, réflexe de redressement …etc.) est actuellement en cours, ainsi qu'une analyse par microarray de l'expression des gènes dans différentes régions du cerveau. Nous avons également isolé des fibroblastes embryonnaires de souris (MEF) à partir des embryons homozygotes pour la mutation MED17/p.L371P. Les résultats préliminaires montrent que la mutation perturbe la composition du complexe Médiateur et l'expression de certains gènes.

## PERSPECTIVES

Mon travail permet ainsi de définir les prochains axes de recherche sur le Médiateur et ses partenaires. Des expériences de microarrays pourront être réalisées afin de mieux évaluer la spécificité des diverses mutations de MED12 et MED17 sur l'expression des gènes en réponse à certaines activations. De plus, les expériences sur les modèles murins seront poursuivies afin de mieux comprendre l'effet des mutations sur le développement d'un organisme.

Notre laboratoire a montré que la formation d'une boucle d'ADN entre le promoteur et le terminateur du gène *RAR$\beta$* est nécessaire pour la synthèse optimale de son ARNm (Le May et al., 2012). De plus, Kagey *et al.* ont montré que le complexe cohésine, composé de NIPBL, interagit avec le Médiateur afin de faciliter la formation de ces boucles d'ADN (Kagey et al.,

2010). Des expériences de q3C (quantitative Chromatin Conformation Capture) devront être entreprises afin d'analyser l'organisation spatiale du gène *RAR*$\beta$ dans les cellules du patient atteint du syndrome Cornelia de Lange.

Nous pourrons ainsi mieux percevoir comment chaque mutation dans MED12, MED17 ou NIPBL affectent l'expression de certains gènes et ainsi être à même d'expliquer certains phénotypes.

## CONCLUSION

Mes travaux de thèse ont permis de mieux comprendre certains des mécanismes moléculaires des maladies associées à une mutation dans les sous-unités MED12 ou MED17 du Médiateur ou dans son partenaire, NIPBL. J'ai pu montrer que chaque mutation dans MED12 ou MED17 conduit à une dérégulation spécifique de l'expression des gènes, expliquant ainsi qu'en fonction de la position de la mutation dans le gène MED12, la pathologie associée est différente. Dans un même temps, cette étude a également permis de mieux appréhender l'action des sous-unités MED12 et MED17 du complexe Médiateur et de son partenaire NIPBL lors de la transcription.

# INTRODUCTION

# INTRODUCTION

There is nearly half a century that DNA has been described as the molecule that encodes the genetic instruction of all living organisms. Cells use two processes in series to convert the coded information inside the DNA into proteins. In the first, called transcription, the coding region of a gene is copied into RNA molecule. The second process called translation allows the production of proteins from RNA molecules. Proteins are essential for the development and the functioning of cells life.

## I. Transcription of class II genes

The process of DNA transcription is carried out almost exclusively by multisubunit DNA-dependent RNA polymerase (Pol). To date, four different RNA polymerases (Pol I or A, Pol II or B, Pol III or C and Pol IV) have been identified in higher eukaryotes, whereas only one is found in prokaryotes and *archaea* (Roeder and Rutter, 1969). Pol I are responsible for the synthesis of major ribosomal RNA (rRNA, excepted 5S rRNA), whereas Pol III are implicated in the transcription of small RNAs (including 5S rRNA and transfer RNA). Pol IV, identified only in plants, is involved in RNA-directed DNA methylation, transcriptional silencing and formation of heterochromatin (Herr et al., 2005; Kanno et al., 2005; Onodera et al., 2005). Finally, Pol II catalyzes the formation of messenger RNA (mRNA) allowing proteins synthesis. In this manuscript, I will detail RNA polymerase II machinery, which is the most intricate, consisting of more than 60 polypeptides.

### A. RNA polymerase II

RNA polymerase II is the core of the transcription machinery. On its own, it can polymerize RNA and proofread the nascent transcript.

#### A.1. Subunit composition

RNA polymerase II contains 12 subunits, designated RPB1 to RPB12 (RNA Polymerase B 1-12), which represented a total mass of >0,5MD (Young, 1991). In general, the 12 subunits of Pol II are highly conserved in sequence, function and architecture between human, bacteria

and yeast. Indeed, RPB1, RPB2, RPB3, RPB6 and RPB11 are related respectively to β', β, α, ω and α subunits of bacterial RNA polymerase (Tan et al., 2000; Minakhin et al., 2001) (Table 1). Moreover, 5 subunits of Pol II (RPB5, RPB6, RPB8, RPB10 and RPB12) are commonly shared between Pol I, Pol II and Pol III, whereas 4 subunits (RPB1, RPB2, RPB3 and RPB11) have sequence-homology counterparts in Pol I and Pol III (Woychik and Young, 1990; Hampsey, 1998).

### A.2. Structure of Pol II

The resolution of Pol II structure by X-ray crystallography has allowed a better understanding of its organization (Armache et al., 2003, 2005; Bushnell and Kornberg, 2003). We can distinguish two main parts: the core formed by 10 subunits and the stalk composed by the heterodimer RPB4 and RPB7 (Table 1). This heterodimer is important for transcription initiation but not for elongation. The two largest subunits RPB1 and 2 created a positively charged cleft in which the DNA enters in order to be transcribed (Cramer, 2000, 2001; Gnatt et al., 2001). The other subunits (RPB3, 6, 10, 11 and 12) will be assembled around to maintain the structure.

| | Pol II subunits | Size (kDa) | Orthologs in *E. Coli* | Notes |
|---|---|---|---|---|
| Core | **RPB1** | 220 | β' | Contains CTD<br>involved in the selection of the initiation and elongation site |
| | **RPB2** | 140 | β | contains the active site<br>involved in the selection of the initiation and elongation site |
| | **RPB3** | 33 | α | |
| | **RPB5** | 25 | | Common to all Pol |
| | **RPB6** | 14.5 | ω | Common to all Pol |
| | **RPB8** | 17.1 | | Common to all Pol |
| | **RPB9** | 14.4 | | involved in the selection of the initiation site |
| | **RPB10** | 7.6 | | Common to all Pol |
| | **RPB11** | 13.2 | α | |
| | **RPB12** | 7 | | Common to all Pol |
| Stalk | **RPB4** | 16.2 | | Involved in the initiation, form a subcomplex with RPB7 |
| | **RPB7** | 19.2 | | Involved in the initiation, form a subcomplex with RPB4 |

**Table 1: Composition of RNA polymerase II and function of the different subunits.**
*Presentation of the 12 Pol II subunits and their functions*

### A.3. CTD and its modifications

The RPB1 subunit of Pol II contains a unique structural element, not found anywhere else: the C-terminal domain (CTD), which is very important in the regulation of Pol II. The CTD consist of a tandem repeat of heptapeptide: Tyr-Ser-Pro-Thr-Ser-Pro-Ser ($Y^1S^2P^3T^4S^5P^6S^7$). The number of repeats is depending upon the species: 52 times in humans, 42 times in *Drosophila* and 26 to 29 times in yeast. CTD is an essential element for life. Indeed, cells containing a Pol II with partial truncation of the repeat structure are not viable (Nonet and Young, 1989). However, the CTD is not necessary for *in vitro* transcription (Zehring et al., 1988).

CTD is a substrate for different post-translational modifications that contribute to the regulation of Pol II activity. Indeed, CTD is subject to a cycle of phosphorylation and dephosphorylation, throughout the transcriptional process (Dahmus, 1995). By consequence, depending on the phosphorylation state, two form of human Pol II (IIO and IIA) can be distinguished. The IIA form has a hypo- or unphosphorylated CTD and is normally involved in the assembly of the preinitiation complex. The IIO form is highly phosphorylated, mainly at serine residues 2, 5 and 7 and is implicated in transcription steps occurring after preinitiation complex assembly.

## B. Regulatory sequence in the heart of DNA

The promoters of Pol II transcribed genes have a great variability (Gershenzon and Ioshikhes, 2005). Eukaryotic promoters can be divided into core and regulatory elements. Core promoter elements define the site for assembly and orientation of the main transcription proteins along with Pol II itself and encompass the transcription start site (TSS). In contrast, regulatory elements are gene-specific sequences that are located usually upstream of the core promoter and control the rate of transcription initiation. Both elements are reviewed below.

### B.1. The core promoter

The core promoter spans about 40 base pairs (pb) up- and downstream of the TSS. Studies on eukaryotic promoters have thus far identified eight core promoter elements (Figure 3). However, there are no universal core promoter elements.

### B.1.a. The TATA-box

The TATA-box was the first core promoter motif that was discovered. The metazoan TATA box contains the consensus sequence TATA(A/T)A(A/T)(A/G), located approximately 25 to 30 nucleotides upstream of the TSS in human. It is recognized by the TBP subunit of the TFIID complex which is responsible for the positioning of the other transcription factors on the promoter. Interestingly, although TATA box is a well known core promoter motif, it is present in only 10-20 % of the mammalian core promoters (Kim et al., 2005; Carninci et al., 2006; Cooper et al., 2006).

### B.1.b. The initiator element (Inr)

The initiator (Inr) contains a pyrimidine-rich sequence (C/T)(C/T)**A**N(T/A)(C/T)(C/T), surrounding the transcription start site (Corden et al., 1980; Javahery et al., 1994). The first A nucleotide usually becomes the first transcribed nucleotide. Inr is sufficient for (i) determining the start site location in a promoter that lacks a TATA box and (ii) enhancing the strength of a promoter that contains a TATA box. TAF1/TAF2 components of TFIID have been implicated in Inr recognition (Chalkley and Verrijzer, 1999).

### B.1.c. The TFIIB recognition element (BREu and BREd)

The TFIIB recognition element (BRE) is a disjoint binding element of TFIIB transcription factor. It consists of two sequences: BREu and BREd, located respectively upstream and downstream of the TATA box (Lagrange et al., 1998; Deng and Roberts, 2005). Both the BREu and BREd function in conjunction with a TATA box and can have positive or negative effects on transcription depending of the promoter(Evans, 2001; Deng et al., 2009).

### B.1.d. The Downstream promoter element (DPE)

The DPE is a core promoter element located downstream (+28 to +33) of the Inr (Burke and Kadonaga, 1996). The DPE consensus sequence in *Drosophila* is (A/G)G(A/T)CGTG while the DPE consensus in humans has yet to be determined. It is recognized by TAF6 and TAF9 subunit of TFIID complex and functions cooperatively with the Inr (Burke and Kadonaga, 1997). The spacing between the Inr and DPE is important for the optimal transcription (Kutach and Kadonaga, 2000).

### B.1.e. The Motif ten element (MTE)

The MTE is another core promoter element situated downstream of the TSS. It contains the consensus sequence C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C) and is located immediately upstream of the DPE (+18 to +27) (Lim et al., 2004). Both MTE and DPE may work in synergy with each other and with Inr.

### B.1.f. The downstream core element (DCE)

In contrast to the sequence continuity seen with other core promoter elements, DCE contains three discontinuous subelements: SI, SII and SIII. TAF1 component of TFIID is implicated in their recognition. The presence of DCE and DPE seems to be mutually exclusive (Lee et al., 2005).

### B.1.g. The X core promoter element 1 (XCPE1)

The XCPE1 is a rare promoter element that encompasses the TSS (Tokusumi et al., 2007). It has a consensus sequence of (G/A/T)(G/C)G(T/C)GG(G/A)A(G/C)(A/C) and is present in about 1% of human core promoters, most of which are TATA-less. The factor that recognises this element has not been identified. XCPE1 do not act by itself. Instead, it acts in conjunction with some sequence-specific activators, such as NRF1, NF-1, and Sp1.

## B.2. Regulatory response element

Although all cells of an organism have the same genetic information, the expression level of a given gene is not similar in each differentiated tissues. By consequence, genes possess regulatory sequences which allow the control of transcriptional level. There are two types of regulatory elements, classified according to their distance from the TSS: the proximal sequences and the distal sequences (Figure 3).

The proximal sequences are located between 40 and 200 pb upstream of the TSS. It is recognized by sequence-specific binding proteins that activate or inhibit transcription.

Distal regulatory elements are located several thousand base pairs upstream or downstream of the TSS. They are named “enhancer” when they activate transcription or “silencer” when they repress it. These elements will interact with the promoter thanks to the formation of a DNA loop (Cook, 2003; Saiz et al., 2005). Enhancers could act either in *cis*

(promoter and enhancer on the same DNA molecule) or in *trans* (promoter and enhancer on separate chromosome) (Goldsborough and Kornberg, 1996).

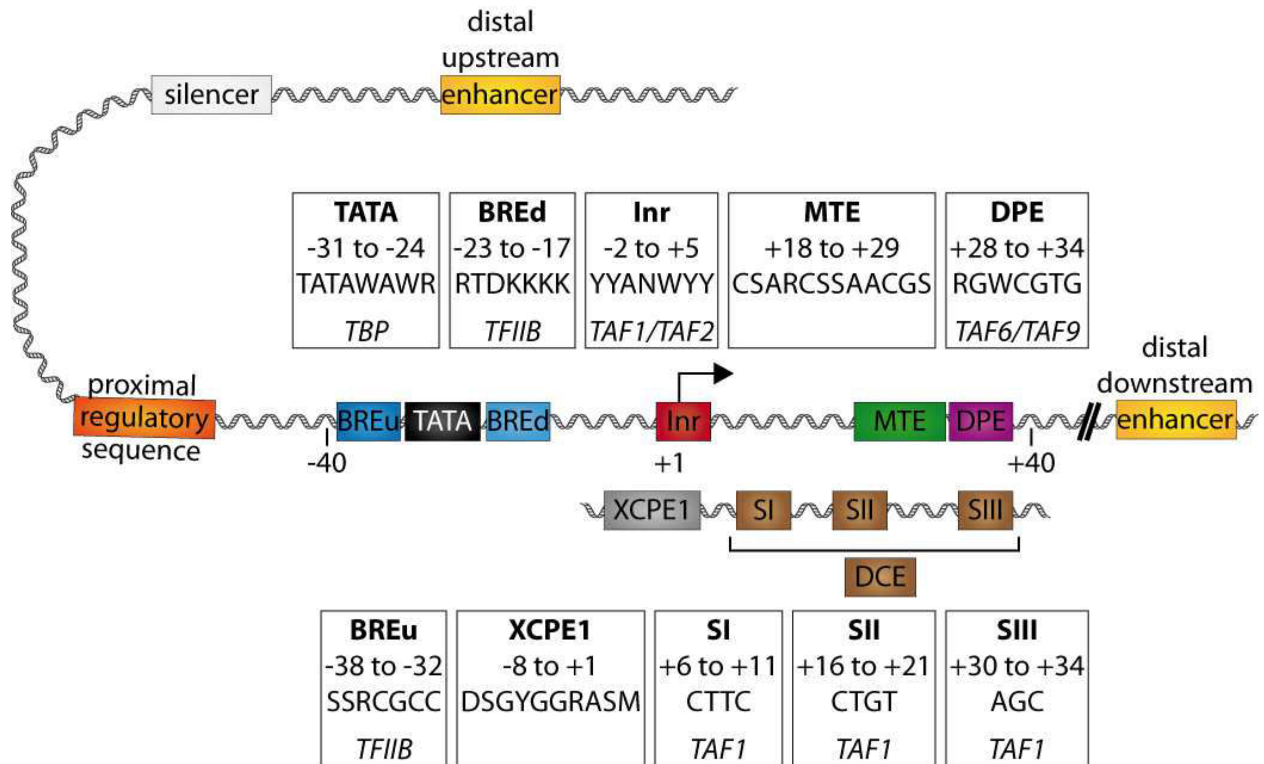



**Figure 3: Eukaryotic promoter motifs.**
*This diagram shows some promoter elements, their respective consensus sequences, their position from the transcription start site and the factors implicated in their recognition. The following elements can be found in a core promoter: the TATA box (TATA), the initiator (Inr), the TFIIB recognition element located either upstream (BREu) or downstream (BREd) of the TATA box, the motif ten element (MTE), the downstream promoter element (DPE) and the downstream core element (DCE). It is likely that additional core promoter motifs remain to be discovered. This diagram is roughly to scale. (S=C/G; W = A/T; Y=C/T; R=A/G; M=C/A; K=T/G; D=T/G/A; N=A/C/G/T)*
*Figure adapted from Juven-Gershon and Kadonaga, 2010*

## C. General transcription factors

Purified RNA polymerase II can synthesize RNA from a DNA template but is not able to recognize the core promoter (Weil et al., 1979). This process requires additional factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH) that are called general transcription factors (GTFs) (Table 2). They were named using the following nomenclature: TF represents Transcription Factor, the Roman numeral II indicates Pol II-driven transcription, and the “letter” generally corresponds to which chromatographic fraction the specific GTF was isolated from (Matsui et al., 1980; Samuels et al., 1982).

Mediator (MED) complex is not considered as GTFs but essential for transcription. It will be extensively presented in Chapter II, since Mediator is a major subject of my thesis.

| Factor | Protein composition | Size (kDa) | Function in Pol II transcription | References |
|---|---|---|---|---|
| TFIIA | α | 35 | Antirepression | (Merino et al., 1993; Ge and Roeder, 1994) |
| | β | 19 | Stabilizes TBP-TATA complex | (Buratowski et al., 1989; Lee et al., 1992) |
| | γ | 12 | Coactivation | (Shykind et al., 1995) |
| TFIIB | | 33 | Selects TSS | (Li et al., 1994) |
| | | | Stabilizes TBP-TATA complex | (Maldonado et al., 1990) |
| | | | Recruits Pol II/TFIIF | (Ha et al., 1993) |
| TFIID | TBP | 38 | Binds Core promoter | (Sawadogo and Roeder, 1985) |
| | | | Coactivation | (Lavigne et al., 1999) |
| | TAF1-14 | 15-250 | Protein kinase activity | (Dikstein et al., 1996) |
| | | | Ubiquitin-activating/conjugating activity | (Pham and Sauer, 2000) |
| | | | Histone acetyltransferase activity | (Mizzen et al., 1996) |
| TFIIE | α | 56 | Recruits and regulates TFIIH | (Ohkuma and Roeder, 1994; Okuda et al., 2004) |
| | | | Facilitates formation of an initiation-competent Pol II | (Holstege et al., 1996) |
| | β | 35 | Involved in promoter clearance | (Maxon et al., 1994; Watanabe et al., 2003) |
| TFIIF | 2 x RAP30 | 30 | Recruits Pol II to the promoter | (Flores et al., 1991) |
| | | | Recruits TFIIE and TFIIH | (Buratowski et al., 1989) |
| | 2 x RAP74 | 74 | Selects TSS | (Killeen et al., 1992; Ghazy et al., 2004) |
| | | | Facilitates Pol II promoter escape | (Yan et al., 1999) |
| | | | Increase elongation efficiency | (Zhang and Burton, 2004) |
| TFIIH | Core: XPB | 89 | ATPase activity for transcription initiation and promoter clearance | (Goodrich and Tjian, 1994; Dvir et al., 1997; Kumar et al., 1998) |
| | Core: p62 | 62 | Helicase activity for promoter opening | (Holstege et al., 1996; Tirode et al., 1999) |
| | Core: p52 | 52 | Kinase activity for phosphorylation of Pol II CTD and nuclear receptors | (Lu et al., 1992; Rochette-Egly et al., 1997; Drané et al., 2004) |
| | Core: p44 | 44 | Nucleotide Excision repair | (Schaeffer et al., 1993; Compe and Egly, 2012) |
| | Core: p34 | 34 | | |
| | Core: p8/TTDA | 8 | | |
| | XPD | 80 | | |
| | CAK: CDK7 | 39 | | |
| | CAK: Cyclin H | 37 | | |
| | CAK: MAT1 | 35 | | |
| MED | MED1-31, CDK8, CycC | 15-250 | See chapter II | (Poss et al., 2013) |

**Table 2: General transcription factors.**
*Presentation of the GTFs, their composition and their functions.*
*Table adapted from Thomas and Chiang, 2006.*

## D. The transcription cycle

Eukaryotic transcription is a precisely timed event, which can be divided into a number of distinct steps: promoter binding and pre-initiation complex assembly, initiation and promoter clearance, elongation and finally termination (Shandilya and Roberts, 2012).

### D.1. Promoter binding

The first step of RNA polymerase II (Pol II) transcription is the binding around the transcription initiation site of the transcription pre-initiation complex (PIC) that includes the general transcription factors and RNA polymerase II. PIC formation may occur by two different pathways: the sequential assembly pathway or Pol II holoenzyme pathway (Figure 4).

The sequential assembly start from the binding of TFIID to the promoter. It is followed by the entry of TFIIA and TFIIB that stabilise promoter-bound TFIID. Then, Pol II, together with TFIIF is recruited. This drives the association of TFIIE and the subsequent entry of TFIIH (Buratowski et al., 1989; He et al., 2013).

In Pol II holoenzyme pathway, a preassembled complex containing Pol II and Mediator complex with GTFs (excepted TFIID and TFIIA) is recruited in one step to the promoter. Holoenzyme complexes of different compositions have been reported (Kim et al., 1994; Ossipow et al., 1995; Maldonado et al., 1996). Although both pathways were identified *in vitro*, it is likely that both exist *in vivo*.

### D.2. Initiation and promoter clearance

To progress further, DNA strands must be separated around the TSS. This promoter opening occurs due to the ATP-dependent helicase activity of XPB, a TFIIH subunit (Holstege et al., 1996; Coin et al., 1999). The template strand is then placed in the active site of Pol II (Figure 5B). Usually, the transcription starts from several abortive runs and transcripts of less than 10 nucleotides are released (Goldman et al., 2009). However, once RNA products are longer than 10 nucleotides, Pol II can clear the promoter. This step requires the phosphorylation of serine 5 of Pol II CTD by CDK7 subunit of TFIIH (Ohkuma and Roeder, 1994). The phosphorylated Ser5 CTD repeat is further recognized by the capping enzyme, which then catalyze the addition of a methylguanosine cap to the 5′ end of nascent mRNA. It is a signal for productive transcription initiation (Komarnitsky, 2000).

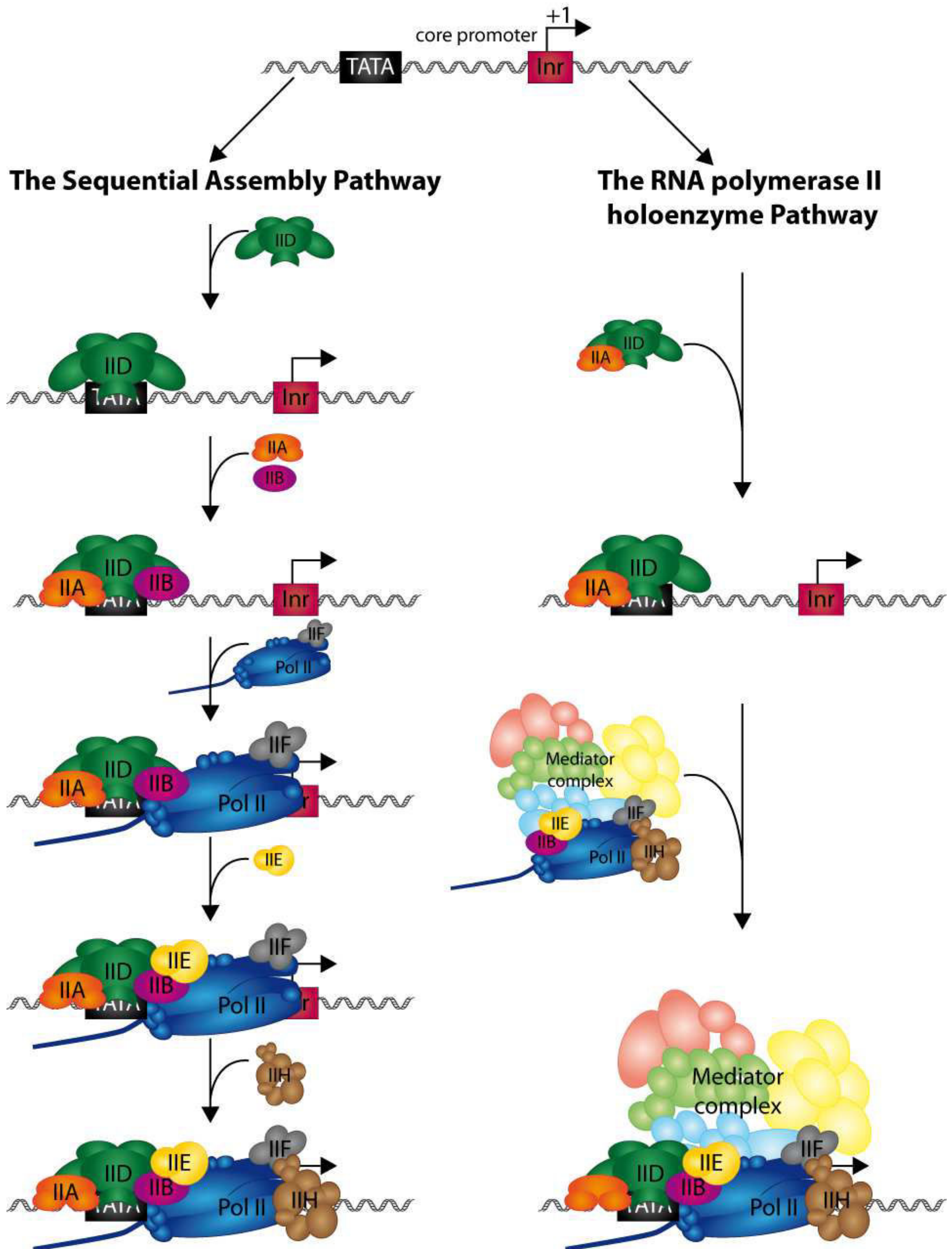



**Figure 4: Pre-initiation complex assembly.**

*PIC formation may occur by stepwise recruitment of the general transcription machinery (left panel, the sequential assembly pathway) or by recruitment of a preassembled Pol II holoenzyme and TFIID complexes (right panel, the Pol II holoenzyme pathway).*

Following promoter clearance, Pol II proceeds for elongating the transcript while a part of the PIC components remains associated at the promoter, forming a re-initiation scaffold complex. This complex consists at least of TFIIA, TFIID, TFIIE and TFIIH and the Mediator. TFIIF is the only factor staying with Pol II (Bengal et al., 1991; Lei et al., 1999).

### D.3. Proximal pausing and elongation

Successful initiation does not guarantee productive elongation. Immediately following initiation, Pol II soon enters transcriptional arrest and if nothing else happens, terminates the transcription. This arrest is mediated by DSIF (DRB Sensitivity Inducing Factor) and NELF (Negative elongation factor), which bind Pol II and inhibit its function. Their negative effect can be relieved by the positive transcription elongation factor, P-TEFb (Yamaguchi et al., 1999). The kinase subunit of P-TEFb, CDK9, phosphorylates DSIF and NELF, as well as the serine 2 of Pol II CTD (Marshall et al., 1996) (Figure 5D). This event allows the transition of Pol II from promoter proximal pausing to productive elongation for efficient full-length mRNA synthesis (Cheng and Price, 2007). Recent studies suggest that stable pausing of polymerase provide a temporal window of opportunity for recruitment of factors to modulate gene expression (Henriques et al., 2013; Buckley et al., 2014).

As the Pol II progresses towards the 3' of the gene, there is an increase in the phosphorylation status of Ser2 and gradual loss of Ser5 phosphorylation of the CTD repeats (Komarnitsky, 2000). Besides P-TEFb, Pol II is further assisted by elongation factors like TFIIS, the ELL phosphatase, Elongin, histone chaperone complex FACT and histone deacetylases (HDACs). The two last factors are required for Pol II progression through nucleosomes. In addition, the elongation complexes serve as a platform for downstream RNA processing (Zhou et al., 2012).

### D.4. Termination

Termination is the last step of transcription cycle, which allows the release of the transcript and Pol II dissociation. This step is one of the least understood processes in gene expression. There are two known mechanisms that are linked to processing of the transcript: poly(A)-dependent pathway and Sen1-dependent pathway (Birse et al., 1998). The choice of the pathways depends on the RNA 3'-end processing signals and the termination factors that are present at the end of a gene.

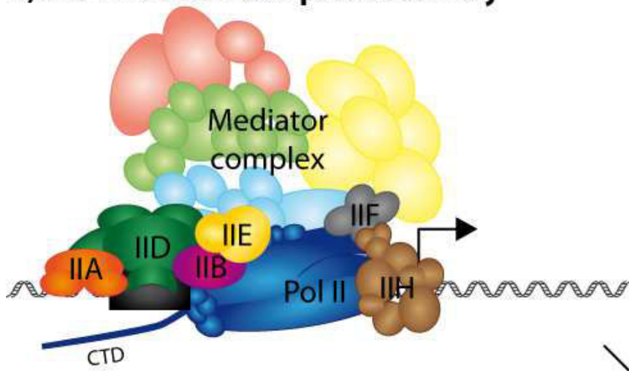
Transcripts of most protein-coding genes are polyadenylated. The gene itself has a highly conserved poly (A) signal, 5′-AAUAAA-3′ which is followed by a G/U-rich sequence towards the 3′ end. During transcription, the Ser2-phosphorylated CTD repeats recruit several factors required for the termination (McCracken et al., 1997; Ahn et al., 2004). The first recruited factor is the cleavage and polyadenylation specificity factor (CPSF), which bind to the poly(A) signal (Nag et al., 2007) (Figure 5E). This binding reduces the rate of Pol II progress and causes its pausing. Another factor called the cleavage stimulatory factor (CstF) binds to the downstream GU-rich signal. Interaction between CPSF and CstF leads to the RNA cleavage between poly(A) signal and GU-rich region. This cleavage is then followed by the polyadenylation of the upstream cleavage product, the degradation of the downstream cleavage product by XRN2 and release of the Pol II. Finally, RNA is spliced by the spliceosome to be translated into protein.

For genes that do not contain poly(A) signal, the termination mechanism is via the Sen1-dependent pathway. In this case, the Senatoxin protein is responsible for the unwinding of the RNA-DNA hybrid inside the active site of Pol II and thus transcription termination (Steinmetz et al., 2001, 2006).

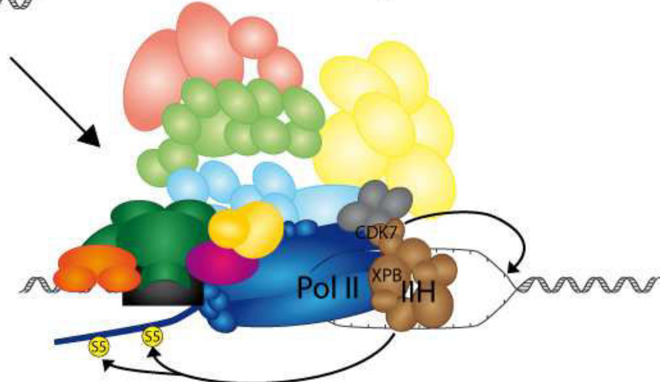
### D.5. Re-initiation and gene looping

There is a body of evidence for a physical interaction between the terminal and promoter regions of active genes. This interaction involves chromosome looping and facilitates transcription reinitiation by the same Pol II complex (O'Sullivan et al., 2004; Ansari and Hampsey, 2005). Release of Pol II at the end of transcription requires a dynamic reversal of the associated covalent marks on the CTD repeats. This will bring Pol II to its original hypo-phosphorylated state for a subsequent round of transcription. Indeed, the remaining promoter bound GTFs form a scaffold that allows reinitiation (Yudkovsky et al., 2000). Moreover, transcription activator has been shown to be involved in transcription reinitiation by facilitating promoter-terminator association (El Kaderi et al., 2009). TFIIB directs the assembly of such reinitiation by interacting with the transcription termination complexes CPSF and CstF (Singh and Hampsey, 2007). This interaction is regulated by the phosphorylation of TFIIB (Wang et al., 2010) (Figure 5F).

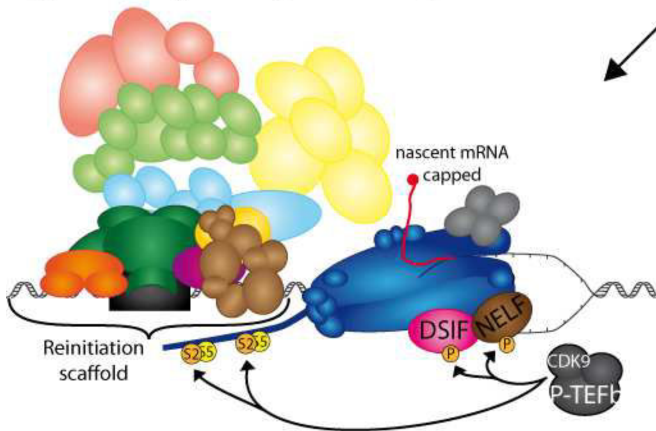
**A) Pre-initiation complex assembly**



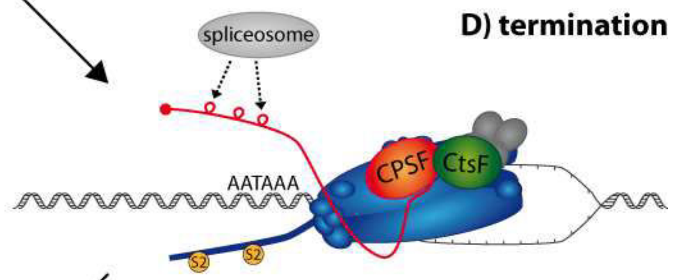

**B) Initiation and promoter clearance**




**C) proximal pausing and elongation**




**D) termination**




**E) Re-initiation**

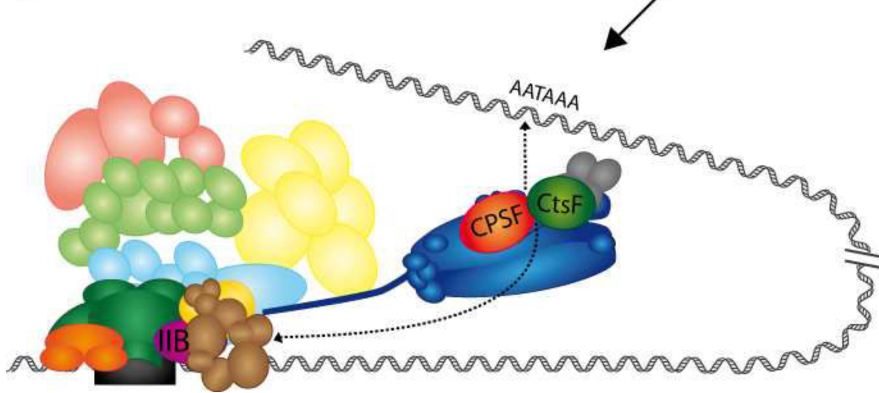


**Figure 5: RNA polymerase II transcription cycle**
**(A***)* *The stability of an assembled PIC is characterized by the presence of essential GTFs, Pol II and the mediator complex at the promoter.* ***(B)*** *The ATPase-dependent helicase activity of XPB within TFIIH allows promoter opening. The cyclin dependent kinase 7 (CDK7) subunit of TFIIH phosphorylates Ser5-CTD.* ***(C)*** *The Ser5-CTD phosphorylation recruits capping enzyme to the 5' region of nascent mRNA which triggers Pol II-escape from the promoter. Presence of negative factors such as NELF and DSIF inhibits productive transcription resulting in paused transcription. The kinase activity of the CDK9 subunit of the pTEFb complex alleviates this repression* via *phosphorylation of NELF and DSIF which results in their dissociation. Following promoter clearance, Pol II proceeds for elongating the transcript while a part of the PIC components remains associated at the promoter forming a reinitiation scaffold. The CTD repeat of the elongating Pol II is progressively phosphorylated at serine 2 by CDK9, while the Ser5-CTD phosphorylation is removed. Ser2-CTD phosphorylation recruits mRNA splicing complex for co-transcriptional splicing of nascent mRNA.* ***(D)*** *Once Pol II reaches a pause signal (poly A) at the gene terminal, 3' end processing and termination specific complexes such as CPSF and CstF are recruited.* ***(E)*** *TFIIB, in the reinitiation scaffold complex, interacts with Pol II and the termination complexes CPSF and CstF. It mediates promoter–terminator contacts known as gene looping and thereby increases the efficiency of reinitiation by Pol II.*

Transcription is a process even more complex than what I have described. Lately, multiple roles of the ubiquitin-proteasome system in transcription regulation have emerged (Geng et al., 2012). Moreover, chromatin structure (histone modifications, chromatin remodeling, histone variant incorporation…etc) is also important in gene expression (Li et al., 2007). These two subjects are too vast to be discussed in this manuscript.

## E. The activated transcription

During the development and life of an organism, several genes are required for the maintenance of basic cellular function, and thus they are expressed in all cells under normal and patho-physiological conditions. These genes are called housekeeping genes (for example: GAPDH, β-actin...). However, a large number of genes are regulated by endogenous or exogenous stimulus (cellular differentiation, stress response), for example immediate early genes and nuclear receptors regulated genes.

### E.1. Immediate early response genes

Immediate early genes (IEGs) are genes which are activated transiently and rapidly in response to a wide variety of cellular stimuli. The term IEG was originally created in reference to viral genes that were rapidly transcribed following invasion of a host (Clements et al., 1977). IEGs have important roles in processes such as brain development, learning, and responses to drug abuse (Pérez-Cadahía et al., 2011). They represent the first round of response to stimuli and thus are rapidly expressed without new protein synthesis. They encode different factors that act in a combinatorial fashion to differentially affect a distinct second wave of genes expression.

About 40 cellular IEGs have been identified so far. The earliest known and best characterized IEGs include C-FOS, EGR1, and C-JUN. The regulation of IEGs is a complex affair. Indeed, multiple signals and intracellular pathways can influence IEG-mediated transcription activation. For example, JUN can be activated by cellular stress, such as UV, heat or serum but depending on the stress, its activation pathway will be different.

### E.2. Nuclear receptors regulated genes

Nuclear receptors are one of the most abundant classes of transcriptional regulators in animals. Humans have 48 different nuclear receptors, which regulate diverse functions, such as homeostasis, reproduction, development and metabolism. Nuclear receptors function as ligand-activated transcription factors and thus provide a direct link between signalling molecules that control these processes and transcriptional responses (Olefsky, 2001).

### E.2.a. Nuclear receptors structure

Nuclear receptors share a common structural organization (Figure 6) containing five functional regions:

**(A/B) N-terminal regulatory domain (NTD).** This region is highly variable and contains at least one ligand-independent transactivation domain (AF-1) and several autonomous transactivation domains (AD), necessary for recruiting transcriptional co-activators

**(C) DNA-binding domain (DBD).** This highly conserved domain contains two zing fingers that recognize specific NR-responsive elements. It is also involved in dimerization of nuclear receptors including homodimers as well as heterodimers.

**(D) Hinge region.** Between the DNA-binding domain and ligand-binding domain, this less conserved region, behaves as a flexible hinge and contains the nuclear localization signal (NLS).

**(E) Ligand binding domain (LBD).** This domain plays a crucial role in ligand-mediated nuclear receptor. The secondary structure of 12 α-helixes is better conserved than the primary sequence. LBD is responsible for many functions, such as the ligand-regulated transcriptional activation function (AF-2), a strong dimerization interface, another NLS, and often a repression function.

**(F) C-terminal domain (CTD):** This domain is not always present. Its structure and function are vaguely known and its sequence is extremely variable.

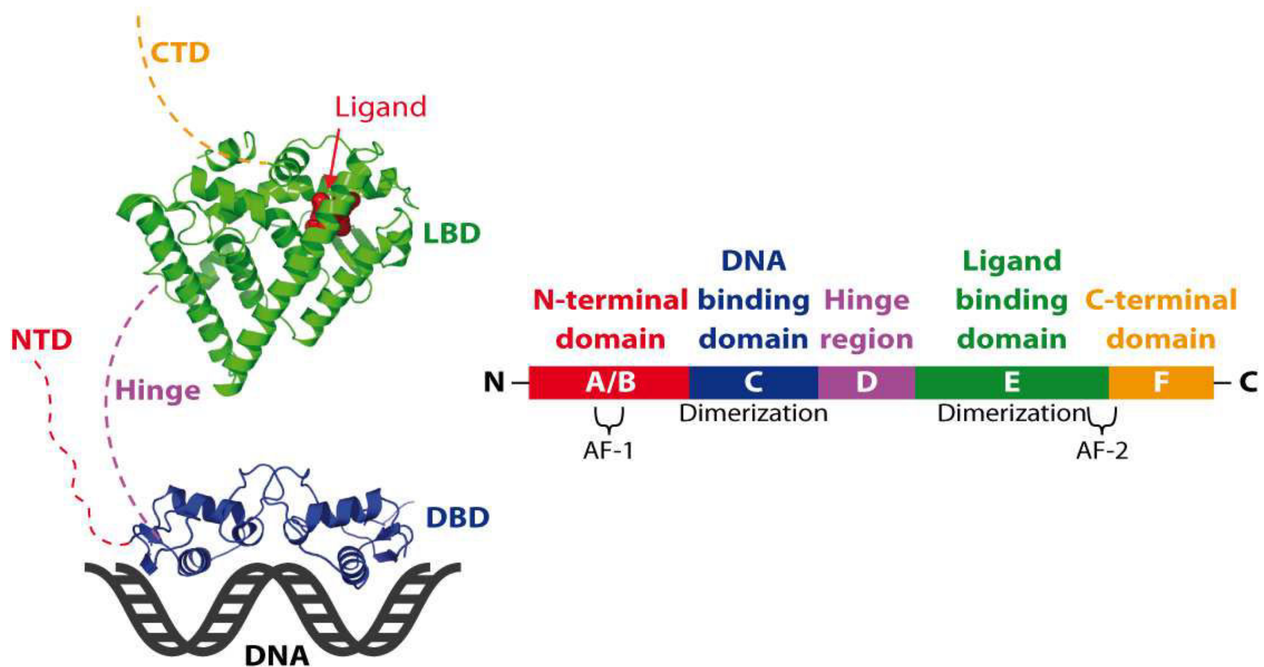
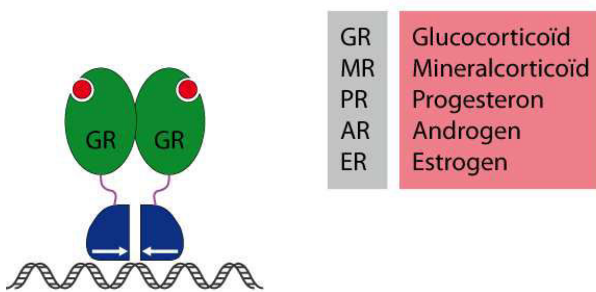
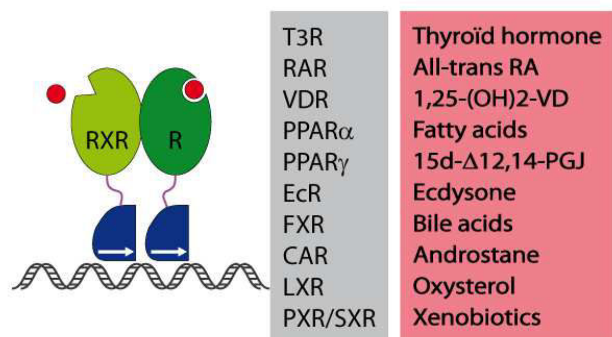


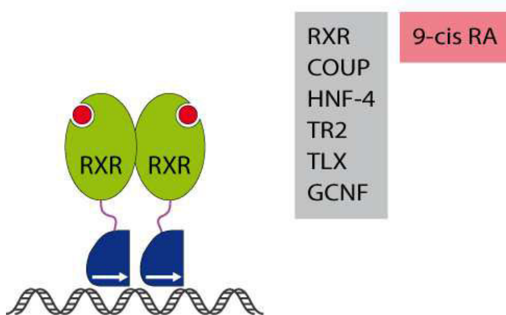

**Figure 6: structural organization of Nuclear Receptors**
*The six domains (A–F) of nuclear receptors: the N-terminal domain (NTD), the DNA binding domain (DBD) and the ligand binding domain (LDB), plus the hinge region and the C-terminal domain.*

### E.2.b. Nuclear receptor classification

The nuclear receptor superfamily is divided into four classes based on the nature of the ligand and on their dimerization and DNA binding properties (Mangelsdorf et al., 1995) (Figure 7). Class I receptors include the steroid receptor which function as ligand induced homodimers and bind to DNA half-sites organized as inverted repeats. It includes the progesterone receptor (PR), the estrogen receptor (ER), the glucocorticoid receptor (GR) or the androgen receptor (AR). Class II receptors heterodimerize with RXR and characteristically bind to direct repeats. It includes the thyroid receptor (TR), vitamin D receptor (VDR), the retinoic acid receptor (RAR) and the peroxisome proliferator-activated receptor (PPAR). Class III receptors bind primarily to direct repeats as homodimers. Class IV receptors typically bind to extended care sites as monomers. Most of the orphan receptors (no ligand identified yet) fail into the last two classes.



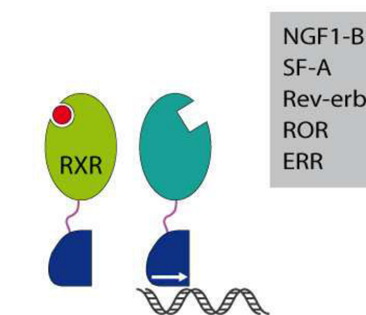




**Figure 7: Classification of Nuclear Receptors**
*Nuclear receptors can be grouped into four classes according to their ligand binding, DNA binding and dimerization properties. Shown are representative receptors for each group.*

### E.2.c. Retinoic acids receptors

Nuclear retinoic acid receptors (RARs) are transcriptional regulators controlling the expression of specific subsets of genes, in a ligand-dependent manner. RARs consist of three

subtypes (RARα, RARβ, RARγ) encoded by separate genes. They heterodimerize with retinoid X receptors (RXRs), which also exist as three subtypes (RXRα, RXRβ, RXRγ).

Compounds that bind RARs and modulate their activity are referred to as retinoids. Retinoids includes both natural and synthetic derivatives of vitamin A (Sporn et al., 1976). They are hydrophobic, lipid-soluble and small molecules that can easily cross the lipid bi-layer of cell membranes. Vitamin A regulates a wide variety of essential biological processes, such as embryonic morphogenesis and organogenesis, cell growth arrest, differentiation and apoptosis. By consequence, deficiency in Vitamin A can leads to neonatal growth retardation and a large number of congenital malformations (Sommer, 2008).

### E.2.d. Transcription of RAR target genes

The transcriptional regulation of retinoic acid (RA) target genes involves several dynamic, sequential, and coordinated steps (Dilworth and Chambon, 2001). In the absence of ligand, RAR/RXR heterodimers binds to retinoic acid response elements (RAREs), located in the regulatory region of the target gene. Classical RAREs are composed of two direct repeats of a core hexameric motif, PuG(G/T)TCA separated by 5 base pairs (Leid et al., 1992). The heterodimer RXR/RAR is associated, via the helix H12 of the RAR LBD, with co-repressors complexes, such as NCoR (Nuclear receptor corepressor) and SMRT (silencing mediator of retinoic acid and thyroid hormone). These co-repressors serve as a platform for the recruitment of the histone deacetylases (HDAC) which deacetylate the lysine of histones, thus maintaining the repressed chromatin state (Perissi et al., 2010).

Upon ligand binding, RAR undergoes conformational changes which results in H12 helix reorientation. This induces co-repressors dissociation and co-activators recruitment. In turn, co-activators facilities the recruitment of several other complexes implicated in chromatin remodeling, such as histone acetyl-transferases (HAT) or histone methyl-transferase (HMT). These complexes further allow the histone modifications leading to chromatin decompaction, for example H3K4 methylation and H3K9 acetylation. Once activated, RAR then recruits the transcriptional machinery which includes Mediator complex, RNA polymerase II, general transcription factors to start the transcription of the target gene (Figure 8).

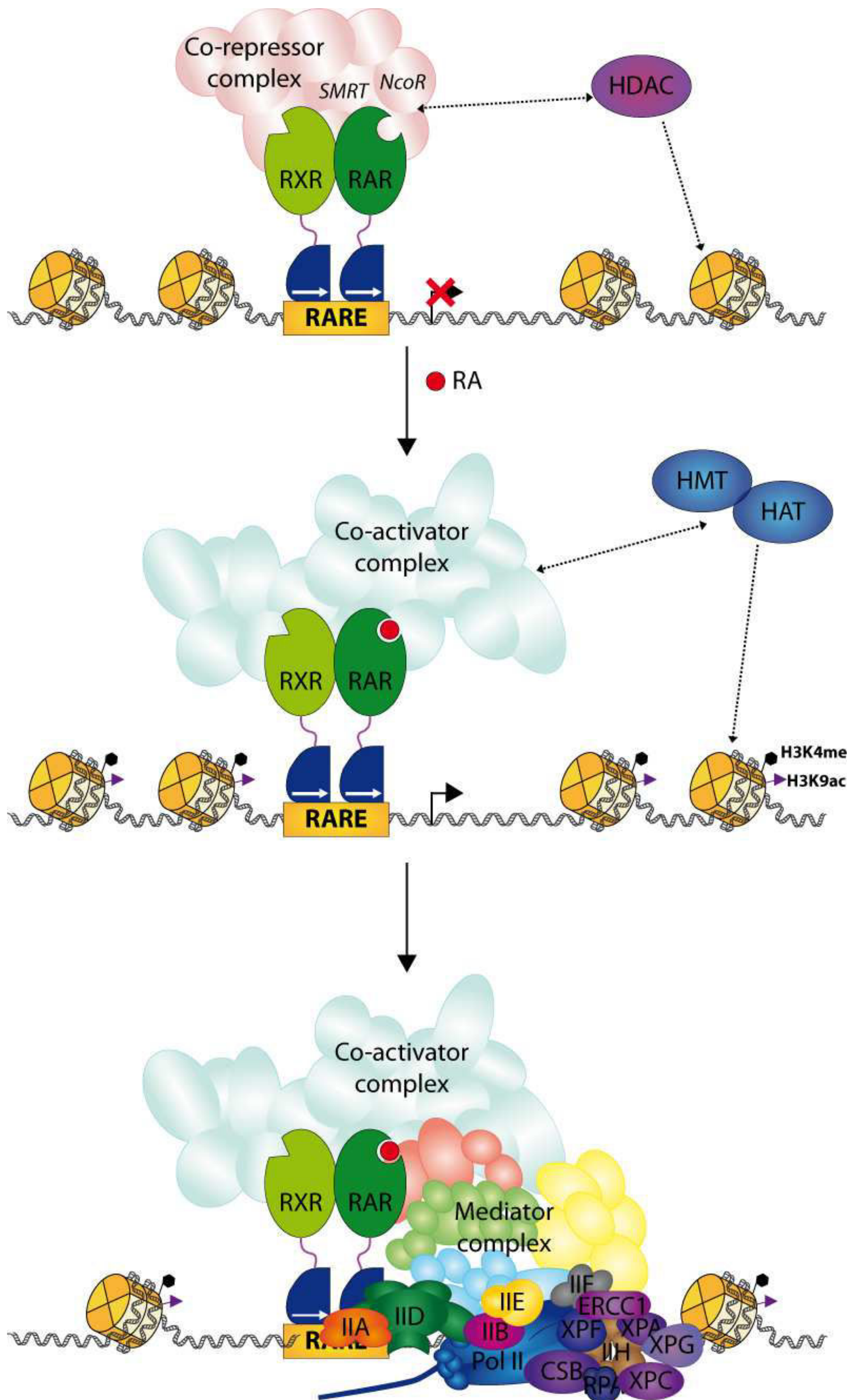



**Figure 8: Classical model of the activation of retinoic acid target genes**
*In the absence of ligand, RAR / RXR heterodimer is present on the DNA and associated with corepressor complexes, thus repressing gene expression. The RA binding induces the dissociation of corepressors and the recruitment of coactivators. Mediator, general transcription factor (GTFs), NER factors and Pol II are recruited and the chromatin is unpacked, thus allowing the initiation of transcription.*

Our group recently demonstrated that NER (Nucleotide Excision Repair) factors, initially characterized as part of DNA repair, are also recruited together with transcriptional apparatus on the promoter of nuclear receptor target genes. These factors are sequentially recruited in the following order: CSB/XPC, XPA/RPA, XPG and XPF/ERCC1 (Le May et al., 2010). They are required for optimal chromatin remodeling including histone posttranslational modifications (PTMs) as well as DNA demethylation and chromatin looping (Schmitz et al., 2009; Le May et al., 2012).

In addition to the above classical genomic effects, RA also has a number of nongenomic effects through the activation of several kinase cascades (Al Tanoury et al., 2013). Indeed, in response to RA, p38MAPK is activated, then translocates into the nucleus and phosphorylates MSK1. Activated MSK1 phosphorylates RARα at a serine located in the LBD and phosphorylation of this residue induces a structural conformation change of RARα. Subsequently, the Cyclin H subunit of TFIIH is recruited, allowing the phosphorylation of the NTD by the CDK7 kinase (Bastien et al., 2000; Gaillard et al., 2006; Bruck et al., 2009). This phosphorylation cascade is followed by the phosphorylation of H3 at serine 10 by MSK1. Consequently, the chromatin is reorganised, allowing the recruitment of phosphorylated RARα to response elements located in the promoter of target genes (Figure 9).

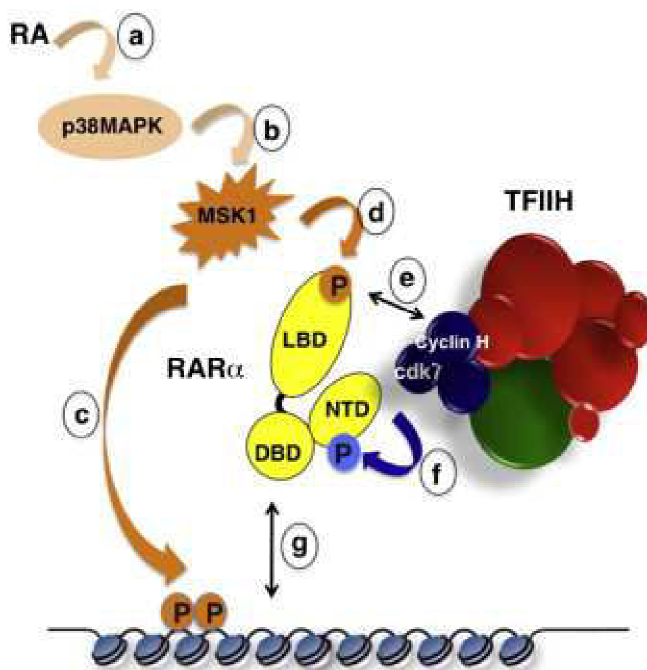



**Figure 9: Recapitulation of the phosphorylation cascade induced by RA.**
*In response to RA, p38MAPK (a) and the downstream protein kinase MSK1 (b) are activated. MSK1 phosphorylates histones H3 at Serine 10 (c) as well as RARα at a serine located in the LBD (d). Subsequently, the cyclin H subunit of the CAK subcomplex of TFIIH is recruited to an adjacent domain (e), allowing the formation of a RARα/TFIIH complex and the phosphorylation of the NTD by the CDK77 kinase (f). Finally, RARα phosphorylated and associated with TFIIH is recruited to response elements located in the promoter of target genes (g).*
*Figure from Duong and Rochette-Egly, 2011*

### E.2.e. Vitamin D receptors

Vitamin D is important in a variety of biological processes such as calcium homeostasis, cell proliferation and cell differentiation (Gonzalez-Parra et al., 2012; Goltzman et al., 2014). By consequence, vitamin D deficiency causes rickets and is associated with cardiovascular mortality, hypertension and immunity disorders. Vitamin D can be ingested from the diet as vitamin $D_3$ from animal or as vitamin $D_2$ from plants. It can also be produced from cholesterol by UV light on the skin (MacLaughlin et al., 1982). Vitamin D is then converted to its active form 1,25-dihydroxyvitamin D [1,25$(OH)_2$D] through a series of metabolic transformations.

The biological actions of vitamin D are exerted through the nuclear vitamin D receptor (VDR), a ligand-regulated transcription factor. VDR forms heterodimers with RXR which bind to vitamin D response elements (VDRE). Nuclear receptors from the same class have a similar mode of action and both RARs and VDRs belongs to RXR heterodimers class. Therefore, the classical genomic mechanism describe for RARs can also be applied to VDRs.

# II. REVIEW on the Mediator complex

*Review In preparation*

## The Mediator complex and associated human genetic disorders

DONNIO Lise-Marie[1], Baptiste BIDON[1], Jean-Marc EGLY*

**[1]The authors contributed equally to this review.**

IGBMC, Department of Functional Genomics and Cancer biology, CNRS/INSERM/Université de Strasbourg, BP 163, 67404 Illkirch Cedex, C.U. Strasbourg, FRANCE

*Corresponding authors
EGLY Jean-Marc : jean-marc.egly@igbmc.fr

Key words: Mediator complex, genetic disorders, Transcription, X-linked intellectual disability.

## ABSTRACT

Mediator is an evolutionary conserved multi-subunit complex that plays a central role in the regulation of RNA polymerase II transcribed genes. One important function of the MED is to convey essential information from DNA-bound transcription factors to the basal RNA polymerase II transcriptional machineries. This last decade, numerous studies have shown that genetic mutations in Mediator subunits or its partners cause various diseases, such as congenital malformations or neurodevelopmental disorders. After a brief description of the basic features of Mediator complex (discovery, structure and function in transcription initiation), we recapitulate the current body of knowledge concerning association of Mediator or its partner with specific genetic disorders. In some particular cases, the molecular etiologies underlying genotype-phenotype correlations are addressed.

## INTRODUCTION

Mediator (MED) is an evolutionarily conserved multi-protein complex that is a key regulator of protein-coding genes. MED contain 25 subunits in yeast and 30 or more in higher organisms. They are organized into three core modules (Head, Middle and Tail) and a dissociable Kinase module (Figure 1). MED regulate not only transcription initiation but also elongation [1], termination [2], mRNA processing [3] as well as chromatin remodeling [4,5]. In this review, we begin with a brief description of the discovery and structure of Mediator. We also summarize well-studied examples of Mediator interactions, outlining its function in transcription initiation. We then highlight examples in which genetic mutations of human Mediator subunits and its partners have been linked to specific pathological disorders.

## THE MEDIATOR COMPLEX

### *Discovery of the Mediator complex*

In the early 1990s, biochemical and genetic studies in *S. cerevisiae* led to the discovery of the MED complex [6]. The term ‘‘Mediator’’ was first proposed for its requirement in transcriptional activation of a reconstituted system [7]. Further studies demonstrated a general role in both basal and regulated transcription and, in the latter case, in both coactivation and corepression. After the first description in yeast, counterparts in other organisms, including mammals, have been identified. The first mammalian MED complex isolated was the human TRAP (Thyroid hormone Receptor-Associated Protein) complex. It was purified from HeLa cells as a protein complex that associates with the Thyroid hormone Receptor $\alpha$ (TR $\alpha$) in a ligand-dependent manner, and able to potentiate TR $\alpha$ -mediated transcription *in vitro* [8]. This was followed by isolation of other Mediator-like complexes [9–13]. Later, a proteomic analysis using Multidimensional Protein Identification Technology (MudPIT) showed all of these Mediator-like complex as representative of the same complex [14] and thus led to a unified nomenclature for MED subunits [15].

### *Composition and structure of the Mediator complex*

Comprehension of the mechanisms by which MED regulates transcription requires the understanding of its conformational behaviour. However, given its size, intrinsic flexibility, and composition heterogeneity, the high-resolution structure of the complete MED is still a challenge. Nonetheless, a number of structures of single MED subunits or subunit segments

have been resolved [16–19]. Moreover, a partial model of *S. cerevisiae* Middle module has been obtained by combining protein cross-linking information with partial crystal structures [20]. Until recently, the largest MED subcomplex characterized with high resolution was the yeast head module, which includes seven MED proteins [21,22].

Lately, an accurate electron microscopy (EM) map of the yeast MED (yMED) has been obtained [23]. The authors optimize specimen preparation and image analysis protocol to eliminate the problem of heterogeneity in MED conformation and/or composition. In addition, the localization of all yMED subunits into the EM map has been determined. Using similar approaches, they also calculate a precise EM map of human Mediator, which includes a number of subunit not found in yMED. Interestingly, the overall structure of MED appears to be largely conserved between yeast and human, and probably across eukaryotes.

## MEDIATOR COMPLEX IN TRANSCRIPTION INITIATION

Mediator complex interact with diverse components of the transcription machinery including several transcription factors (TFs), the RNA Polymerase II (Pol II) and the General Transcription Factors (GTFs) [24]. These interactions ultimately allow the MED to deliver outputs that range from maximal activation of genes to modulation of basal transcription.

### *Pol II*

Early work shows that the yMED associates closely with Pol II in an assembled complex termed Pol II holoenzyme [25–27]. EM studies then suggest that several subunits of Pol II, including Rpb1, Rpb2, Rpb3, Rpb6, and Rpb11, contact the middle or head module of MED [28]. Recently, an *in vivo* photo-crosslinking approach complemented by genetic analysis has identified a direct contact between Rpb3 and Med17, interaction essential for genome-wide Pol II recruitment [29]. Consistent with these findings, MED loss-of-function mutants have been found to compromise Pol II loading on the promoters of both induced and constitutively active genes [30,31].

### *TFIIA, TFIIB and TFIID.*

The first functional synergy between MED and TFIID in transcriptional activation was demonstrated in the early 2000s by Roeder lab [32,33]. Later, using immobilized template assays and extract depleted or supplemented with purified factors, MED was revealed to coordinate TFIID binding to promoter [34,35]. Similar experiments showed that TFIIB

recruitment is MED-dependent but this can be overcome by high level of TFIIB [36]. Moreover, Carey lab demonstrated that purified TFIIA/TFIID/MED bounded to promoter DNA generates a platform that supports active levels of PIC assembly and transcription, regardless of the presence of an activator [37]. A direct interaction between TFIID and MED26 was identified using a combination of biochemical and proteomics experiments. Interestingly, this interaction was not essential for TFIID recruitment, but rather appeared to regulate the timing of MED26 interaction with elongation factors [38].

***TFIIE and TFIIH.***

Mediator was shown to enhance the phosphorylation of the Pol II CTD by TFIIH in a yeast reconstituted transcription system, containing Pol II and basal factors [26]. This activity was later established to be a key event for promoter clearance and disruption of MED-Pol II interaction [39]. Interactions of MED with TFIIE and TFIIH have been revealed in yeast by two different reports. The first study demonstrated that the tail module subunit Med15 (Gal11) of Mediator complex binds stably to TFIIE and TFIIH [40,41], whereas the second study found a direct interaction between Med11 subunit and the Rad3 subunit of TFIIH [42].

***TFIIF.***

To this day, no direct interaction between MED and TFIIF has been convincingly demonstrated. However, the head module of yMED stably associates with a Pol II-TFIIF complex, but not with the polymerase or TFIIF alone [43]. Moreover, a cryo-EM analysis showed that both the presence of TFIIF and an activator-bound Mediator are required for a stable orientation of Pol II within a MED–Pol II–TFIIF assembly [44,45]. Recently, two reports revealed that MED can overcome the repressive effect of Pol II(G) complex, a distinct form of Pol II containing an additional tightly associated polypeptide called Gdown1 [46,47]. Pol II(G), unlike the normal Pol II, fails to activate transcription in the absence of MED [48]. The two reports differ in some important aspects [49], but both agree that Gdown1 represses transcription by inhibiting TFIIF functions and this effect is overcome by recruitment of activator-bound Mediator.

***Transcription factors (TFs).***

EM studies of human Mediator complexes revealed that the structure of the MED change markedly upon TFs binding. The structural comparison of complexes purified using

either a FLAG-tagged MED26 Mediator subunit, the VP16 activation domain or the SREBP-1a activator showed substantial differences in size and shape between complexes [50]. A recent study extend these finding by showing that two domains of p53, the C-terminus and the activation domain, interact with different MED subunits and thus the Mediator structure is differentially affected. Only the p53 activation domain elicit the conformation able to activate the stalled Pol II into a productive elongating state [51]. In addition, MED association with co-regulatory factors may diverge depending on the activator bound [52].

Besides, 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 MED1 and MED14 [53]. 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 [54].

***Others Mediator functions in transcription regulation***

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 block imposed by DSIF, a negative transcription elongation factor [55]. Furthermore, MED26 interacts first with TFIID in Pol II initiation complex and then exchanges TFIID for elongation complexes containing ELL/EAF and P-TEFb, in order to facilitate transition of Pol II from a stalled state into an elongation state [38].

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 gene loop formation. The Young lab show that Mediator interacts with NIPBL and cohesin complex to facilitate enhancer-promoter loops [4]. This interaction is important for cell-type-specific chromosome structure and gene expression [56]. The role of MED in gene looping has been further confirmed with the discovery of super-enhancers, which consist of clusters of enhancers. These domains depend upon MED and other transcription factors to control mammalian cell identity [57]. In addition to cohesin, MED can interact with ncRNA-a (non-coding RNA-activating) to regulate local gene expression and chromatin looping [58].

Mediator has also been reported to be important for transcription termination. Mukunda and Ansari demonstrated that Srb5/Med18 facilitate the recruitment of cleavage factor 1 (CF1) complex at the 3’ end of genes and thus the depletion of this MED subunit leads to an accumulation of Pol II near the 3’ end of genes [2]. These results indicate that MED is needed for termination of transcription through both the recruitment of termination factors and the release of Pol II. Consequently, as Med18 subunit was found at both the 5’ and 3’ ends of genes, it was postulated a role for MED as a bridge between the promoter and terminator region.

Finally, the Wang lab reported an association between MED23 and the RNA binding splicing regulator hnRNP L using a combination of affinity purification and mass spectrometry analysis [3]. The authors demonstrated that MED23 is involved in the regulation of a subset of hnRNP L-dependent alternative splicing events.

All these results showed that MED is an essential part of the transcription machinery, playing role in every step from loading of the PIC to the termination of transcription and splicing of mRNA.

## GENETIC DISORDERS ASSOCIATED WITH MEDIATOR

This last year, more and more studies have shown that mutations in MED subunit are associated with a wide range of human genetic disorders leading to congenital malformation and/or intellectual disability (Figure 1) [59–68].

### ***Med17 and infantile cerebral and cerebellar atrophy.***

Microcephaly is defined as a small cranium with a significant reduction of the occipital-frontal head circumference compared with age, sex and ethnicity matched controls. Microcephaly can be present at birth or may appear later, in the first years of life (source: National Institutes of Neurological Disorders and Stroke). Few years ago, a specific form of microcephaly within the Caucasus Jewish community has been associated with mutation in MED17 [64]. This association was discovered through the study of five infants from four unrelated families who manifested shortly after birth progressive microcephaly, spasticity, epilepsy and severe developmental retardation. Brain scans revealed cerebral and cerebellar atrophy with severe myelination defect, small thalami and a thin brainstem. By genetic analysis,

a homozygous missense mutation in MED17 (p.L371P) was found to segregate with the disease state.

MED17, a head module's Mediator subunit, has a central role in Mediator architecture and function. It is critical for head module assembly [21,43] and overall Mediator integrity by forming the largest contact between the Head and the rest of Mediator [23]. Moreover, as described in introduction, the interaction of Med17 with Rpb3 Pol II subunit of *S. Cerevisiae* is essential for genome-wide Pol II recruitment in vivo [29]. This may explain transcription impairment and lethality of Med17 inactivation in yeast and Drosophila [69–71]. Additionally, MED17 is an established physical target of the transcription factors p53 and NF-κB [72,73] and also of the DNA repair proteins Rad2/XPG [74].

Further studies will be required to understand how Med17 mutation might impact these critical functions and lead to infantile cerebral and cerebellar atrophy. Considering the normal prenatal development of patient and postnatal white matter deficiency, L371P missense mutation is speculate to disrupt a critical function of MED17 in controlling genes important for oligodendrocyte development, a process beginning only after birth in human.

***Med23 and nonsyndromic intellectual disability.***

Intellectual disability (ID) is defined as a significantly reduced ability to understand new or complex information and to learn and apply new skills. It begins before adulthood and results in a reduced ability to cope independently (world health organization definition). The majority of patient with ID have no other clinical abnormality. To date, only 15% of ID can be attributed to environmental factor and only 30-35% to known genetic abnormalities. A recent work in our lab has uncovered a direct link between a nonsyndromic intellectual disability and MED23, one of the tail module's MED subunits [66]. 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 cosegregate perfectly with the disease and was not found in control chromosomes. A second family in United States with a pair of brother affected by ID was recently discovered to have also a MED23 gene defect (Children's Neurological Solutions Foundation).

Med23 was originally identified as a genetic suppressor of a hyperactive ras phenotype in *C. Elegans* [75] and mediates the response of *EGR1* gene, an immediate early gene (IEG), to serum mitogens [54,76–78]. In line with such a function, our group found that the mutation

p.R617Q in MED23 alters the interaction between enhancer-bound transcription factors and MED, leading to transcriptional dysregulation of mitogen-responsive IEGs. These genes are important for brain development and functioning. However, further characterization of the serum response pathway will be required to understand the precise mechanism of MED23 induced nonsyndromic intellectual disability.

***Med25 and Charcot–Marie–Tooth disease.***

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 [79]. CMT, first described in 1886, was named after the three physicians who first identified it (Jean-Martin Charcot and his pupil Pierre Marie in France, and Howard Henry Tooth in England) [80,81]. Two major CMT forms are distinguishable based on electrophysiological and pathological criteria: the demyelinating (CMT1) and the axonal (CMT2) type. Among all forms of CMT, the autosomal recessive axonal is very rare (ARCMT2) and at this time three causative genes have been identified: Lamin A/C, GDAP1 and MED25. The association of the MED subunit with ARCMT2 was established through investigation of an extended Costa Rican family with Spanish and Amerindian ancestor [63]. 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.

MED25 contains several notable domains: a von Willebrand factor type A domain (VWA, residues 1–228), a conserved region containing a prostate tumour over-expressed protein 1 domain (PTOV, residues 395–545) and a NR box (LXXLL motif, residues 646–650). These domains allow MED25 to interact with multiple proteins, such as the histone acetyltransferase CBP through the PTOV domain and RAR (retinoic acid receptor) in a ligand-dependent manner through the NR box. These both interactions are important for recruitment of MED complex to retinoic acid (RA)-responsive genes [82]. The p.A335V mutation probably alters the structure of the protein and could maybe compromise this process. Consequently, considering that RA is involved in the maintenance of adult neurons [83], the axonal degeneration symptom observed in CMT disease could be explained by a dysregulation of RA-target genes. But evidence to

support this hypothesis is currently lacking. The molecular basis underlying the appearance of ARCMT2 caused by the p.A335V mutation in MED25 thus remains to be established.

## MED12 AND X-LINKED MENTAL RETARDATION

MED12, a 230kDa Mediator subunit located at Xq13.1, belongs to the kinase module, also composed of MED13, CDK8 and Cyclin C (CycC). By consequence, RNAi-mediated MED12 depletion in HeLa cells lead to a corresponding reduction in the steady-state levels of CDK8 and CycC proteins as well as their stable incorporation into Mediator [84]. At this time, the structure of MED12 protein is not solved but the amino acid sequence 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. Mutations in MED12 gene have been found to cause various disorders (Figure 2).

### *Opitz-Kaveggia and Lujan-Fryns syndrome*

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, imperforate anus and hypotonia [85]. This syndrome has been linked to a recurrent missense mutation (c.2881C>T, p.R961W) in MED12 gene [61]. Currently, this mutation has been found in 10 families with FG syndrome, including the original family (leading to a total of 23 affected males) [86]. Another MED12 missense mutation (c.2873G>A, p.G958E) has been reported in a family with three cousins affected by Opitz-Kaveggia syndrome [67].

Few years later, Lujan and Fryns independently described a X-linked mental retardation syndrome (commonly called Lujan-Fryns or Lujan syndrome, NIM #309520) characterized by intellectual disability, dysgenesis of the corpus callosum, macrocephaly, hypotonia and behavioural disturbance [87,88]. Later on, Schwartz et al. discovered a different sequence alteration (c.3020A>G, p.N1007S) in MED12 gene as causing Lujan syndrome [62].

Although both FG and Lujan syndrome are allelic and share several overlapping clinical manifestation, neither syndrome was originally considered in the differential diagnosis of the other (Table 1).

### *Ohdo syndrome*

Ohdo syndrome (NIM #300895) comprises a heterogeneous group of disorders characterized by intellectual disability and typical facial feature, including narrowing of the eye opening (blepharophimosis) [89]. The Maat-Kievit-Brunner type (OSMKB or X-linked Ohdo syndrome) was initially distinguished from the other type of Ohdo syndrome due to its X-linked inheritance [90]. The facial characteristics are 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 was performed in two families with the OSMKB type. In these two families, two different MED12 missense mutations (c.3443G>A, p.R1148H, or c.3493T>C p.S1165P) segregated with the phenotype. Subsequent analysis of a cohort of nine males with Ohdo syndrome, revealed another de novo missense mutation (c.5185C>A p.H1729N) in MED12 [68].

***Other MED12 disorders***

Recently, sequencing of all X-chromosome exons identified a novel mutation in MED12 (c.5898insC frameshift, p.S1967Qfsx84) in a large family with profound X-linked intellectual disability (10 males and 1 female affected) [91]. Dysmorphic features common to most affected males were long narrow face, high forehead, short philtrum and absent or severely-limited language (Table 1). Unlike the other MED12-related syndromes previously described, variable cognitive impairment was noted in the heterozygous females. The truncating mutation in this family seems to have a more severe effect on MED12 function than previous missense mutations.

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 [92,93]. Further analyses revealed a second, rare deletion polymorphism within the MED12 OPA domain (HOPA−15 bp) that appears also to be related with psychosis [94].

Finally, our lab is currently studying three new missense mutations (c.617G>A, p.R206Q; c.2692A>G, p.N898D and c.3884G>A, p.R1295H), which have been found in patients with X-linked mental retardation (paper under submission). Although individuals share some symptoms with previously described syndromes, they cannot be associated with one of them (Table 1).

How mutations of the same gene but at different position can lead to different disorders with overlapping symptoms?

***MED12 functions***

The etiological basis of Med12 associated disorders, while not fully resolved, is nonetheless suggested by studies that implicate MED12 in critical aspects of development. MED12 has been linked biochemically and genetically with the Notch, Wnt, and Sonic hedgehog signalling pathways that control key aspects of brain development and function, from initial patterning to neuronal plasticity [95–97]. Furthermore, Med12-deficient zebrafish embryos show defects in the development of brain, neural crest and ear, among other organs [98–100]. 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 [100]. This may produce some explanations to the neurological features observed in the MED12-related patients. Another work has also identified a role for Med12 during endoderm development. Defects occurring during this step of development may lead to the craniofacial characteristics and the digestive system defects observed in Med12-patients [101].

Mediator was shown to be involved in a protein network required for extraneuronal gene silencing. Indeed, MED12 within the Mediator have been demonstrated to link REST (RE1-silencing transcription factor) with the enzymatically active form of G9a in order to silence REST-target genes, in non-neuronal cells. G9a is a histone methyltransferase, which catalyzes histone H3K9 mono- and di-methylation [102]. Previous reports showed this modification as a platform for HP1 (Heterochromatin Protein 1) protein arrival. This event in turn induces the recruitment of the DNA-methylating enzyme DNMT1, leading through its activity, dimethylation of H3K9 histone, to long-term epigenetic gene silencing. Boyer lab has demonstrated that MED12 mutations (FG/p.R961W, Lujan/p.N100S and Ohdo/p. R1148H, /p.S1165P) compromised the ability of MED12 to mediate REST-direct recruitment of G9a and the imposition of the transcriptionnally repressive histone mark $H3K9me^2$. This deficiency is not due to a lower interaction between MED12 and G9a, but to impaired Mediator recruitment to REST-target genes [68,102].

In parallel, MED12 have been shown to interact via its PQL domain with AICD (Amyloid Precursor Protein Intracellular Domain) [103]. 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 suggests a role of MED12 in neuron maintenance.

Med12 modulates also Gli3-dependent Sonic hedgehog (SHH) signalling [104]. Mediator complex constraints the transcription of Shh target genes. Shh-induced Gli3 interacts with MED12, via its PQL domain, within the Mediator, resulting in the suppression of Mediator-constraint. Later, Zhou et al. found that the FG and Lujan MED12 mutations disrupt this constraint, thereby leading to enhanced Sonic hedgehog pathway activation [105].

## TRANSCRIPTION COMPLEX ASSOCIATED DISORDERS

A new set of evidences imputes a number of human diseases to genetic defects in Mediator binding partners. In this section, we will focus on disorders associated with mutations in the general transcription factor TFIIH, cohesin complex and transcription factors TCF4 or SOX9 (Figure 3 and Table 2). Although mutations arise in diverse genes of the transcription machinery, they lead to some common symptoms, suggesting that they have, at least partially, a similar effect on gene transcription.

### *Disorders associated with mutation in TFIIH and NER factors.*

As described at the beginning of this review, multiple studies have demonstrated interaction between Mediator and TFIIH [26,41,42]. Recently, a direct interaction between Med17 and XPG, a Nucleotide Excision Repair (NER) factor, has also been discovered [74]. NER factors were initially described to be involved in DNA repair. Mutations in TFIIH or NER factors lead to the autosomal recessive disorders Xeroderma Pigmentosum (XP), Trichothiodystrophy (TTD) or Cockayne Syndrome (CS). Xeroderma Pigmentosum is characterized by photosensitivity, premature skin aging, pigmentary changes and increased risk of skin cancers. In addition, patients with XP syndrome can develop progressive neurological degeneration, immature sexual development and dwarfism [106]. The clinical symptoms of Cockayne Syndrome are growth failure, impaired development of the nervous system and sun-sensitivity. CS is also characterized by a typical faces, ophthalmic and auditory disorders [107]. Patients with Trichothiodystrophy typically have dry and easily brittle hair and develop sterility, short stature and various neurological defects, including mental retardation [108].

For many years, XP, CS and TTD diseases were defined as DNA repair syndrome, as the NER pathway is reduced or sometimes absent in cells isolated from patients. However, some of the clinical features (including neurological and developmental defects) are difficult to explain as only DNA repair related disorders. The transcription factor TFIIH is a ten-subunit complex that has a fundamental role during the DNA nucleotide excision repair pathway as well as in transcription [109]. In addition, our team recently demonstrates that NER factors are also involved in active forms of transcription [110–112], forcing us to reconsider these diseases as both DNA repair and transcription related diseases.

***Cohesin complex and Cornelia de Lange syndrome.***

Cohesin complex, together with Mediator, facilitates DNA looping [4,113]. Mutations in cohesin complex are responsible for the rare developmental disorder Cornelia de Lange syndrome (CdLS), also known as Brachmann-de Lange syndrome. This syndrome is characterized by typical facial features, growth and mental retardation, upper limb defects, hirsutism and gastrointestinal dysfunction [114]. The first description have been made in 1849 by Vrolik, followed by Brachmann in 1916, but the diagnostic criteria were established by de Lange in 1933 [115–117]. More than half of individuals with CdLS present heterozygous mutations in the gene encoding the cohesin loader NIPBL [118,119]. However, mutations in the core structural components of the cohesin complex, SMC1 and SMC3 subunits, were also found at lower frequency (~5% and ~1% respectively) [120,121]. More recently, mutations in HDAC8, a SMC3 deacetylase, have been identified in six CdLS probands [122]. The mutations cause loss of HDAC8 activity, leading to SMC3 hyperacetylation and inefficient released of the cohesin complex from chromatin.

Cohesin is a dynamic multiprotein complex, which was identified for its role in the regulation of sister chromatids segregation, during both mitosis and meiosis. Recently, cohesin has also been demonstrated to play a critical role in DNA repair and gene expression [123]. In mammals, experiments of chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) reveal that a significant number of cohesin binding sites overlap with cell type-specific transcription factors binding sites and/or sites for CTCF (CCCTC-binding factor), a protein which help the formation of chromatin 3D structure during transcription [124–127].

Cells from CdLS individuals do not display sister chromatid cohesion defects, but present an increased DNA damage sensitivity [128] and a dysregulated gene expression [129,130]. Indeed, genome-wide assessment of transcription revealed a specific set of dysregulated genes

that correlates with disease severity. This dysregulation appear to be due to significantly decreased binding of cohesin on the promoter region [130]. These studies confirm the role of cohesin complex in gene expression and associate CdLS to transcription related diseases.

***TCF4 and Pitt-Hopkins syndrome.***

In our hands, Mediator was showed to interact with TCF4 [66]. 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 [131].

Transcription factor TFC4 (alias ITF2, SEF2 or E2-2) is a member of the basic Helix-Loop-Helix (bHLH) protein family that binds DNA at E-box motif CANNTG. It has important roles in number of developmental processes [132]. Although widely expressed, TCF4 is particularly high in the brain [133,134]. It was demonstrated to dimerize with several proneural transcription factors, such as MATH1 (also called Atoh1), ASCL1, NEUROD1 (alias BETA2) and NEUROD2 (alias NDRF) [135–137]. Proneural factors play important roles in the development of the nervous system by coordinating neuronal differentiation programs. As an illustration, homozygous Tcf4 knockout (Tcf4−/−) mice have reduced number of neurons forming the pontine nucleus [138]. Moreover, TCF4 is responsible for transcription regulation of IEGs, which are dysregulated in different Mediator related neurological disorder ([66], and paper under submission).

***SOX9 and campomelic dysplasia***

A recent study demonstrated that Med25 interact with Sox9 and this interaction augments Sox9 transcriptional activity [139]. Mutations in or near the SOX9 gene cause Campomelic Dysplasia (CMD), a severe disorder that affects development of the skeleton and reproductive system [140]. The name is derived from the Greek root ‘campo’ (or campto) meaning bent and ‘melia’ meaning limb. It is frequently lethal in the neonatal period due to respiratory insufficiency. Affected individuals have bowed lower limbs, external genitalia and distinctive facial features. Among survivors of CMD, neurological defects, including intellectual disability, are often seen [141].

SOX9 is a transcription factors that plays a pivotal role in the development of the skeleton and reproductive organs [142,143]. In addition, a recent study established a central role for SOX9 in neural stem cells specification and maintenance [144]. Further works will be required to better understand the diverse functions of SOX9 and thus improved our comprehension of CMD disorders etiology.

## CONCLUSION

Although a number of studies have greatly expanded our knowledge of Mediator complex, we are only beginning to understand the diversity of its role on transcription process. By consequence, further characterization of Mediator and its different binding partners would be necessary to improve our comprehension of the complex mechanisms that regulate the expression of protein coding genes at the right time and right amount in each cell types. Considering the number of neurological and neurodevelopmental disorder related to its subunits or its interacting partners, we could consider the Mediator as an essential player of the brain development. Advancing our knowledge on gene expression will then be a prerequisite to provide explanations for the phenotypes of patients baring mutations in the components of the transcription machinery.

## REFERENCES


1. Conaway RC, Conaway JW (2013) The Mediator complex and transcription elongation. *Biochim Biophys Acta BBA - Gene Regul Mech* **1829:** 69–75.
2. Mukundan B, Ansari A (2013) Srb5/Med18-mediated Termination of Transcription Is Dependent on Gene Looping. *J Biol Chem* **288**: 11384–11394.
3. Huang Y, Li W, Yao X, Lin Q, Yin J, Liang Y, Heiner M, Tian B, Hui J, Wang G (2012) Mediator Complex Regulates Alternative mRNA Processing via the MED23 Subunit. *Mol Cell* **45**: 459–469.
4. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**: 430–435.
5. Zhu X, Zhang Y, Bjornsdottir G, Liu Z, Quan A, Costanzo M, Davila Lopez M, Westholm JO, Ronne H, Boone C, et al. (2011) Histone modifications influence mediator interactions with chromatin. *Nucleic Acids Res* **39**: 8342–8354.
6. Kelleher RJ, Flanagan PM, Kornberg RD (1990) A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* **61**: 1209–1215.
7. Flanagan PM, Kelleher RJ, Sayre MH, Tschochner H, Kornberg RD (1991) A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **350**: 436–438.

8. Fondell JD, Ge H, Roeder RG (1996) Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc Natl Acad Sci U S A* **93**: 8329–8333.
9. Malik S, Gu W, Wu W, Qin J, Roeder RG (2000) The USA-Derived Transcriptional Coactivator PC2 Is a Submodule of TRAP/SMCC and Acts Synergistically with Other PCs. *Mol Cell* **5**: 753–760.
10. Näär AM, Beaurang PA, Zhou S, Abraham S, Solomon W, Tjian R (1999) Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature* **398**: 828–832.
11. Rachez C, Suldan Z, Ward J, Chang CP, Burakov D, Erdjument-Bromage H, Tempst P, Freedman LP (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.
12. Ryu S, Zhou S, Ladurner AG, Tjian R (1999) The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* **397**: 446–450.
13. Sun X, Zhang Y, Cho H, Rickert P, Lees E, Lane W, Reinberg D (1998) NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol Cell* **2**: 213–222.
14. Sato S, Tomomori-Sato C, Parmely TJ, Florens L, Zybailov B, Swanson SK, Banks CAS, Jin J, Cai Y, Washburn MP, et al. (2004) A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol Cell* **14**: 685–691.
15. Bourbon H-M, Aguilera A, Ansari AZ, Asturias FJ, Berk AJ, Bjorklund S, Blackwell TK, 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.
16. Baumli S, Hoeppner S, Cramer P (2005) A conserved mediator hinge revealed in the structure of the MED7.MED21 (Med7.Srb7) heterodimer. *J Biol Chem* **280**: 18171–18178.
17. Hoeppner S, Baumli S, Cramer P (2005) Structure of the Mediator Subunit Cyclin C and its Implications for CDK8 Function. *J Mol Biol* **350**: 833–842.
18. Larivière L, Geiger S, Hoeppner S, Röther S, Strässer K, Cramer P (2006) Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat Struct Mol Biol* **13**: 895–901.
19. Koschubs T, Seizl M, Larivière L, Kurth F, Baumli S, Martin DE, Cramer P (2009) Identification, structure, and functional requirement of the Mediator submodule Med7N/31. *EMBO J* **28**: 69–80.
20. Larivière L, Plaschka C, Seizl M, Petrotchenko EV, Wenzeck L, Borchers CH, Cramer P (2013) Model of the Mediator middle module based on protein cross-linking. *Nucleic Acids Res* **41**: 9266–9273.
21. Imasaki T, Calero G, Cai G, Tsai K-L, Yamada K, Cardelli F, Erdjument-Bromage H, Tempst P, Berger I, Kornberg GL, et al. (2011) Architecture of the Mediator head module. *Nature* **475**: 240–243.
22. Larivière L, Plaschka C, Seizl M, Wenzeck L, Kurth F, Cramer P (2012) Structure of the Mediator head module. *Nature* **492**: 448–451.
23. Tsai K-L, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Asturias FJ (2014) Subunit Architecture and Functional Modular Rearrangements of the Transcriptional Mediator Complex. *Cell* **157**: 1430–1444.
24. Borggrefe T, Yue X (2011) Interactions between subunits of the Mediator complex with gene-specific transcription factors. Semin *Cell Dev Biol* **22**: 759–768.

25. Thompson CM, Koleske AJ, Chao DM, Young RA (1993) A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**: 1361–1375.
26. Kim YJ, Björklund S, Li Y, Sayre MH, Kornberg RD (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**: 599–608.
27. Koleske AJ, Young RA (1994) An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**: 466–469.
28. Davis JA, Takagi Y, Kornberg RD, Asturias FA (2002) Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol Cell* **10**: 409–415.
29. Soutourina J, Wydau S, Ambroise Y, Boschiero C, Werner M (2011) Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo. *Science* **331**: 1451–1454.
30. Cantin GT, Stevens JL, Berk AJ (2003) Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc Natl Acad Sci U S A* **100**: 12003–12008.
31. Ansari SA, He Q, Morse RH (2009) Mediator complex association with constitutively transcribed genes in yeast. *Proc Natl Acad Sci U S A* **106**: 16734–16739.
32. Guermah M, Malik S, Roeder RG (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.
33. Guermah M, Tao Y, Roeder RG (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.
34. Johnson KM, Wang J, Smallwood A, Arayata C, Carey M (2002) TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev* **16**: 1852–1863.
35. Black JC, Choi JE, Lombardo SR, Carey M (2006) A mechanism for coordinating chromatin modification and preinitiation complex assembly. *Mol Cell* **23**: 809–818.
36. Baek HJ, Kang YK, Roeder RG (2006) Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J Biol Chem* **281**: 15172–15181.
37. Johnson KM, Carey M (2003) Assembly of a Mediator/TFIID/TFIIA Complex Bypasses the Need for an Activator. *Curr Biol* **13**: 772–777.
38. Takahashi H, Parmely TJ, Sato S, Tomomori-Sato C, Banks CAS, Kong SE, Szutorisz H, Swanson SK, Martin-Brown S, Washburn MP, et al. (2011) Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **14**6: 92–104.
39. Søgaard TMM, Svejstrup JQ (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.
40. Sakurai H, Fukasawa T (1997) Yeast Gal11 and Transcription Factor IIE Function through a Common Pathway in Transcriptional Regulation. *J Biol Chem* **272**: 32663–32669.
41. Sakurai H, Fukasawa T (2000) Functional Connections between Mediator Components and General Transcription Factors of Saccharomyces cerevisiae. *J Biol Chem* **275**: 37251–37256.
42. Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, Holstege F, Werner M (2008) Mediator-Dependent Recruitment of TFIIH Modules in Preinitiation Complex. Mol Cell **31**: 337–346.

43. Takagi Y, Calero G, Komori H, Brown JA, Ehrensberger AH, Hudmon A, Asturias F, Kornberg RD (2006) Head Module Control of Mediator Interactions. *Mol Cell* **23**: 355–364.
44. Bernecky C, Grob P, Ebmeier CC, Nogales E, Taatjes DJ (2011) Molecular architecture of the human Mediator-RNA polymerase II-TFIIF assembly. *PLoS Biol* **9**: e1000603.
45. Bernecky C, Taatjes DJ (2012) Activator–Mediator Binding Stabilizes RNA Polymerase II Orientation within the Human Mediator–RNA Polymerase II–TFIIF Assembly. *J Mol Biol* **417**: 387–394.
46. Cheng B, Li T, Rahl PB, Adamson TE, Loudas NB, Guo J, Varzavand K, Cooper JJ, Hu X, Gnatt A, et al. (2012) Functional association of Gdown1 with RNA polymerase II poised on human genes. *Mol Cell* **45**: 38–50.
47. Jishage M, Malik S, Wagner U, Uberheide B, Ishihama Y, Hu X, Chait BT, Gnatt A, Ren B, Roeder RG (2012) Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. *Mol Cell* **45**: 51–63.
48. Hu X, Malik S, Negroiu CC, Hubbard K, Velalar CN, Hampton B, Grosu D, Catalano J, Roeder RG, Gnatt A (2006) A Mediator-responsive form of metazoan RNA polymerase II. *Proc Natl Acad Sci U S A* **103**: 9506–9511.
49. Espinosa JM (2012) Get back TFIIF, don't let me Gdown1. *Mol Cell* **45**: 3–5.
50. Taatjes DJ, Näär AM, Andel F, Nogales E, Tjian R (2002) Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* **295**: 1058–1062.
51. Meyer KD, Lin S-C, Bernecky C, Gao Y, Taatjes DJ (2010) p53 activates transcription by directing structural shifts in Mediator. *Nat Struct Mol Biol* **17**: 753–760.
52. Ebmeier CC, Taatjes DJ (2010) Activator-Mediator binding regulates Mediator-cofactor interactions. *Proc Natl Acad Sci U S A* **107**: 11283–11288.
53. Chen W, Rogatsky I, Garabedian MJ (2006) MED14 and MED1 differentially regulate target-specific gene activation by the glucocorticoid receptor. *Mol Endocrinol Baltim Md* **20**: 560–572.
54. Balamotis MA, Pennella MA, Stevens JL, Wasylyk B, Belmont AS, Berk AJ (2009) Complexity in transcription control at the activation domain-mediator interface. *Sci Signal* **2**: ra20.
55. Malik S, Barrero MJ, 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.
56. Phillips-Cremins JE, Sauria MEG, Sanyal A, Gerasimova TI, Lajoie BR, Bell JSK, Ong C-T, Hookway TA, Guo C, Sun Y, et al. (2013) Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. *Cell* **153**: 1281–1295.
57. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **153**: 307–319.
58. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R (2013) Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494**: 497–501.
59. Berti L, Mittler G, Przemeck GKH, 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.
60. Muncke N (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.

61. Risheg H, Graham JM, Clark RD, Rogers RC, Opitz JM, Moeschler JB, Peiffer AP, May M, Joseph SM, Jones JR, et al. (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat Genet* **39**: 451–453.
62. Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, Friez MJ, Futreal PA, 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.
63. Leal A, Huehne K, Bauer F, Sticht H, Berger P, Suter U, Morera B, Valle G, Lupski JR, 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.
64. 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.
65. Mukhopadhyay A, Kramer JM, Merkx G, Lugtenberg D, Smeets DF, Oortveld MAW, Blokland EAW, 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.
66. Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, Egly J-M, Colleaux L (2011) MED23 Mutation Links Intellectual Disability to Dysregulation of Immediate Early Gene Expression. *Science* **333**: 1161–1163.
67. Rump P, Niessen R, Verbruggen K, Brouwer O, de Raad M, Hordijk R (2011) A novel mutation in MED12 causes FG syndrome (Opitz-Kaveggia syndrome). *Clin Genet* **79**: 183–188.
68. Vulto-van Silfhout AT, de Vries BBA, van Bon BWM, Hoischen A, Ruiterkamp-Versteeg M, Gilissen C, Gao F, van Zwam M, Harteveld CL, van Essen AJ, et al. (2013) Mutations in MED12 Cause X-Linked Ohdo Syndrome. *Am J Hum Genet* **92**: 401–406.
69. Boube M (2000) Drosophila homologs of transcriptional mediator complex subunits are required for adult cell and segment identity specification. *Genes Dev* **14**: 2906–2917.
70. Linder T, Zhu X, Baraznenok V, Gustafsson CM (2006) The classical srb4-138 mutant allele causes dissociation of yeast Mediator. *Biochem Biophys Res Commun* **349**: 948–953.
71. Thompson CM, Young RA (1995) General requirement for RNA polymerase II holoenzymes in vivo. *Proc Natl Acad Sci U S A* **92**: 4587–4590.
72. Van Essen D, Engist B, Natoli G, Saccani S (2009) Two Modes of Transcriptional Activation at Native Promoters by NF-κB p65. *PLoS Biol* **7**: e73.
73. Ito M, Yuan C-X, Malik S, Gu W, Fondell JD, Yamamura S, Fu Z-Y, Zhang X, Qin J, Roeder RG (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.
74. Eyboulet F, Cibot C, Eychenne T, Neil H, Alibert O, Werner M, Soutourina J (2013) Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment. Genes Dev 27: 2549–2562.
75. Singh N, 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.
76. Stevens JL (2002) Transcription Control by E1A and MAP Kinase Pathway via Sur2 Mediator Subunit. *Science* **296**: 755–758.

77. Wang G, Balamotis MA, Stevens JL, Yamaguchi Y, Handa H, Berk AJ (2005) Mediator Requirement for Both Recruitment and Postrecruitment Steps in Transcription Initiation. Mol *Cell* **17**: 683–694.
78. Wang W, Yao X, Huang Y, Hu X, Liu R, Hou D, Chen R, Wang G (2013) Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. *Transcription* **4**: 39–51.
79. Reilly MM, Murphy SM, Laurá M (2011) Charcot-Marie-Tooth disease. *J Peripher Nerv Syst JPNS* **16**: 1–14.
80. Charcot JM, Marie P (1886) Sur une forme particulière d'atrophie musculaire progressive, souvent familiale débutant par les pieds et les jambes et atteignant plus tard les mains. *Rev Médicale* **6**: 97–138.
81. Tooth HH (1886) The peroneal type of progressive muscular atrophy. *Dissertation.*
82. Lee H-K, Park U-H, Kim E-J, Um S-J (2007) MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J* **26**: 3545–3557.
83. Maden M (2007) Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nat Rev Neurosci* **8**: 755–765.
84. Kim S, Xu X, Hecht A, Boyer TG (2006) Mediator Is a Transducer of Wnt/beta-Catenin Signaling. *J Biol Chem* **281**: 14066–14075.
85. Opitz JM, Kaveggia EG (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.
86. Clark RD, Graham JM, Friez MJ, Hoo JJ, Jones KL, McKeown C, Moeschler JB, Raymond FL, Rogers RC, Schwartz CE, 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.
87. Fryns JP, Buttiens M (1987) X-linked mental retardation with marfanoid habitus. *Am J Med Genet* **28**: 267–274.
88. Lujan JE, Carlin ME, Lubs HA (1984) A form of X-linked mental retardation with marfanoid habitus. *Am J Med Genet* **17:** 311–322.
89. Ohdo S, Madokoro H, Sonoda T, Hayakawa K (1986) Mental retardation associated with congenital heart disease, blepharophimosis, blepharoptosis, and hypoplastic teeth. *J Med Genet* **23**: 242–244.
90. 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.
91. Lesca G, Moizard M-P, Bussy G, Boggio D, Hu H, Haas SA, Ropers H-H, Kalscheuer VM, 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.
92. DeLisi LE, Smith AB, Razi K, Stewart J, Wang Z, Sandhu HK, Philibert RA (2000) Investigation of a candidate gene for schizophrenia on Xq13 previously associated with mental retardation and hypothyroidism. *Am J Med Genet* **96**: 398–403.
93. Philibert RA, Bohle P, Secrest D, Deaderick J, Sandhu H, Crowe R, Black DW (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.

94. Beyer KS, Klauck SM, Benner A, Poustka F, Poustka A (2002) Association studies of theHOPA dodecamer duplication variant in different subtypes of autism. *Am J Med Genet* **114**: 110–115.
95. Moghal N (2003) A component of the transcriptional mediator complex inhibits RAS-dependent vulval fate specification in C. elegans. *Development* **130**: 57–69.
96. 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.
97. Yoda A (2005) Components of the transcriptional Mediator complex are required for asymmetric cell division in C. elegans. *Development* **132**: 1885–1893.
98. Hong S-K, Haldin CE, Lawson ND, Weinstein BM, Dawid IB, Hukriede NA (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.
99. Rau MJ, Fischer S, Neumann CJ (2006) Zebrafish Trap230/Med12 is required as a coactivator for Sox9-dependent neural crest, cartilage and ear development. *Dev Biol* **296**: 83–93.
100. Wang X, Yang N, Uno E, Roeder RG, Guo S (2006) A subunit of the mediator complex regulates vertebrate neuronal development. *Proc Natl Acad Sci* **103**: 17284–17289.
101. Shin CH, Chung W-S, Hong S-K, Ober EA, Verkade H, Field HA, Huisken J, Y. R. Stainier D (2008) Multiple roles for Med12 in vertebrate endoderm development. *Dev Biol* **317**: 467–479.
102. Ding N, Zhou H, Esteve P-O, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S, et al. (2008) Mediator Links Epigenetic Silencing of Neuronal Gene Expression with X-Linked Mental Retardation. *Mol Cell* **31**: 347–359.
103. Xu X, Zhou H, Boyer TG (2011) Mediator is a transducer of amyloid-precursor-protein-dependent nuclear signalling. *EMBO Rep* **12**: 216–222.
104. Zhou H, Kim S, Ishii S, Boyer TG (2006) Mediator Modulates Gli3-Dependent Sonic Hedgehog Signaling. *Mol Cell Biol* **26**: 8667–8682.
105. Zhou H, Spaeth JM, Kim NH, Xu X, Friez MJ, Schwartz CE, Boyer TG (2012) MED12 mutations link intellectual disability syndromes with dysregulated GLI3-dependent Sonic Hedgehog signaling. *Proc Natl Acad Sci* **109**: 19763–19768.
106. Kraemer KH, Levy DD, Parris CN, Gozukara EM, Moriwaki S, Adelberg S, Seidman MM (1994) Xeroderma pigmentosum and related disorders: examining the linkage between defective DNA repair and cancer. *J Invest Dermatol* **103**: 96S – 101S.
107. Nance MA, Berry SA (1992) Cockayne syndrome: review of 140 cases. *Am J Med Genet* **42**: 68–84.
108. Hashimoto S, Egly JM (2009) Trichothiodystrophy view from the molecular basis of DNA repair/transcription factor TFIIH. *Hum Mol Genet* **18**: R224–R230.
109. Compe E, Egly J-M (2012) TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* **13**: 343–354.
110. Le May N, Mota-Fernandes D, Vélez-Cruz R, Iltis I, Biard D, Egly JM (2010) 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.
111. Le May N, Egly J-M, Coin F (2010) True Lies: The Double Life of the Nucleotide Excision Repair Factors in Transcription and DNA Repair. *J Nucleic Acids* **2010**: 1–10.
112. Le May N, Fradin D, Iltis I, Bougnères P, 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.

113. Muto A, Ikeda S, Lopez-Burks ME, Kikuchi Y, Calof AL, Lander AD, Schilling TF (2014) Nipbl and mediator cooperatively regulate gene expression to control limb development. *PLoS Genet* **10**: e1004671.
114. Liu J, Krantz I (2009) Cornelia de Lange syndrome, cohesin, and beyond. *Clin Genet* **76**: 303–314.
115. Brachmann W (1916) Ein Fall von symmetrischer Monodactylie durch Ulnadefect, mit symmetrischer Flughautbildung in den Ellenbeugen, sowie anderen Abnormalitaten. *Jarb Kinder Phys Erzie* **84**: 225–235.
116. De Lange C (1933) Sur un type nouveau de dégénération (Typus amstelodamensis). *Arch Médecine Enfants* 713–719.
117. Vrolik W (1849) Tabulae ad illustrandam embryogenesin hominis et mammalium tam naturalem quam abnormem.
118. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, et al. (2004) Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nat Genet* **36**: 631–635.
119. Tonkin ET, Wang T-J, Lisgo S, Bamshad MJ, Strachan T (2004) NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet* **36**: 636–641.
120. Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodríguez C, Arnedo M, Loeys B, Kline AD, et al. (2007) Mutations in Cohesin Complex Members SMC3 and SMC1A Cause a Mild Variant of Cornelia de Lange Syndrome with Predominant Mental Retardation. *Am J Hum Genet* **80**: 485–494.
121. Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L (2006) X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. *Nat Genet* **38**: 528–530.
122. Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y, Clark D, et al. (2012) HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* **489**: 313–317.
123. Mehta GD, Kumar R, Srivastava S, Ghosh SK (2013) Cohesin: Functions beyond sister chromatid cohesion. *FEBS Lett* **587**: 2299–2312.
124. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, et al. (2008) Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell* **132**: 422–433.
125. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT (2010) A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res* **20**: 578–588.
126. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, et al. (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**: 796–801.
127. Zuin J, Franke V, van IJcken WFJ, van der Sloot A, Krantz ID, van der Reijden MIJA, Nakato R, Lenhard B, Wendt KS (2014) A Cohesin-Independent Role for NIPBL at Promoters Provides Insights in CdLS. *PLoS Genet* **10**: e1004153.
128. Vrouwe MG, Elghalbzouri-Maghrani E, Meijers M, Schouten P, Godthelp BC, Bhuiyan ZA, Redeker EJ, Mannens MM, Mullenders LHF, Pastink A, et al. (2007) Increased DNA damage sensitivity of Cornelia de Lange syndrome cells: evidence for impaired recombinational repair. *Hum Mol Genet* **16**: 1478–1487.
129. Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, et al. (2009) Multiple Organ System Defects and Transcriptional

Dysregulation in the Nipbl+/− Mouse, a Model of Cornelia de Lange Syndrome. *PLoS Genet* **5**: e1000650.
130. Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, et al. (2009) Transcriptional Dysregulation in NIPBL and Cohesin Mutant Human Cells. *PLoS Biol* **7:** e1000119.
131. Peippo M, Ignatius J (2011) Pitt-Hopkins Syndrome. *Mol Syndromol*. **2**, 171–180
132. Forrest MP, Hill MJ, Quantock AJ, Martin-Rendon E, Blake DJ (2014) The emerging roles of TCF4 in disease and development. *Trends Mol Med* **20**: 322–331.
133. Sepp M, Kannike K, Eesmaa A, Urb M, Timmusk T (2011) Functional Diversity of Human Basic Helix-Loop-Helix Transcription Factor TCF4 Isoforms Generated by Alternative 5′ Exon Usage and Splicing. *PLoS ONE* **6**: e22138.
134. Soosaar A, Chiaramello A, Zuber MX, Neuman T (1994) Expression of basic-helix-loop-helix transcription factor ME2 during brain development and in the regions of neuronal plasticity in the adult brain. *Mol Brain Res* **25**: 176–180.
135. Brzózka MM, Radyushkin K, Wichert SP, Ehrenreich H, Rossner MJ (2010) Cognitive and Sensorimotor Gating Impairments in Transgenic Mice Overexpressing the Schizophrenia Susceptibility Gene Tcf4 in the Brain. *Biol Psychiatry* **68**: 33–40.
136. Persson P, Jögi A, Grynfeld A, Påhlman S, Axelson H (2000) HASH-1 and E2-2 are expressed in human neuroblastoma cells and form a functional complex. *Biochem Biophys Res Commun* **274**: 22–31.
137. Ravanpay AC, Olson JM (2008) E protein dosage influences brain development more than family member identity. *J Neurosci Res* **86**: 1472–1481.
138. Flora A, Garcia JJ, Thaller C, Zoghbi HY (2007) The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proc Natl Acad Sci U S A* **104**: 15382–15387.
139. Nakamura Y, Yamamoto K, He X, Otsuki B, Kim Y, Murao H, Soeda T, Tsumaki N, Deng JM, Zhang Z, et al. (2011) Wwp2 is essential for palatogenesis mediated by the interaction between Sox9 and mediator subunit 25. *Nat Commun* **2**: 251.
140. Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Bricarelli FD, Keutel J, Hustert E, et al. (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* **79**: 1111–1120.
141. Mansour S (2002) The phenotype of survivors of campomelic dysplasia. *J Med Genet* **39**: 597–602.
142. Bi W, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B (1999) Sox9 is required for cartilage formation. *Nat Genet* **22**: 85–89.
143. Kobayashi A, Chang H, Chaboissier M-C, Schedl A, Behringer RR (2005) Sox9 in testis determination. *Ann N Y Acad Sci* **1061**: 9–17.
144. Scott CE, Wynn SL, Sesay A, Cruz C, Cheung M, Gaviro M-VG, Booth S, Gao B, Cheah KSE, Lovell-Badge R, et al. (2010) SOX9 induces and maintains neural stem cells. *Nat Neurosci* **13**: 1181–1189.

# FIGURE 1

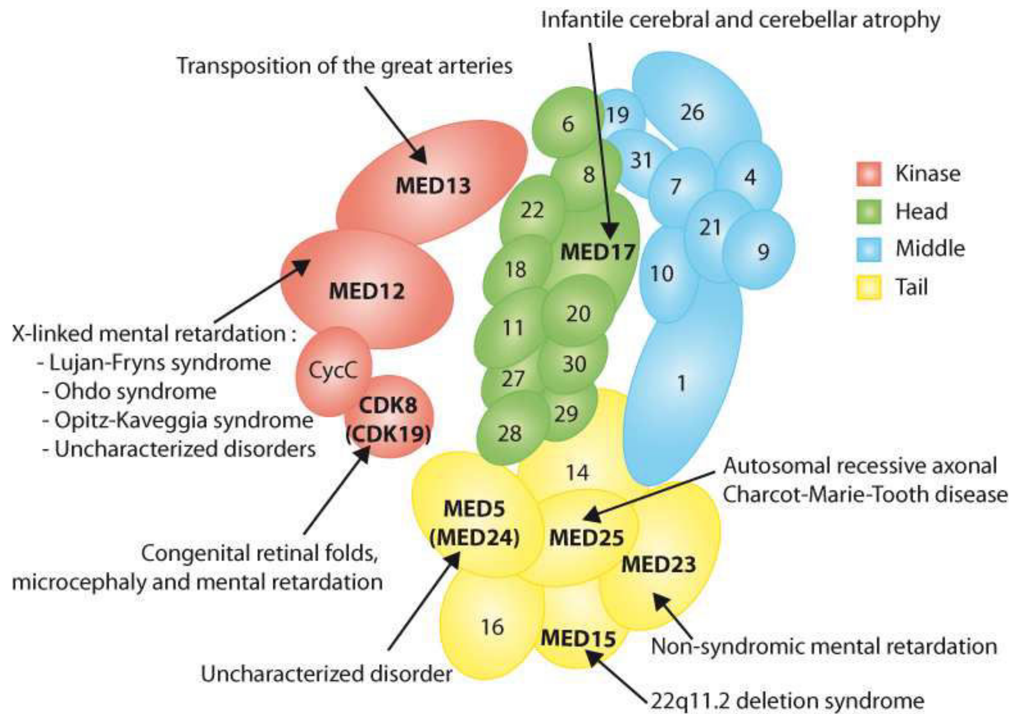



**Figure 1: Modular structure of human Mediator and genetic disorders related to mutation in MED subunits**

A composite depiction of the subunit structure of the human Mediator complex is shown. Note that the relative placement of the subunits in the subcomplexes is based on published binary interaction and partial structural data [23]. It is primarily for illustration; in this 2D plan, some of the localization, interaction and size of proteins can not be respect. Genetic disorders associated with mutation in some Mediators subunits are also indicated.

# FIGURE 2

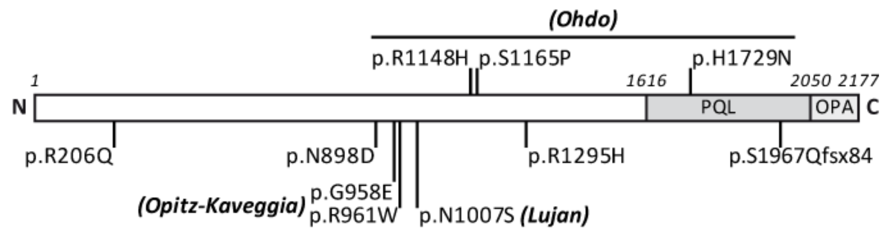



**Figure 2: Schematic overview of MED12 proteins and associated mutations.**

The different domains of MED12 are indicated: the PQL (proline-, glutamine- and leucine-rich) domain, which is involved in proteins interaction, and the OPA (glutamine-rich) domain. The identified amino acid changes leading to X-linked mental retardation syndrome are showed.

# FIGURE 3

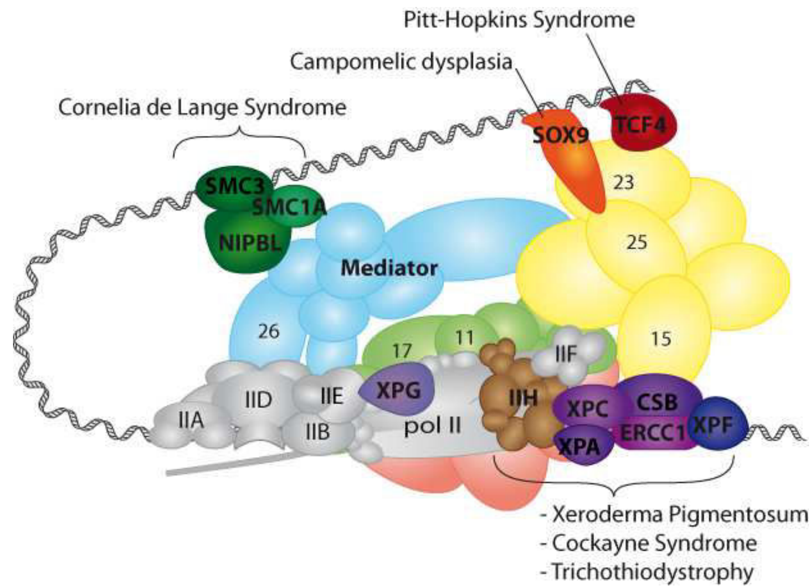



**Figure 3: Transcription factors interacting with Mediator and their associated genetic disorders**

A schematic representation of transcription machinery including Mediator complex, Pol II, the General Transcription Factor (IIA-F), NER factors, the cohesin complex and the transcription factor TCF4. Genetic disorders associated with mutations in the transcription machinery are indicated, as well as the known interaction between some MED subunits and transcription machinery.

# TABLE 1

**Table 1: Comparison of Clinical Finding in Med12-associated disorders**

| | ? | ? | FG | Lujan | Ohdo | ? | |
|---|---|---|---|---|---|---|---|
| | R206Q | N898D | G958E<br>R961W | N1007S | R1148H<br>S1165P<br>H1729N | R1295H | p.S1967Q<br>fsx84 |
| **Growth** | | | | | | | |
| Tall Stature | | | - | + | - | + | |
| Macrocephaly | + | + | + | + | - | | - |
| **Neurological** | | | | | | | |
| Intellectual disability | + | + | + | + | + | + | + |
| Agenesis of corpus callosum | | | + | + | - | | - |
| Hypotonia | | + | + | + | + | | - |
| Behavioural disturbance | + | | + | + | + | | + |
| Speech abnormalities | | | - | - | + | | + |
| **Craniofacial** | | | | | | | |
| Long narrow face | | | - | + | - | | + |
| Tall prominent forehead | | + | + | + | + | + | + |
| Triangular face | | | - | - | + | | - |
| Blepharophimosis | | | - | - | + | | - |
| Downslanting palpebrae | | | + | + | + | | - |
| Strabismus | + | | + | + | + | | - |
| Hypertelorim | | | + | - | - | | - |
| Small ears | | + | + | - | + | | - |
| Philtrum | | | | short | long | | |
| Maxillary hypoplasia | | | + | + | - | + | + |
| Micrognathia | | | + | + | + | | - |
| High narrow palate | | | + | + | + | + | - |
| Open mouth | | | + | + | + | | + |
| Dental anomalies | | | + | + | - | | - |
| **Extremity** | | | | | | | |
| Foetal finger pads | | | + | - | - | | - |
| Syndactyly | | | + | - | - | | - |
| Broad thumbs | | + | + | + | - | | - |
| Horizontal palmar crease | | | + | - | - | | - |
| Long hyperextensible digits | | | - | + | + | | - |
| **Cardiovascular** | | | | | | | |
| Congenital heart defect | Left ventricular hypertrophy | | - | - | - | | |
| **Gastrointestinal** | | | | | | | |
| Constipation | | | + | - | + | | - |
| Anal anomalies | | | + | - | - | | - |
| **Genitourinary** | | | | | | | |
| Genital anomalies | | | + | - | + | + | |
| **Others** | Extra Nipples | | | | | | |

# TABLE 2

**Table 2: Principal clinical features of transcription machinery disorders**

| | **Neurological defect** | **Development delay** | **Distinctive facial feature** | **Gastro-intestinal dysfunction** | **Genital anomalies** | **Photo-sensitivity** | **Breathing problem** |
|---|---|---|---|---|---|---|---|
| **MED12** *XLID* | + | - | + | +/- | +/- | - | - |
| **MED17** | + | - | - | - | - | - | - |
| **MED23** | + | - | - | - | - | - | - |
| **MED25** *CMT* | - | - | - | - | - | - | - |
| *CdLS* | + | + | + | + | + | - | +/- |
| **TCF4** *PTHS* | + | + | + | + | +/- | - | + |
| **SOX9** *CMD* | + | + | + | - | - | - | + |
| *XP* | +/- | - | + | - | - | + | - |
| *CS* | + | + | - | - | + | + | - |
| *TTD* | + | + | - | - | - | +/- | - |

# III. Supplemental information

***Discovery of the Mediator complex***

Several Mediator-like complexes were identified in transcription systems derived from fly, worm, mouse, rat and human cells. Each complex was named according to how it was discovered: TRAP (Thyroid hormone Receptor-Associated Protein) (Fondell et al., 1996), ARC (Activator Recruited Factor) (Näär et al., 1999), DRIP (vitamin D receptor interacting protein) (Rachez et al., 1998), CRSP (Cofactor Required for Sp1 activation) (Ryu et al., 1999), PC2 (Positive Cofactor) (Malik et al., 2000) and NAT (Negative regulator of Activated Transcription) (Sun et al., 1998). The unified nomenclature proposed by Bourbon et al. (2004) is showed in Table 3.

***Composition and structure of the Mediator complex***

The Figure 10 represent a data collection of the accurate EM map of the yeast and human obtained by Tsai et al.

***The Mediator complex: a general transcription factor***

The first definition of general transcription factors is their aptitude to induce *in vitro* basal transcription by Pol II. Mediator was first discovered based on its ability to increase basal transcription in the presence of activators, and therefore fails to satisfy this definition of a GTF. However, another definition of GTFs is their requirement for (almost) all mRNA transcription *in vivo*. A genome-wide expression analysis using temperature-sensitive mutants show that, like Pol II and TFIIH, the Mediator controls the transcription of almost all genes of *S. cerevisiae* (Holstege et al., 1998). This is satisfying the second definition of GTF.

However, contrary to other GTFs, Mediator has a highly structural flexibility and a variable subunit composition. It may be better to recognize the Mediator not as a GTFs but instead as a nearly universally required coactivator that contributes to transcriptional activation through a variety of mechanisms (Ansari and Morse, 2013).

| New Name | *S. cerevisiae* | C. *elegans* | *D. melanogaster* | *H.sapiens* | | |
|---|---|---|---|---|---|---|
| **MED1** | Med1 | MDT-1.1 (SOP-3) | Trap220 | TRAP220 | ARC/DRIP205 | CRSP200 |
| **MED2** | Med2 | MDT-1.2 | | | | |
| **MED3** | Pgd1/Hrs1/Med3 | | | | | |
| **MED4** | Med4 | MDT-4 | Trap36 | TRAP36 | ARC/DRIP36 | |
| **MED5/24** | Nut1 | | Trap100 | TRAP100 | ARC/DRIP100 | CRSP100 |
| **MED6** | Med6 | MDT-6 (LET-425) | Med6 | hMed6 | ARC/DRIP33 | |
| **MED7** | Med7 | MDT-7 (LET-49) | Med7 | hMed7 | ARC/DRIP34 | CRSP33 |
| **MED8** | Med8 | MDT-8 | Arc32 | | ARC32 | |
| **MED9** | Cse2/Med9 | | CG5134 | | | |
| **MED10** | Nut2/Med10 | MDT-10 | Nut2 | hNut2 | hMed10 | |
| **MED11** | Med11 | MDT-11 | Med21 | | | |
| **MED12** | Srb8 | MDT-12 (DPY-22/SOP-1) | Kto | TRAP230 | ARC/DRIP240 | |
| **MED12L** | | | | | | |
| **MED13** | Ssn2/Srb9 | MDT-13 (LET-19) | Skd/Pap/Bli | TRAP240 | ARC/DRIP250 | |
| **MED13L** | | | | PROSIT240 | | |
| **MED14** | Rgr1 | MDT-14 | Trap170 | TRAP170 | ARC/DRIP150 | CRSP150 |
| **MED15** | Gal11 | MDT-15 | Arc105 | PCQAP | ARC105 | |
| **MED16** | Sin4 | | Trap95 | TRAP95 | DRIP92 | |
| **MED17** | Srb4 | MDT-17 | Trap80 | TRAP80 | ARC/DRIP77 | CRSP77 |
| **MED18** | Srb5 | MDT-18 | p28/CG14802 | | | |
| **MED19** | Rox3 | MDT-19 | CG5546 | | | |
| **MED20** | Srb2 | MDT-20 | Trfp | hTRFP | | |
| **MED21** | Srb7 | MDT-21 | Trap19 | hSrb7 | hSrb7 | |
| **MED22** | Srb6 | MDT-22 | Med24 | | | |
| **MED23** | | MDT-23 | Trap150b | Trap150b | ARC/DRIP130 | CRSP130 |
| **MED25** | | | Arc92 | | ARC92 | |
| **MED26** | | | Arc70 | | ARC70 | CRSP70 |
| **MED27** | | MDT-27 | Trap37 | TRAP37 | | CRSP34 |
| **MED28** | | MDT-28 | Med23 | | | |
| **MED29** | | MDT-29 | Intersex | | | |
| **MED30** | | | Trap25 | TRAP25 | | |
| **MED31** | Soh1 | MDT-31 | Trap18 | hSoh1 | | |
| **CDK8** | Srb10/Ssn3/Ume5 | | Cdk8 | hSrb10 | CDK8 | |
| **CycC** | Srb11/Ssn8/Ume3 | CIC-1 | CycC | hSrb11 | CycC | |

**Table 3: Mediator subunits found in different organisms**
*Table adapted from Bourbon et al., 2004.*

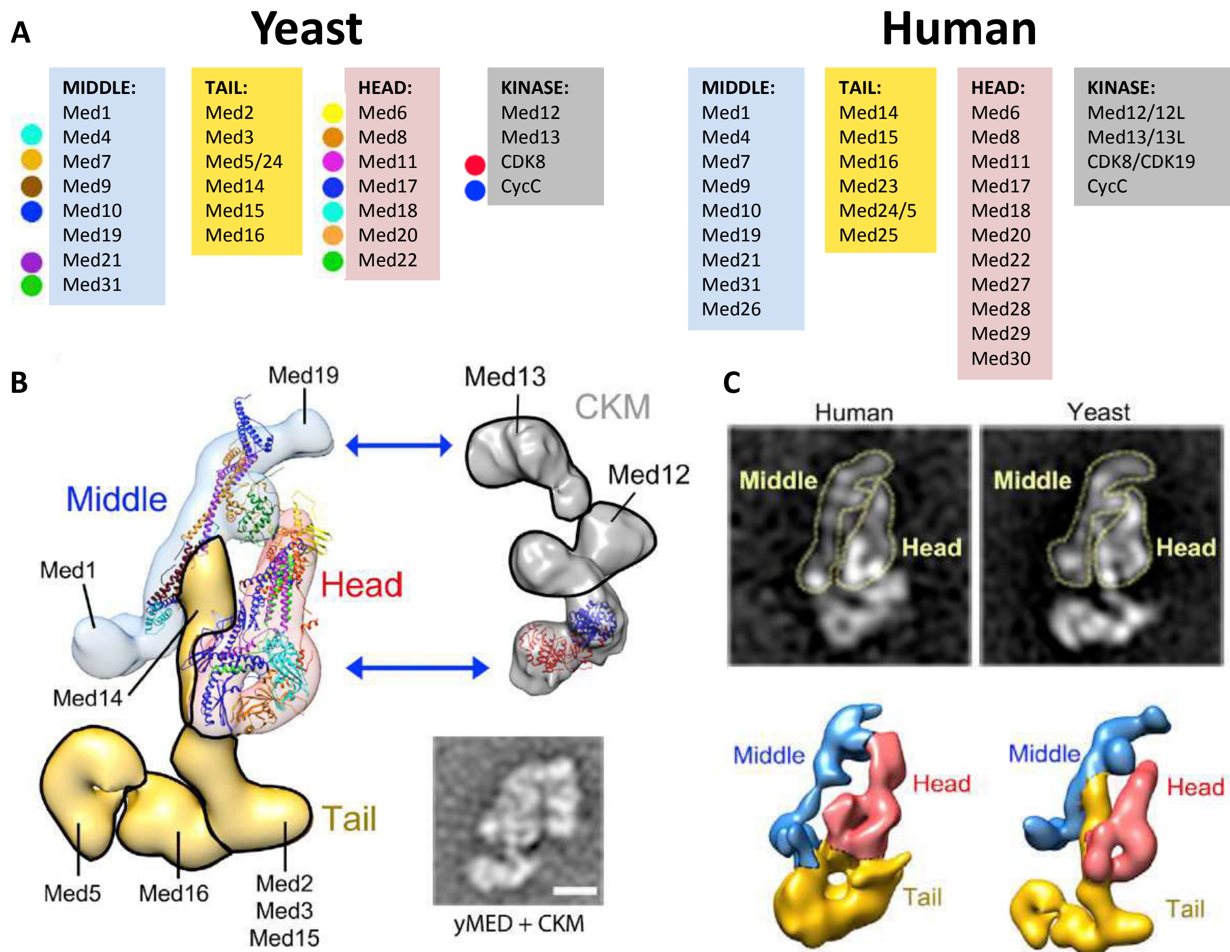



**Figure 10: Modular and subunit organization of yeast Mediator and comparison with human Mediator**
***(A)*** *Subunit organizations of yeast and human Mediators (25 and 30 different protein components, respectively).* ***(B)*** *Position and relative arrangement of all 25 yMED subunits (including those in the dissociable Kinase module). Available X-ray structures are shown docked into the yMED cryo-EM map.* ***(C)*** *Comparing the structures of yeast and human Mediator (hMED) highlights a similarity between them and can be used to tentatively identify modules and module boundaries in hMED. The overall structure and interactions of the Head, Middle, and Tail modules appear to be conserved between yeast and human Mediators.*
*Figure from Tsai et al., 2014.*

### *Cohesin complex*

Four evolutionarily conserved subunits form the core structural components of the cohesin complex: SMC1, SMC3, RAD21 and SA1/SA2. SMC1 and SMC3 belong to the structural maintenance of chromosomes (SMC) family which is conserved from yeast to human. SMC1 and SMC3 interact through their central hinge regions, while their respective paired amino- and carboxyl-terminal globular domains are further bridged by RAD21 (Figure 11). High-resolution microscopy and biochemical studies revealed that cohesin form a ring structure that topologically encircles DNA (Anderson, 2002; Haering et al., 2002; Gruber et al., 2003). Human

cohesin requires NIPBL (Nipped-B-like or Delangin) and its partner MAU2 for chromatin loading (Seitan et al., 2006; Watrin et al., 2006).

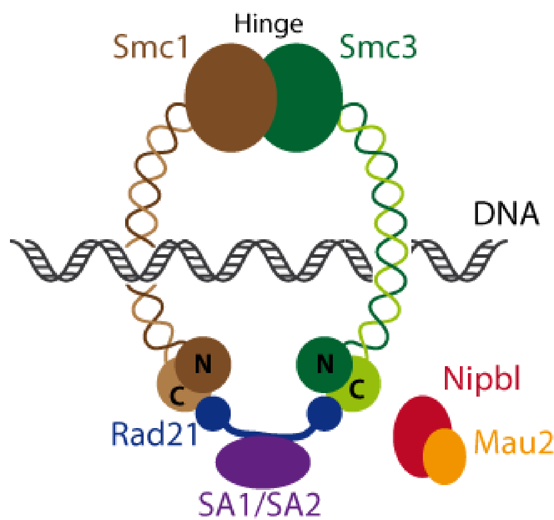



**Figure 11: The cohesin complex and its regulators**
*Ring model of cohesin complex formed by four subunits SMC1, SMC3, RAD21 and SA1/SA2. The NIPBL/MAU2 dimer loads cohesin onto DNA.*

Recently, cohesin has been demonstrated to play a critical role in gene expression. In yeast, cohesin contributes to this process by facilitating the subnuclear organization of chromatin, such as nucleolar morphology or clustering of tRNA genes (Gard et al., 2009). In *S. pombe*, it also regulates termination of transcription (Gullerova and Proudfoot, 2008).

In *Drosophila*, cohesin and its loader, Nipbl, binds preferentially transcribed regions where it colocolizes with Pol II (Misulovin et al., 2008). More recently, cohesin appears to both positively and negatively affect the transition from paused RNA polymerase to transcription elongation (Fay et al., 2011; Schaaf et al., 2013).

In mammals, experiments of chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) demonstrates that NIPBL binding site do not overlap with cohesin-binding site, but colocalize with specific transcription factors at active promoter (Zuin et al., 2014). These observations suggest a cohesin-independent role of NIPBL for transcription.

In mouse embryonic stem cells, cohesin and MED facilitate DNA looping between the enhancers and promoters of genes required to maintain pluripotency (Kagey et al., 2010). In addition, a recent paper demonstrates a crucial role for Nipbl, together with Mediator, on regulation of long-range chromosomal interaction, necessary for zebrafish and mice limb development (Muto et al., 2014). Thus, cohesin is probably a novel kind of coactivator that contributes to the three-dimensional organization of active genes.

Cornelia de Lange Syndrome (CdLS)

The prevalence of CdLS has been estimated between 1:50 000 and 1:100 000 births (Barisic et al., 2008). Almost all cases are sporadic and dominant. Mutations in NIPBL, HDAC8, SMC1A and SMC3 may explain approximately 65% of CdLS patients but the cause of the remaining 35% remains unclear.

Haploinsufficient NIPBL mutations (protein truncating mutations, i.e. frameshift or nonsense mutations) often exhibit more severe phenotype compared to missense mutations (Gillis et al., 2004). For SMC1 or SMC3, no truncating mutations were identified and patients often show mental retardation, with other abnormalities being fewer and/or milder. Recently, mutations in RAD21 were found in patient with a CdLS-like disorder (Deardorff et al., 2012). Unlike patient with mutations in NIPBL, SMC1, or SMC3, these individuals have much milder cognitive impairment.

Nipbl heterozygous mutant (Nipbl$^{+/-}$) mice which recapitulate several features of CdLS have a significant transcriptional dysregulation of many genes (Kawauchi et al., 2009). Surprisingly, Nipbl transcript levels in CdLs mice are decreased of only 25-30%, suggesting compensatory upregulation of the intact allele. Consistent with this, a 15% decrease in NIPBL expression was observed in a mild form of CdLS, whereas a more important decrease was observed in severe forms of CdLS (Borck et al., 2006; Liu et al., 2009). These observations indicate the importance of Nipbl dosage for proper development.

# MATERIALS
# &
# METHODS

# MATERIALS & METHODS

### *Cells culture*

Human primary fibroblasts isolated from patients with MED12 mutation were grown in Dulbecco's modified Eagle's medium (DMEM) and Ham's F-10 Nutrient Mix, supplemented with 12% fetal calf serum (FCS) and penicillin-streptomycin. The patient was followed at CHU Clermont Ferrant by Dr Christine FRANCANNET. Human primary fibroblasts isolated from patients with NIPBL mutation were grown in DMEM (1g/l Glucose) supplemented with 10% FCS and gentamicine. The patient was followed at Hôpital Necker Enfants Malades, Paris by Dr Valérie CORMIER-DAIRE.

Lymphoblastoïdes cells were generated by Epstein-Barr Virus (EBV) transformation of the peripheral blood lymphocytes from patients. Cells were provided by Dr Charles SCHWARTZ from Greenwood Genetic Center, USA, excepted cells with Ohdo syndrome provided by Dr Hans BRUNNER and Dr Arjan BROUWER from Nijmegen, the Netherlands. Cells were grown in RPMI without Hepes supplemented with 15% FCS and gentamicine.

All the cells were maintained at 37°C in 5% $CO_2$ environment.

| | Mutation | Cell |
|---|---|---|
| **Lymphoblastoid cells** | WT | 2184 |
| | MED12/R206Q | cms2458 |
| | MED12/N898D | cms13404 |
| | MED12/R961W | cms14176 |
| | MED12/N1007S | 7439 |
| | MED12/R1148H | MED12 R1148H son |
| | Mother of patient MED12/R1148 | MED12 R1148H mother |
| | MED12/S1165P | MED12 S1165P |
| | MED17/L371P | MED17 L371P |
| **Fibroblasts** | WT | FB789 |
| | MED12/R1295H | MED12 R1295H (2) |
| | Mother of patient MED12/R1295H | MED12 / R1295H mère |
| | NIPBL | CdL - NIPBL1 |
| **MEFs** | WT | MEF WT2 (CLT M17L371P) |
| | Med17/L371P | MEF M17L371P |

***Isolation of mouse embryonic fibroblasts***

Embryos around 11 days of gestation were dissected from the uterus and separated from its placenta and surrounding membranes. The head were removed and used for genotyping. Dark red organs were also cut away and the remainder embryo was carefully washed with PBS and then transferred in 1ml of MEFs medium (DMEM containing 4,5g/l glucose supplemented with GLUTAMAX-I, 10% FCS and gentamicine). The embryo carcass was pipette up and down several times with a syringe fitted with an 18G needle to break up tissue chucks and get cells into suspension. Tissues debris was removed using a 40µm cell strainer and then uniform single-cell suspension was plated in 35mm dishes.

***Treatments***

For treatment by all-trans retinoic acid (tRA; Biomol) or 1α,25-Dihydroxyvitamin D3 (VitD; Sigma-Aldrich), cells were incubated in red phenol-free medium containing charcoal treated FCS and antibotics during 12 hr prior to the treatment. The induction start with replacement of this medium by the same medium containing 10µM of tRA or 100nM of VitD

For serum treatment, cells were incubated in medium without serum and red phenol for 24h. Then, cells were treated by addition of serum (20% final concentration) directly into the medium.

For RNA experiment, we used around $15.10^4$ fibroblasts per 35-mm dishes and, for ChIP experiment, we used around $250.10^4$ fibroblasts per 150-mm dish.

For UV irradiation, lymphoblastoid cells were quickly rinsed with PBS. For each point, 4 millions of cells were resuspended in 2 ml of PBS and spread on a 10-cm dish, forming a small layer. Cells were then exposed to UV irradiation with a Philips TUV lamp (predominantly 254 nm) at a dose of 20 $J/m^2$. Subsequently, the medium was added back, and cells were returned to culture conditions for definite times.

***RNA extraction, reverse transcription and real-time qPCR***

Total RNA was extracted with RNAeasy mini kit (QIAgen) and reverse transcribed with poly-dT or hexamer 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 provided in the table below. mRNA levels represent the ratio between

values obtained from treated cells compared to untreated cells after normalization against the housekeeping GAPDH or 18S mRNA.

***Chromatin Immunoprecipitation (ChIP)***

After treatment, cells were subjected to crosslinking at room temperature for 15 min with 1% formaldehyde, followed by addition of glycine to terminate crosslinking reactions. Cells were lysed in shearing buffer (50 mM Tris-HCl pH8, 10mM EDTA, 1% SDS) with PIC at 4°C for 15 min. Nucleus were pelleted and resuspended in sonication buffer (50mM Hepes pH7.6, 140mM NaCl, 1mM EDTA, 1% TritonX, 0.1% Na-deoxycholate, 0.1% SDS) with PIC. Chromatin was sheared using 800R sonicator (Qsonica). Samples were incubated with corresponfing antibodies at 4°C overnight. Immune complexes were precipitated with a pre-blocked mix of protein G- and A-Sepharose beads (GE Healthcare Life Sciences) for 3h at 4°C. Bound complexes were sequentially washed with sonication buffer, high salt buffer (50mM Hepes pH7.6, 500mM NaCl, 1mM EDTA, 1% TritonX, 0.1% Na-deoxycholate, 0.1% SDS) and TE buffer (10mM Tris, 1mM EDTA). Protein-DNA complexes were recovered in elution buffer (100mM $NaHCO_3$, 1%SDS) and the cross-linking was reversed by overnight incubation at 65°C. DNA fragments were treated with proteinase K for 2h at 42°C and purified using PCR purification kit (QIAGEN). Quantitative PCR was performed as described above using sets of primers provided in the table below. All the results are presented as 'fold recruitment' and represent the ratio of input percentage between treated and non-treated cells.

***Western blot and co-immunoprecipitation***

Patients' cells were harvest in RIPA buffer (20mM Tris HCl pH7.5, 120mM NaCl, 1mM EDTA, 1% NP-40, 0.1% SDS, 0.5% Na-deoxycholate and PIC). Lysate were separated by SDS-PAGE, transferred to nitrocellulose membrane and revealed by immunoblotting using the ECL Western blotting detection system (GE Healthcare).

For co-immunoprecipitation, cells were harvest in lysis buffer (50mM Tris HCl pH7.5, 1mM EDTA, 1% Triton and PIC). 10µl of protein G magnetic bead (dynabead, invitrogen) were used par IP. 2µg of antibodies were bound to the beads in PBS with BSA (5mg/ml) during 2h at 4°C with rotation. 200µg of whole cell extract were then incubated with beads-antibodies complex for 2h at 4°C with rotation. After 2 washes at 100mM salt, 2 washes at 300mM and 1

wash at 100mM, beads were boiled in Laemmli buffer and proteins were separated by SDS-PAGE, transferred to nitrocellulose membrane and probed with antibodies.

***Antibodies***

For Western blot, the following antibodies were used: MED6 (santacruz, sc-9434, 1:500), MED12 (Béthyl, A300-774A, 1:2000), MED16 (abcam, AB28520, 1:5000), MED17 (santacruz, sc-48777, 1:500), MED22 (santacruz, sc-393738, 1:500), MED23 (BD Pharmingen, 550429, 1:2000), CDK8 (abcam, ab64940, 1:1000), Cyclin C (santacruz, sc-1061, 1:500), β-tubulin (millipore, MAB3408, 1:5000) and TBP (IGBMC, 3TF1-3G3, 1:10000).

For immunoprecipitation, the following antibodies were used: monoclonal antibodies against RNA polymerase II (IGBMC, 1BP 7C2), RARα (IGBMC, 9α9A6), ELK3 (IGBMC, 5NE 2F3A2), and polyclonal antibodies against BSA (santacruz, sc50528), TCF4 (santacruz, sc13027), ELK1-P (santacruz, sc8406), TFIIB (santacruz, sc-225), MED1 (santacruz, sc-8998), MED6 (santacruz, sc-9434), MED12 (Béthyl, A300-774A), NIPBL (Béthyl, A301-779A), SMC1 (Béthyl, A300-055A), H3K9ac (cell signaling, #9671) and $H3K9me^2$ (cell signaling, #9753).

***Primers***

| HUMAN | | |
|---|---|---|
| **mRNA primers** | **Forward** | **Reverse** |
| GAPDH | TCGACAGTCAGCCGCATCTTCTTT | ACCAAATCCGTTGACTCCGACCTT |
| RARβ | CCAGCAAGCCTCACATGTTTCCAA | TACACGCTCTGCACCTTTAGCACT |
| TMG2 | GCCACTTCATTTTGCTCTTCAA | TCCTCTTCCGAGTCCAGGTACA |
| RARα | GAAGATTACTGACCTGCGAAGC | CCCTCTGAGTTCTCCAACATTTC |
| PDK4 | ACCCAAGCCACATTGGAAGCA | AACTGTTGCCCGCATTGCATT |
| SMAD3 | TTGTCCAGTCTCCCAACTGTAAC | GTCAACTGGTAGACAGCCTCAAA |
| NRIP1 | GTGGAACAAAGGTCATGAGTGA | CTCGAGAATACTGCTGCAAATG |
| MED12 | GCAGAAGAGCATGTCCCTATT | TGGCTGTAGAGGGAGGTAAG |
| MED17 | AGTCCAGTGAAGGGCTTCTGGAAA | CGGCTTGCTAAGCTGTCAATGGTT |
| MED23 | AATGCGCTATGAATGCACGA | GTTTGGAAAGGGACCAGGAGA |
| CDK8 | GGGATCTCTATGTCGGCATGTAG | AAATGACGTTTGGATGCTTAAGC |
| NIPBL | GAACTACAGTTGTGTGCCATTAAG | TCCTCTTGCTGGTTGGTAATC |
| SMC1A | TGCCTTGGATAACACCAACA | CAGTCCCCTTGCTCAGGATA |
| SMC3 | TGGCACGATCAGAAGATTTGGA | TCAATAGCATGCCTTGCCGA |
| RAD21 | GGAAAGGAGGAGAGGCAGATAA | GCGGCTTGGCTCTTCAATAA |
| HDAC8 | CAAACGGGCCAGTATGGT | ATGGAGGCCACTTTAGGCTT |
| OSTEOPONTIN | TATGATGGCCGAGGTGATAG | AGGTGATGTCCTCGTCTGTA |

| | | |
|---|---|---|
| CYP24 | TCTTGACAAGGCAACAGTTC | AAGCCAACGTTCAGGTCTAA |
| EGR1 | AGCACCTGACCGCAGAGTCTTT | CACCAGCACCTTCTCGTTGTT |
| FOS | CAAGCGGAGACAGACCAACT | AGTCAGATCAAGGGAAGCCA |
| JUN | AGCGCCTGATAATCCAGTCC | CTGCTCATCTGTCACGTTCTTG |
| GADD45 | TCAACGTCGACCCCGATAAC | TCGGTCTCCAAGAGCAGGAG |

| **ChIP primers** | **Forward** | **Reverse** |
|---|---|---|
| JUN proximal | CCAGAGAAGAATCTTCTAGG | CCCCAAGGCCTTCCCATTGG |
| JUN distal | CCGTCTCACTCTCTTGCTCTTC | CAACTGGACAAAATGGCTCTG |
| FOS promoter | GAGCAGTTCCCGTCAATCC | GCATTTCGCAGTTCCTGTCT |
| EGR1 promoter | CTGCCATATTAGGGCTTCCTGCTT | TATTTGAAGGGTCTGGAACGGCAC |
| NIPBL promoter | GGGTGGTTGTTAGTGTTTGG | TCTCTCTCGTTCCGTCTCT |
| RARβ promoter | TGGTGATGTCAGACTAGTTGGGTC | GCTCACTTCCTACTACTTCTGTCAC |

| | **MICE** | |
|---|---|---|
| **mRNA primers** | **Forward** | **Reverse** |
| 18S | TCAACTTTCGATGGTAGTCGCCGT | TCCTTGGATGTGGTAGCCGTTTCT |
| Med1 | TGGAGGGCATCAGCATTTGG | GCCCAGTCCATTCTGTCTGG |
| Med6 | GACAGCGTGTGGATGCTTTAC | GCTTTTCTCCAGACTTTTGCTGA |
| Med12 | GTTGGAATCCGGTCCTCCTG | TAGCTCCGCATCTCCGAGTA |
| Med17 | GCGAAGTGCCCTTACAGAGA | GAGATCAGCTGCAGCGTTTG |
| Med22 | CGAGATCATCAAGACCGCCA | CTCACCAGCTCGAACGATGT |
| Med23 | CCGCAGACTGCTTTGTTGAG | GCTGCTTGTGCTGCTTATTT |
| Egr1 | CGGCTGCCTCTTCACTCTCT | GCAGGAGATGGGTAGGTGGA |
| Fos | CAGCTATCTCCTGAAGAGGAAG | CTTCTCATCTTCAAGTTGAT |
| Jun | CCTTCTACGACGATGCCCTC | GGTTCAAGGTCATGCTCTGTTT |
| Cyp24 | GACCCCTCGTGGCTTTAGAC | GCTGCAAGGTGCAGTTGTTT |
| Osteopontin | CTGGCAGCTCAGAGGAGAAG | TTCTGTGGCGCAAGGAGATT |

***Sequences alignment***

Sequences alignment was performed with BioEdit software (open source). The following sequence has been used:

| | **MED12** | **MED17** | **MED23** |
|---|---|---|---|
| *Homo sapiens* | NP_005111.2 | NP_004259.3 | NP_004821.2 |
| *Mus musculus* | NP_067496.2 | NP_659182.1 | NP_081623.3 |
| *Rattus norvegicus* | NP_001180221.1 | NP_001100271.1 | NP_001263983.1 |
| *Bos Taurus* | XP_005228076.1 | NP_001029902.2 | NP_001192691.1 |
| *Pongo abelii* | NP_001124553.2 | XP_002822403.1 | XP_009240531.1 |
| *Pan troglodytes* | NP_001009019.1 | XP_009422274.1 | XP_009450287.1 |
| *Danio rerio* | NP_001034550.1 | NP_001071042.1 | NP_001003990.1 |
| *Xenopus tropicalis* | XP_002934949.2 | NP_001016974.1 | XP_002936354.2 |

***Plasmids:***

MED12 expression plasmid was obtained using the Gateway Invitrogen cloning method. Wild-type 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 (NEB) 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, 48h before the experiments.

# RESULTS

*Draft*

# MED12 PROJECT:
# Different X-linked mental retardation disorders depending on the position of MED12 mutation.

Lise-Marie DONNIO*, Satoru Haschimoto*, Baptiste Bidon, Colm Ryan, Christine Francannet, Charles Schwartz et Jean Marc Egly.

*** Co first authors**

IGBMC, Department of Functional Genomics and Cancer biology, CNRS/INSERM/Université de Strasbourg, BP 163, 67404 Illkirch Cedex, C.U. Strasbourg, FRANCE

**Key words:** MED12, Mediator complex, gene expression, Opitz-Kaveggia syndrome, Lujan syndrome, Ohdo syndrome, X-linked intellectual disability

## 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 (Conaway et al., 2005; Kornberg, 2005; Malik and Roeder, 2010; Poss et al., 2013). Mediator is composed of more than 30 subunits, arranged in four different modules named Head, Middle, Tail and Kinase (Figure 1A). To convey essential information from transcription factors bound at DNA responsive elements to the basal transcription machineries, MED physically interact with a collection of transcriptional regulatory proteins, including RNA polymerase II (Pol II) (Whyte et al., 2013; Zhu et al., 2011).

Dysfunction of transcriptional machinery has been shown to elicit broad effects on cell states (proliferation or differentiation) giving rise to diverse pathologies, including cancers (Lee and Young, 2013). Mutations in MED subunits are associated with a wide range of genetic disorders, such as Infantile cerebral and cerebellar atrophy (MED17; (Kaufmann et al., 2010), non-syndromic mental retardation (MED23, (Hashimoto et al., 2011), Charcot-Marie Tooth disease (MED25, (Leal et al., 2009); most of them exhibiting neurological defects. We here focused on MED12 in which mutations are linked with a broad spectrum of genetic disorders with X-linked intellectual disability (Graham and Schwartz, 2013), such as Opitz-Kaveggia syndrome (p.R961W, (Risheg et al., 2007) and p.N898D, (Rump et al., 2011)), Lujan syndrome (p.N1007S, (Schwartz et al., 2007)), Ohdo syndrome (p.S1165P, p.R1248H and H1729N, (Vulto-van Silfhout et al., 2013)) or non-syndromic profound X-linked intellectual disability (p.S1967Qfsx84; (Lesca et al., 2013). However, the underlying basis by which different genetic disruption of MED12 elicits separate and phenotypically distinct syndromes remains unclear.

MED12 is located at Xq13.1, and belongs together with MED13, CDK8 and Cyclin C (CyC) to the Kinase module, that exists in variable association with Mediator. MED12 is required for stable incorporation of CDK8/CyC into Mediator and appear to activate the kinase activity of the CDK8 module (Kim et al., 2006; Knuesel et al., 2009). Depending on the context, the Kinase module can regulate negatively or positively the transcription (Nemet et al., 2014). For example, it functions as a positive regulator at specific p53-regulated genes such as *p21* (Donner et al., 2007). In contrast, Cdk8, as part of the kinase module, phosphorylates Pol II leading to disruption of Mediator–Pol II interaction and transcription inhibition (Hengartner et al., 1998). Similarly, the kinase phosphorylates Cyclin H, a subunit of the transcription/DNA repair factor TFIIH, and thus represses the ability of TFIIH to activate transcription and its CTD kinase activity (Akoulitchev et al., 2000). Moreover, MED12 has been linked biochemically and

genetically with the Notch, Wnt, and Sonic hedgehog signalling pathways that control key aspects of brain development and function, from initial patterning to neuronal plasticity (Moghal, 2003; Rau et al., 2006; Treisman, 2001; Yoda, 2005; Zhou et al., 2006).

We report three new *MED12* missense mutations (p.R206Q, p.N898D and p.R1295H) which cause clinical features that cannot be ascribed to previously described MED12 syndromes. To provide explanations to the pathogenicity of MED12-related disorders, we have investigated the effect of mutations (p.R206Q, p.N898D, p.R961W, p.N1007S, p.R1148H, p.S1165P and p.R1295H) found in *MED12* on the transcriptional activation of some given genes. We found that gene expression in patient's cell lines varies depending on the position of the mutation and the way they are activated. For example, the expression of *JUN*, *FOS* and *EGR1* immediate early genes (IEGs) is impaired after serum addition to serum-starved patient cells compared to control cells. Chromatin immunoprecipitation experiments next showed that the recruitment of some transcription machinery components (including DNA binding transcription factors) at the promoter of activated genes is altered. We also observed that the transition from a closed to an open chromatin state is disturbed.

This study sheds light on how different mutations in *MED12* gene causes distinct expression of activated genes giving rise to different disorders that share some overlapping clinical features.

## RESULTS

### *Discovery of new mutations in MED12 gene*

In addition to the already identified MED12 mutations associated with Opitz-Kaveggia (p.R961W), Lujan (p.N1007S) and Ohdo (p.R1148H and p.S1165P) syndromes, we ascertain three additional mutations, recently discovered (Figure 1B). Genetic analyses identified these p.R206Q, p.N898D and p.R1295H mutations within *MED12* (Figure 1C). Aside, the R206, N898, R961, N1007, R1148, S1165 and R1295 residues are absolutely conserved across all MED12 orthologs, from *Xenopus* to human (Supplemental Figure 1). These patients were not diagnosed *a priori* as having MED12-related syndrome although they share some common clinical features (Table 1).

Here half page of patients description will be filled by Dr C. Schwartz. The differences suggesting that they cannot all of them be classify as Opitz and/or Lujan and/or Ohdo syndromes will be underlined. Some pictures of the patients would also be great if ethically possible. The Table 1 (adapted from (Graham and Schwartz, 2013)) is a draft that should be clearer with the clinician appreciation.

Using EBV-immortalized lymphoblastoid cells from patients, we showed, by reverse transcription-quantitative polymerase chain reaction (RT-qPCR) of total RNA, that the mutations MED12/R206Q, /N898D, /R961W, /N1007S and /R1295H do not affect mRNA expression level of *MED12*, as well as *CDK8*, *MED17* and *MED23* (Figure 1D). Western Blot analysis of whole cell extract reveal that the quantity of MED12 protein, as well as Mediator subunits belonging either to the Kinase, Head or Tail modules, is not significantly different in the five mutated and control (WT) cells (Figure 1E, left panel). Immunoprecipitation (IP) experiments using antibodies directed towards MED6 or MED12 reveal that MED12 mutations do not change the overall composition of the Mediator when compared to WT cells (Figure 1E).

### *MED12 mutated cells respond differently to external stimuli*

Considering certain symptoms of patients, we investigated the effect of MED12 mutations on the expression of some nuclear receptors dependent genes, which play essential roles in development, differentiation, and metabolism (Belakavadi and Fondell, 2006). We first considered *CYP24* gene that is under the control of Vitamin D (VitD) and has a role in maintenance of calcium homeostasis and skeleton architecture. We observed a similar expression of *CYP24* gene in WT, MED12/R206Q, /N1007S and /R1295H cells, whereas in

MED12/N898D and /R961W cells, the gene is up-regulated (Figure 2A, lane 3 and 4). Next we investigated *RARβ* gene induced by trans retinoic acid (tRA), the biologically active form of vitamin A that mediates cellular signalling in embryonic morphogenesis, cell growth and differentiation. Every tRA treated lymphoblastoid cell lines, except MED12/N1007S, accumulate *RARβ* mRNA (Figure 2B lane 5). In this cells line, we repeatedly observed no *RARβ* activation over a 24h induction period (data not shown). We also considered the effect of MED12 mutation on expression of genes induced by environmental stress, such UV irradiation. We found that the expression of *GADD45* is down-regulated in both MED12/R206Q and /N898D cells compared to WT, MED12/R961W, /N1007S and /R1295H cells (Figure 2C, lane 2 and 3).

All together, the above data underline how each MED12 mutation specifically disrupts the expression of certain genes in response to different induction.

***MED12 mutations impair PIC formation***

Mediator being known to mediate the response of Immediate Early Genes (IEGs) to serum mitogens stress (Hashimoto et al., 2011; Stevens, 2002), we next measured the expression of three IEGs, *JUN*, *FOS* and *EGR1*, after serum addition to serum-starved control and patients cells. We observed a dysregulation of those genes that depend firstly on MED12 mutation and secondly on the activated gene (Figure 2 D1, E1 and F1). Indeed, expression of *JUN* is down-regulated in MED12/R206Q, /R961W and /N1007S cells compared to WT, MED12/N898D and /R1295H cells (Figure 2 D1, lanes 2, 4 and 5). In contrast, we observed a similar induction of *FOS* gene after serum addition in WT, MED12/R206Q and /N898D cells, whereas in MED12/R961W, /N1007S and /R1295H cells, *FOS* is up-regulated (Figure 2 E1, lanes 4, 5 and 6). In addition, *EGR1* expression is down-regulated in both MED12/R206Q and /N898D cells compared to WT, MED12/R961W, N1007S and /R1295H cells (Figure 2 F1, lanes 2 and 3)

To determine whether the dysregulation of IEGs expression after serum mitogens in MED12 deficient cells resulted from defective preinitiation complex (PIC) formation (the first step of RNA synthesis), we next monitored the recruitment of the transcriptional machinery to their promoters by using Chromatin Immunoprecipitation (ChIP) coupled to qPCR. We observed a parallel between the expression of *JUN, FOS* and *EGR1* and the recruitment of Pol II and MED12 at their respective promoter. In MED12/R206Q, /R961W and /N1007S cells, the down-regulation of *JUN* correlates with a defective recruitment of Pol II and MED12 (Figure 2, compare panel D1, with panels D2 and D3, lanes 2, 4 and 5). In WT, MED12/R961W, /N1007S

and /R1295H cells, the expression of *FOS* parallels the recruitment of Pol II and MED12 (compare panels E1 with panel E2 and E3, lanes 4-6). *EGR1* down-regulation properly matches with a failure of Pol II and MED12 recruitment in MED12/R206Q and /N898D cells (compare panels F1 with panel F2 and F3, lanes 2 and 3).

We next investigated the effect of the MED12 mutation on the recruitment of DNA binding transcription factors involved in *JUN*, *FOS* and *EGR1* regulation. *JUN* is regulated by several responsive elements, including one that could be targeted by TCF4 (Hazzalin and Mahadevan, 2002; Nateri et al., 2005), a factor mutated in Pitt-Hopkins disorder (Peippo and Ignatius, 2012). We observed that the recruitment of TCF4 at its distal element as well as at the promoter, is impaired when Pol II and MED12 binding is defective (Figure 2 D4 and D5, lane 2, 4 and 5).

*FOS* and *EGR1* expression is regulated by cooperative binding of ELK1, ELK3 (SPA1) and ELK4 (NET) to serum response elements (SREs) (Figure 2, upper schemes) (Buchwalter et al., 2004). In MED12/R961W, /N1007S and /R1295H cells, the up-regulation of FOS parallels a reduced binding of the phosphorylated form of ELK1 (ELK1-P) that is compensated by an increase binding of its related paralog ELK3 on the SRE element (Figure 2, panels E4 and E5, lanes 4, 5 and 6). On the contrary, in MED12/R206Q cells, we observed a higher binding of ELK1-P and a lower binding of ELK3 at *FOS* promoter when compared to WT cells (Panels E4, lane 2). In MED12/N898D cells, neither ELK1-P nor ELK3 are detected at SRE (Panels E4 and E5, lane 3). However, in these two cells lines, *FOS* seems to be normally expressed (Panels E1, lanes 2 and 3). In both MED12/R206Q and /N898D cells in which EGR1 is down regulated neither ELK1-P nor ELK3 are recruited at its promoter (Panels F4 and F5, lanes 2 and 3), while in the three other MED12 deficient cells as well as in WT cells, we observed ELK1-P recruitment (lanes 4-6).

Knowing that each gene can be under the control of different stimuli, we next investigated the behavior of *JUN* and *FOS* when the five MED12 deficient cells were submitted to UV irradiation. In such case, contrary to what was observed above (Figure 2 D1), *JUN* is not down-regulated but even seems to be up-regulated in some of the MED12 cells lines. Moreover, over-expression of FOS after UV exposure is observed in different MED12 mutated cells than after serum treatment (Supplemental Figure 2).

Taken together, our results (performed with lymphoblastoid cells) demonstrate how each MED12 mutation differently disturbs the expression of IEGs by altering the formation of

the pre-initiation complex at their respective promoters. Moreover our data showed that the defect for a given gene depends also on the nature of the activation.

***MED12 mutation also impaired gene expression in heterozygous female***

We further focus on another MED12 patient and his heterozygous mother from whom it was possible to obtain fibroblasts, much easier to handle than lymphoblastoid cells. Patients description, comparing with other patients described on Table 1, will be filled be Dr C. Francannet.

Patient (R1295Hson) and his mother carried a G-to-A transition at nucleotide position c.3884 resulting in Arginine to Histidine amino acid modification. Sanger sequencing revealed that ~80% of mother fibroblasts express the mutant allele (Figure 3A). We proceeded as before and found that upon serum induction, *JUN* activation is deficient in both R1295Hson and mother cells while *FOS* and *EGR1* are expressed similarly to control (WT) cells (Figure 3 B1, C1 and D1).

ChIP experiments next show that the defective recruitment of Pol II, the general transcription factor TFIIB, MED12, as well as TCF4 at *JUN* promoter correlates with the absence of *JUN* induction (Panels B2-B5). In contrast, the high recruitment of Pol II, TFIIB, MED12 and ELK1 parallels the high level of *FOS* and *EGR1* expression (Panels C2-C5 and D2-D5).

Gene activation is accompanied by important chromatin remodeling resulting from histone modifications. Euchromatin, which allows transcription, is characterized by acetylation of H3K9 (H3K9ac) while heterochromatin, which inhibits RNA synthesis, is characterized by a different set of chromatin marks such as dimethylation of H3K9 (H3K9me$^2$) (Bannister and Kouzarides, 2011; Li et al., 2007). We here observed around *FOS* and *EGR1* promoter an increase of H3K9 acetylation concomitantly to a decrease of H3K9 dimethylation in response to serum (Figure 3 C6-D6 and C7-D7). In contrast, *JUN* promoter is in a heterochromatin state in both son and mother cells compared to WT cells, as shown by the absence of H3K9 acetylation (Panel B6).

Following treatment by tRA treatment, we observed a strong over-expression of *RAR$\beta$* and *PDK4* (to a lower extend) in R1295Hson and mother cells compared to control (Figure 3 E and F). Apart of this, in absence of any treatment, we also found an important difference in basal expression of *RAR$\beta$* in R1295Hson and mother cells compared to WT cells, while *PDK4*

basal expression is similar between the three cell lines (Supplemental Figure 3A and B). *CYP24* response to VitD treatment is altered in son fibroblasts while its expression in WT and mother cells occurs similarly (Figure 3G). However, *OSTEOPONTIN* gene is similarly activated in the three cell lines (Figure 3H).

The above data (performed on fibroblasts) does not allow to discriminate between the mother and the son, both exhibit strong dysregulation in the expression of some genes compared to control cells.

## DISCUSSION

Our work revealed that each MED12 mutations disrupt differently the expression of activated genes, thus giving rise to different disorders (Table 2). EM studies show that Mediator complex undergoes distinct structural shift depending on the activator (Meyer et al., 2010; Taatjes et al., 2002). We could speculate that each mutation compromise differently the MED structure, without modifying its composition, and consequently the interaction of Mediator with some transcriptional factors could be altered.

We shown that MED12 mutations impaired the response of *JUN, FOS* and *EGR1* immediate early genes (IEGs) to serum mitogens by altering the transactivation complex formation, as well as chromatin remodeling. It is well known that IEGs expression affects brain development and plasticity (Pérez-Cadahía et al., 2011). We previously proposed that altered IEGs expression might provide a molecular signature for cognitive deficits (Hashimoto et al., 2011). Our results seem to confirm this hypothesis as all MED12 patients have intellectual disability and dysregulated IEGs expression in their cells.

MED12 is located on chromosome X and thus in females one of its two alleles is inactivated. Moreover, X-inactivation occurs randomly during early embryo development on a cellular level, resulting in a mosaic expression, in which patches of cells have a normal allele of MED12, while other patches have a mutated allele. We have observed that most of heterozygous fibroblasts from a patient's mother express the mutated allele. Consequently, both mother and son fibroblasts exhibit dysregulation of some genes expression.

At this date, eight missense mutations have been described within MED12 gene in patients with cognitive deficit and dysmorphic features, but it is likely that more mutations in this gene will be detected in sporadic patients with similar clinical features.

## MATERIALS AND METHODS

***Subjects:****(to be filled by the doctors of patients)*

Physical examination

Genotyping analysis

Mutation screening of the MED12 gene

Informed consent was obtained from each patients ????

***Cell culture:***

Human primary fibroblasts were isolated from patients and grown in Dulbecco's modified Eagle's medium and Ham's F-10 Nutrient Mix, supplemented with 12% fetal calf serum (FCS) and penicillin-streptomycin. Lymphoblastoïdes cells were generated by Epstein-Barr Virus (EBV) transformation of the peripheral blood lymphocytes from patients. They were grown in RPMI without Hepes supplemented with 15% FCS and gentamycin. All the cells were maintained at 37°C in 5% $CO_2$ environment.

***Sequences alignment***

Sequences alignment was performed with BioEdit software (open source). The following sequence 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).

***Coimmunoprecipitation assays.***

Lymphoblastoid cells were harvest in lysis buffer (50mM Tris HCl pH7.5, 150 mM NaCl, 1mM EDTA, 1% Triton) with protease inhibitor Cocktail (cOmplete, ROCHE). 200µg of whole cell extract were incubated with protein G magnetic bead (dynabead, invitrogen) and 2µg of corresponding antibodies (MED6, MED12 or BSA). After washes at 300mM salt, beads were boiled in Laemmli buffer and proteins were separated by SDS-PAGE, transferred to nitrocellulose membrane and probed with antibodies.

***Treatments:***

For treatment by all-trans retinoic acid (tRA; Biomol) or 1α,25-Dihydroxyvitamin D3 (VitD; Sigma-Aldrich), cells were incubated in red phenol-free medium containing charcoal treated FCS and antibotics during 12 hr prior to the treatment. The induction start with replacement of this medium by the same medium containing 10µM of tRA or 100nM of VitD

For serum treatment, cells were incubated in medium without serum and red phenol for 24h. Then, cells were treated by addition of serum (20% final concentration) directly into the medium.

For UV irradiation, lymphoblastoid cells were quickly rinsed with PBS. For each point, 4 millions of cells were resuspended in 2 ml of PBS and spread on a 10-cm dish, forming a small layer. Cells were then exposed to UV irradiation with a Philips TUV lamp (predominantly 254

nm) at a dose of 20 J/m$^2$. Subsequently, the medium was added back, and cells were returned to culture conditions for definite times.

***RNA extraction, reverse transcription and real-time qPCR:***

Total RNA was extracted with RNAeasy 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). All the primers are available upon request. mRNA levels represent the ratio between values obtained from treated cells compared to untreated cells normalized against the housekeeping GAPDH mRNA.

***Chromatin Immunoprecipitation (ChIP):***

After treatment, cells were subjected to crosslinking at room temperature for 15 min with 1% formaldehyde, followed by addition of glycine to terminate crosslinking reactions. Cells were lysed in shearing buffer (50 mM Tris-HCl pH8, 10mM EDTA, 1% SDS) with PIC at 4°C for 15 min. Nucleus were pelleted and resuspended in sonication buffer (50mM Hepes pH7.6, 140mM NaCl, 1mM EDTA, 1% TritonX, 0.1% Na-deoxycholate, 0.1% SDS) with PIC. Chromatin was sheared using 800R sonicator (Qsonica). Samples were incubated with corresponfing antibodies at 4°C overnight. Immune complexes were precipitated with a pre-blocked mix of protein G- and A-Sepharose beads (GE Healthcare Life Sciences) for 3h at 4°C. Bound complexes were sequentially washed with sonication buffer, high salt buffer (50mM Hepes pH7.6, 500mM NaCl, 1mM EDTA, 1% TritonX, 0.1% Na-deoxycholate, 0.1% SDS) and TE buffer (10mM Tris, 1mM EDTA). Protein-DNA complexes were recovered in elution buffer (100mM $NaHCO_3$, 1%SDS) and the cross-linking was reversed by overnight incubation at 65°C. DNA fragments were treated with proteinase K for 2h at 42°C and purified using PCR purification kit (QIAGEN). Quantitative PCR was performed as described above using sets of primers available *upon request*. All the results are presented as 'fold recruitment' and represent the ratio of input percentage between treated and non-treated cells.

***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), CyclinC (santacruz, sc-1061) and β-tubulin (millipore, MAB3408).

For immunoprecipitation, the following antibodies were used: monoclonal antibodies against RNA polymerase 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) and H3K9ac (cell signaling, #9671) and H3K9me$^2$ (cell signaling, #9753).

## REFERENCES


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

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. 21, 381–395.

Belakavadi, M., and Fondell, J.D. (2006). Role of the mediator complex in nuclear hormone receptor signaling. Rev. Physiol. Biochem. Pharmacol. 156, 23–43.

Buchwalter, G., Gross, C., and Wasylyk, B. (2004). Ets ternary complex transcription factors. Gene 324, 1–14.

Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. Trends Biochem. Sci. 30, 250–255.

Donner, A.J., Szostek, S., Hoover, J.M., and Espinosa, J.M. (2007). CDK8 Is a Stimulus-Specific Positive Coregulator of p53 Target Genes. Mol. Cell 27, 121–133.

Graham, J.M., and Schwartz, C.E. (2013). MED 12 related disorders. Am. J. Med. Genet. A. 161, 2734–2740.

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.

Hazzalin, C.A., and Mahadevan, L.C. (2002). MAPK-regulated transcription: a continuously variable gene switch? Nat. Rev. Mol. Cell Biol. 3, 30–40.

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.

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.

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.

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.

Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. Trends Biochem. Sci. 30, 235–239.

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.

Lee, T.I., and Young, R.A. (2013). Transcriptional Regulation and Its Misregulation in Disease. Cell 152, 1237–1251.

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.

Li, B., Carey, M., and Workman, J.L. (2007). The Role of Chromatin during Transcription. Cell 128, 707–719.

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.

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.
Moghal, N. (2003). A component of the transcriptional mediator complex inhibits RAS-dependent vulval fate specification in C. elegans. Development 130, 57–69.
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.
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.
Peippo, M., and Ignatius, J. (2012). Pitt-Hopkins Syndrome. Mol. Syndromol. 2, 171–180.
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.
Poss, Z.C., Ebmeier, C.C., and Taatjes, D.J. (2013). The Mediator complex and transcription regulation. Crit. Rev. Biochem. Mol. Biol. 48, 575–608.
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.
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.
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.
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.
Stevens, J.L. (2002). Transcription Control by E1A and MAP Kinase Pathway via Sur2 Mediator Subunit. Science 296, 755–758.
Taatjes, D.J., Näär, A.M., Andel, F., Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. Science 295, 1058–1062.
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.
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.
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.
Yoda, A. (2005). Components of the transcriptional Mediator complex are required for asymmetric cell division in C. elegans. Development 132, 1885–1893.
Zhou, H., Kim, S., Ishii, S., and Boyer, T.G. (2006). Mediator Modulates Gli3-Dependent Sonic Hedgehog Signaling. Mol. Cell. Biol. 26, 8667–8682.
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.

# FIGURE 1

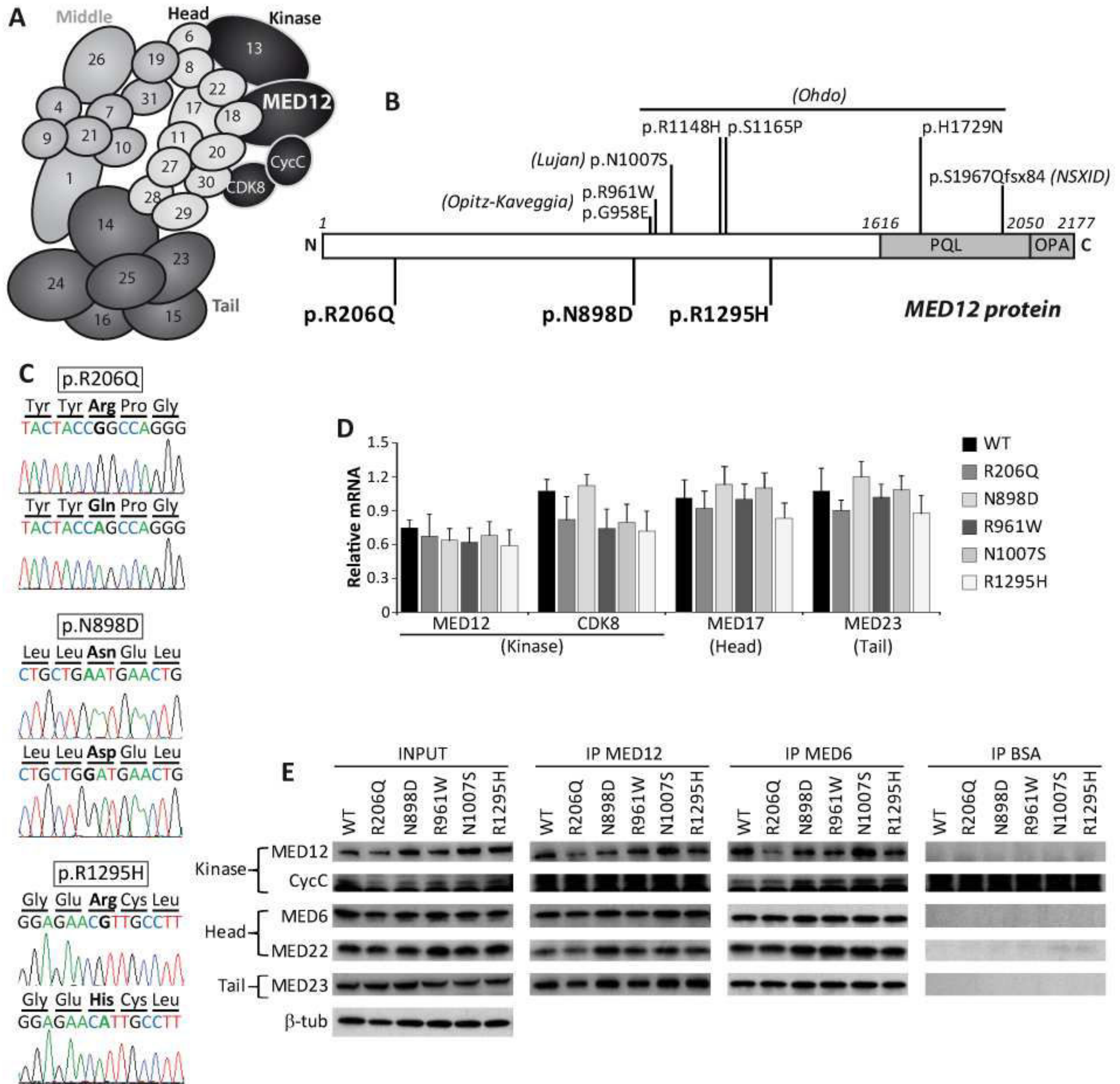


**Figure 1: Effect of mutations on MED12 gene expression and Mediator complex composition.**
**(A)** A composite depiction of the global structure of the human Mediator complex.
**(B)** Schematic overview of MED12 protein, including the PQL (proline-, glutamine- and leucine-rich) and the OPA (glutamine-rich) domain. The three recently identified amino acid changes are indicated (bottom), as well as previously published amino acid (top) changes leading to X-linked mental retardation including Opitz-Kaveggia, Lujan syndrome, the Maat–Kievit–Brunner type of Ohdo syndrome and non-syndromic profound X-linked intellectual disability (NSXID).
**(C)** Electropherograms showing the new discovered mutation (in bold) of affected individuals (bottom) and healthy control (top).
**(D)** Expression of *MED12*, *MED17, MED23* and *CDK8* 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.
**(E)** Analysis of MED complex composition by immunoprecipitation (IP) of MED6, MED12 or a control (BSA). Bound proteins were revealed by Western blot using antibodies against Cyclin C (CycC), MED6, MED12, MED22 and MED23. INPUT corresponds to 20% of the lysate used for IP reactions.

# FIGURE 2

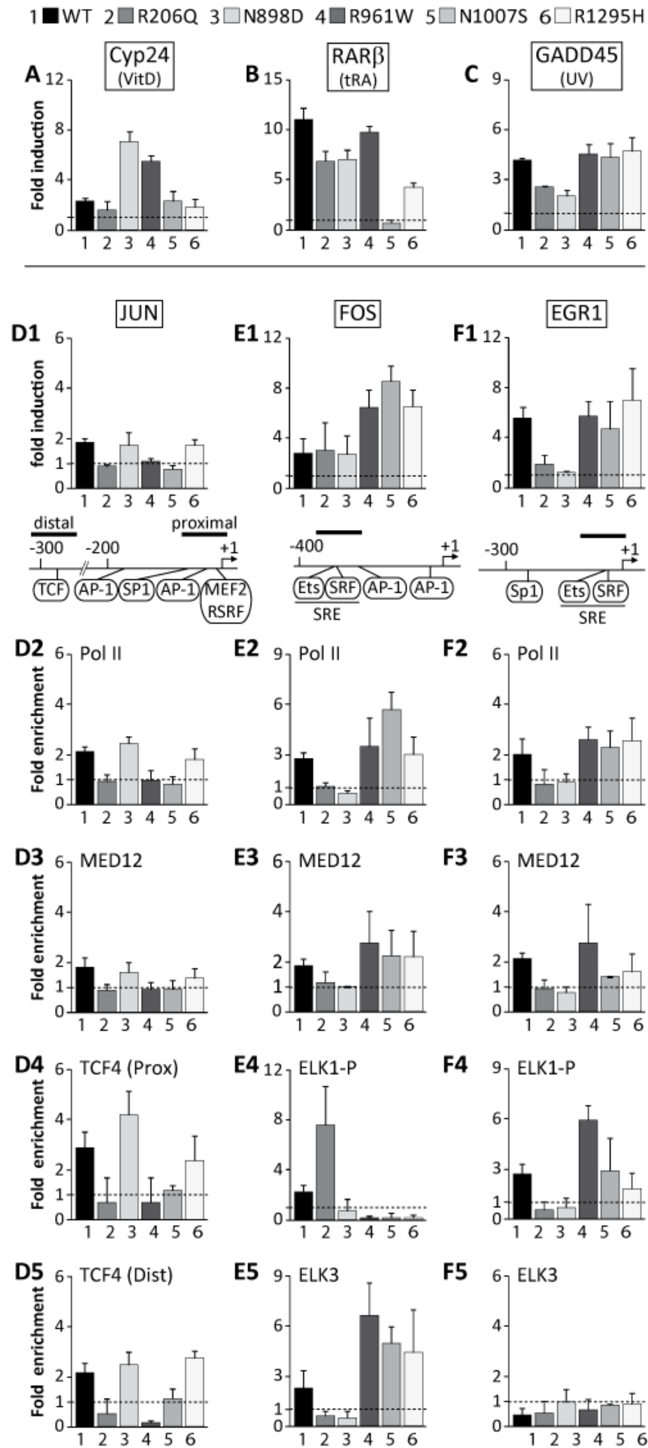


**Figure 2: Dysregulation of gene expression in MED12 lymphoblastoid cells.**

Relative mRNA expression of *CYP24* gene after 9h VitD(10 µM) treatment **(A)**, *RAR*$\beta$ gene after 6h tRA (10 µM) treatment **(B)**, *GADD45* genes after 8h of UV-irradiation (20 J/m²) **(C)**, and *JUN* **(D1)**, *FOS* **(E1)** and *EGR1* **(F1)** genes 30min after serum addition to serum-starved control (WT) or patients lymphoblastoid cells. In this and subsequent figure, values of mRNA expression after induction, 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.

Schematic representations of *JUN*, *FOS* and *EGR1* genes are indicated with the designed PCR amplicons. ChIP monitoring of the serum-dependent recruitment of Pol II **(D2, E2, F2)**, MED12 **(D3, E3, F3)**, TCF4 **(D4, D5)**, ELK1-P **(E4, F4)** or ELK3 **(E5, F5)** on the IEGs promoter was performed on chromatin fraction from WT or patients cells with serum for 10min. In this and subsequent figure, each series of ChIP are representative of at least two independent experiments. Values are expressed as fold enrichment, which represent the ratio of the INPUT percentage between treated and non-treated cells.

# FIGURE 3

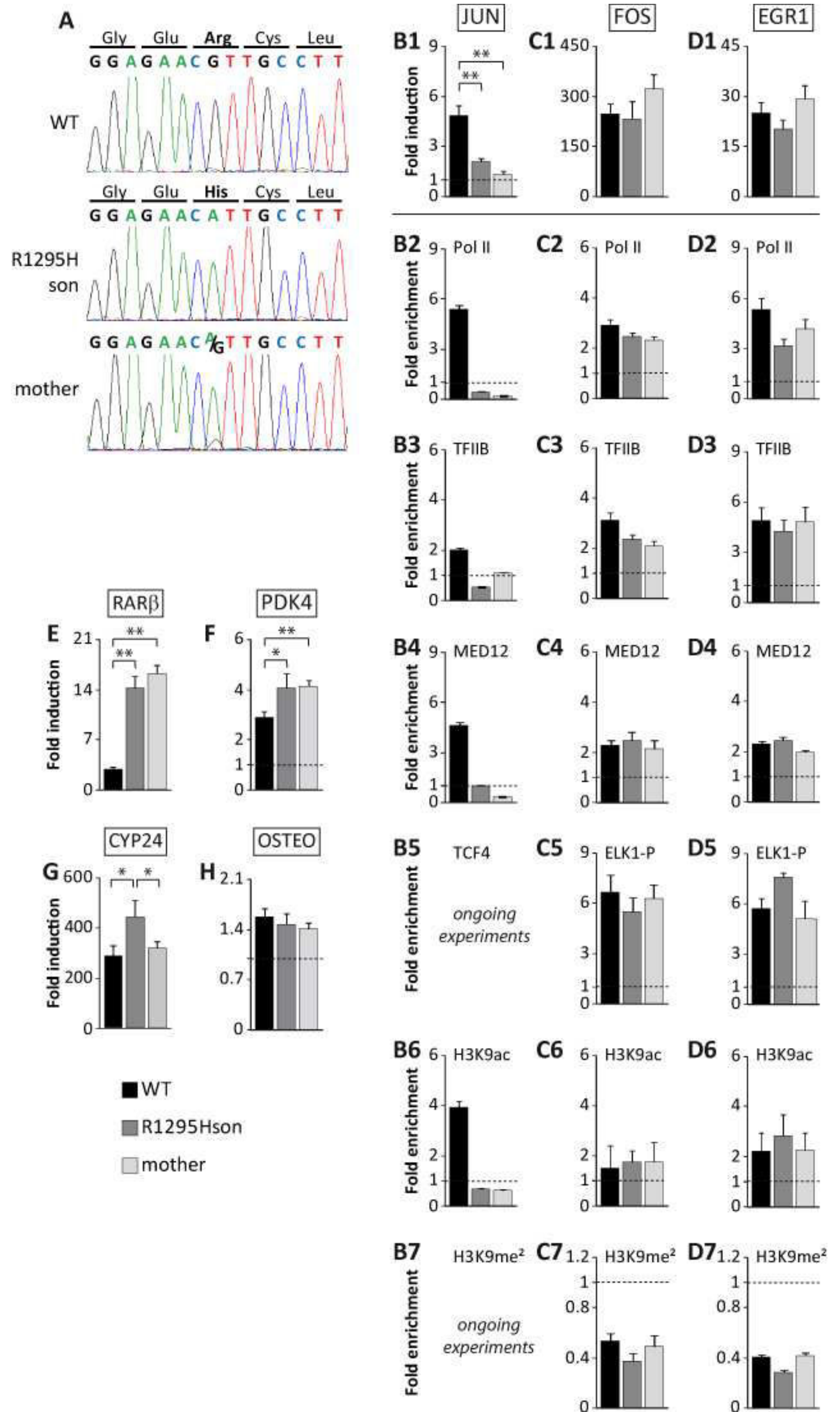


**Figure 3: Dysregulation of gene expression in MED12 fibroblasts.**

**(A)** Electropherograms showing the nucleotide variation c.3884 G>A of the *MED12* sequence in the son (middle), the mother (bottom) and a healthy control (top) fibroblasts.

The fibroblasts were treated with serum after serum starvation (20%, 30min) **(B, C, D)**, with retinoic acid (10µM; 8h) **(E, F)** or with Vitamin D (100nM; 8h) **(G, H,)** and relative mRNA expression of different responsive genes were monitored by RT-qPCR. (*, $P<0.05$; **, $P<0.01$, Student's t-test)

ChIP monitoring the serum-dependent recruitment of Pol II **(B2, D2, C2)**, TFIIB **(B3, D3, C3)**, MED12 **(B4, C4, D4)**, TCF4 **(B5)**, ELK1-P **(C5, D5),** acetylated H3K9 **(B6, C6, D6)** and dimethylated H3K9 **(B7, C7, D7)** on the IEGs promoter are performed on the chromatin fraction from WT or patients cells treated with serum.

# TABLE 1

**Table 1: Comparison of Clinical Features in Med12-associated disorders**

| | **?** | **?** | **FG** | **Lujan** | **Ohdo** | **?** |
|---|---|---|---|---|---|---|
| | R206Q | N898D | R961W<br>G958E | N1007S | R1148H<br>S1165P | R1295H |
| **Growth** | | | | | | |
| Tall Stature | | | - | + | - | + |
| Macrocephaly | + | + | + | + | - | |
| **Neurological** | | | | | | |
| Intellectual disability | + | + | + | + | + | + |
| Agenesis of corpus callosum | | | + | + | - | |
| Hypotonia | | + | + | + | + | |
| Behavioural disturbance | + | | + | + | + | |
| Speech abnormalities | | | - | - | + | |
| **Craniofacial** | | | | | | |
| Long narrow face | | | - | + | - | |
| Tall prominent forehead | | + | + | + | + | + |
| Triangular face | | | - | - | + | |
| Blepharophimosis | | | - | - | + | |
| Downslanting palpebrae | | | + | + | + | |
| Strabismus | + | | + | + | + | |
| Hypertelorim | | | + | - | - | |
| Small ears | | + | + | - | + | |
| Philtrum | | | | short | long | |
| Maxillary hypoplasia | | | + | + | - | + |
| Micrognathia | | | + | + | + | |
| High narrow palate | | | + | + | + | + |
| Open mouth | | | + | + | + | |
| Dental anomalies | | | + | + | - | |
| **Extremity** | | | | | | |
| Foetal finger pads | | | + | - | - | |
| Syndactyly | | | + | - | - | |
| Broad thumbs | | + | + | + | - | |
| Horizontal palmar crease | | | + | - | - | |
| Long hyperextensible digits | | | - | + | + | |
| **Cardiovascular** | | | | | | |
| Congenital heart defect | Left ventricular hypertrophy | | - | - | - | |
| **Gastrointestinal** | | | | | | |
| Constipation | | | + | - | + | |
| Anal anomalies | | | + | - | - | |
| **Genitourinary** | | | | | | |
| Genital anomalies | | | + | - | + | + |
| **Others** | Extra Nipples | | | | | |

# TABLE 2

**Table 2: Comparison of genes response in the different MED12 mutated lymphoblastoid cells.**

| | VitD | tRA | Serum | | | UV | | |
|---|---|---|---|---|---|---|---|---|
| | *CYP24* | *RARβ* | *JUN* | *FOS* | *EGR1* | *JUN* | *FOS* | *GADD45* |
| **MED12/R206Q** | = | = | - | = | - | + | + | - |
| **/N898D** | + | = | = | = | - | = | = | - |
| **/R961W** | + | = | - | + | = | = | + | = |
| **/N1007S** | = | - | - | + | = | = | = | = |
| **/R1148H** | n.d. | n.d. | n.d. | = | = | n.d. | n.d. | n.d. |
| **/S1165P** | n.d. | n.d. | n.d. | = | - | n.d. | n.d. | n.d. |
| **/R1295H** | = | = | = | + | = | = | + | = |

+ : upregulated;
= : similar;
- : downregulated compared to WT.
n.d.=no data

# FIGURE S1

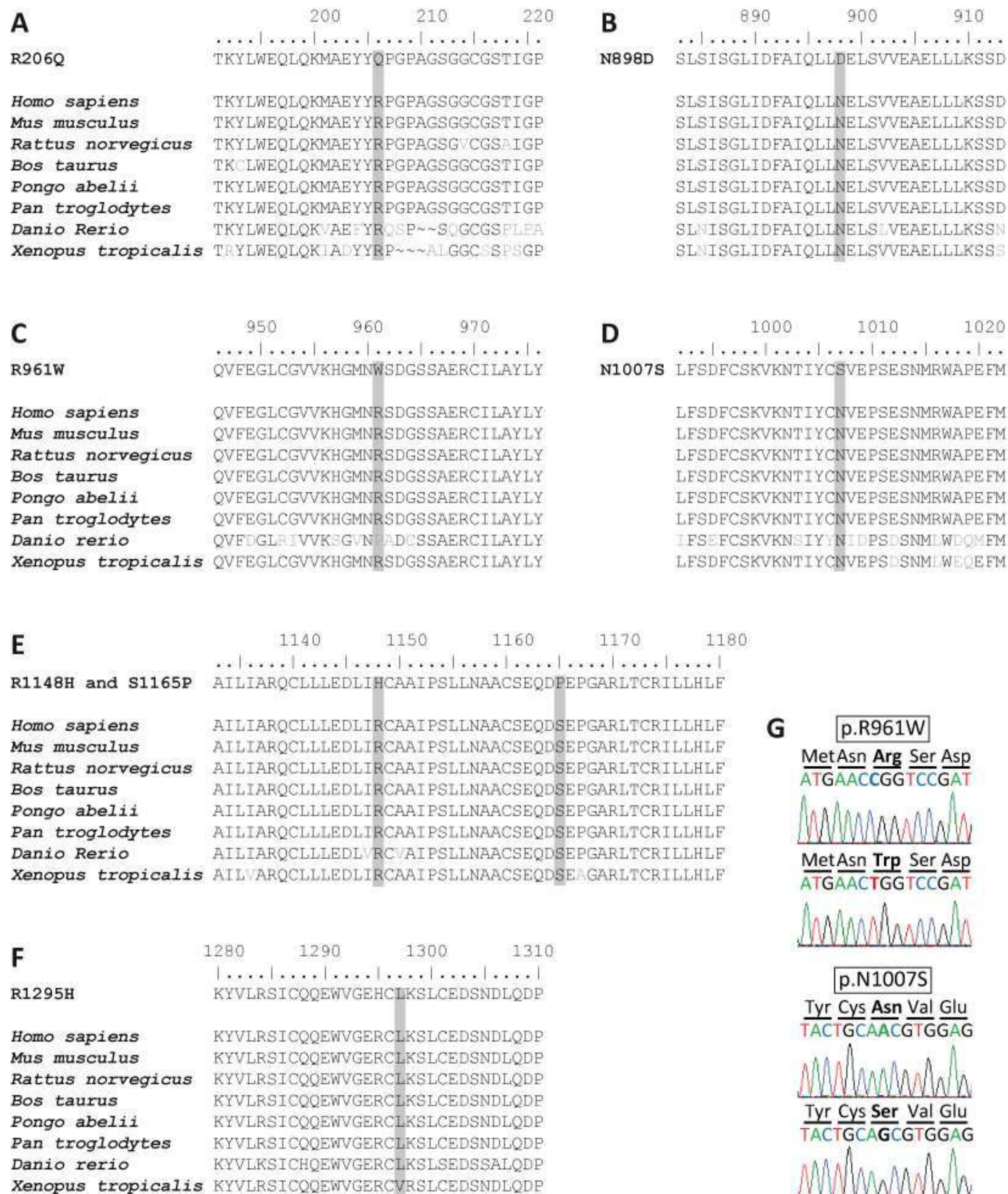



**Figure S1: MED12 sequence conservation**

Amino acid conservation between species at and around the residues R206 **(A)**, N898 **(B)**, R961 **(C)**, N1007 **(D)** R1148 and S1165 **(E)** and R1295 **(F)** in MED12 (residue highlighted in grey).

# FIGURE S2

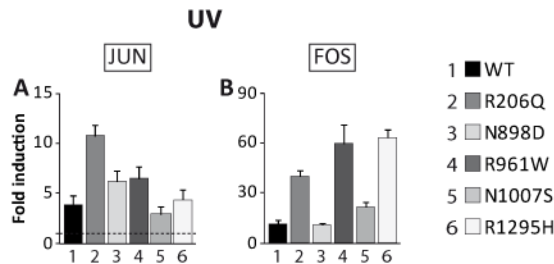



**Figure S2: Dysregulation of gene expression in MED12 lymphoblastoid cells after UV irradiation.**

Relative mRNA expression of *JUN* **(A)** and *FOS* **(B)** genes after 2h and 1h respectively of UV-irradiation (20 J/m²) in WT or patients lymphoblastoid cells. 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 S3

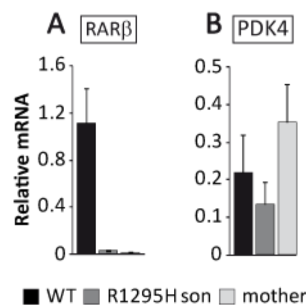



**Figure S3: R1295H mutation in MED12 impairs *RARβ* basal expression.**

Relative mRNA-expression of *RARβ* and *PDK4* in physiological condition. The results are the mean of three independent experiments and represent the relative expression level of the gene versus *GAPDH*.

# FIGURE S4

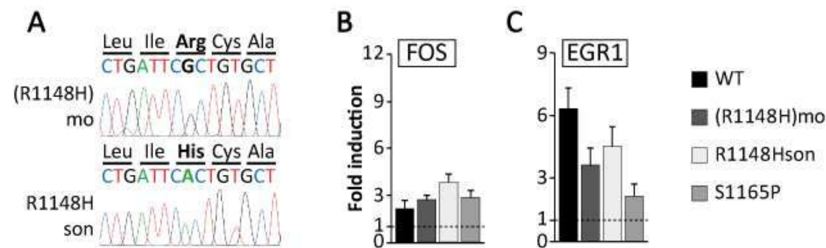



**Figure S4: Effect of MED12 mutations associated with Ohdo syndrome on *FOS* and *EGR1* expression**

**(A)** Electropherograms showing mutation (in bold) of an affected individual (bottom) and his mother expressing the WT allele (top).

mRNA expression level of *FOS* ***(B)*** and *EGR1* **(C)** genes 30min after serum addition to serum-starved control (WT) or patients lymphoblastoid cells. 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.

## PERSPECTIVES

As indicated in the figures, some ChIP experiments are missing and are currently in progress. We also plan to analyze the effect of MED12/R1295H mutation on the binding of Pol II and MED12 genome-wide by ChIP-sequencing experiment.

To verify that genes expression alteration observed in patients fibroblasts is due to MED12/R1295H mutation, cells were transfected with a plasmid engineering to express wild-type MED12 protein. Unfortunately, transfection by its-self affects genes expression. We thus changed our approach. Indeed, mother fibroblasts are composed of two different cells: ~¾ of cells express the mutant allele of MED12, while ~¼ the WT allele. I presently perform a clonal selection of these cells in order to separate the two populations. Later, transcriptional activation of some given genes will be analyzed and compared with previous results.

To generate more proper and robust cellular models for MED12, we are applying CRISPR/Cas9 system to knock-out *MED12* gene in HeLa, neuroblastoma (IMR32, BE(2), Lan1) and oligodendroglioma (HOG), the cellular types affected in patients carrying MED12 mutations. MED12 expression will further be restored via stable transfection of plasmid engineering to express WT or mutated form of MED12 protein fused to a tag B10. This system will allow to study the effect of each MED12 mutation on genes expression in different cell types.

In order to reveal the effect mechanism of MED12 dysfunction during diseases development and progression, we have reprogrammed the fibroblasts (R1295Hson, mother and WT) into Induced Pluripotent Stem cells (hIPSc). We also planned to generated hiPSc from the EBV-immortalized lymphoblastoid cell lines carrying the different MED12 mutations (R206Q, N898D, R961W, N1007S, R1148H, S1165P and R1295H) (Rajesh et al., 2011). Considering the clinical features observed in MED12 patients, the hiPSc will further be differentiated in various cells types (neurons, ...etc) and the effect of MED12 mutations on gene expression and on differentiation efficiency will be analyzed.

# MED17 PROJECT:
# Infantile Cerebral and Cerebellar Atrophy
# and mutation p.L371P in *MED17*

## INTRODUCTION

Besides MED12 subunit, we are also interested in MED17, which belongs to the Head module. A homozygous missense mutation in MED17 (p.L371P) is associated with Infantile cerebral and cerebellar atrophy (Kaufmann et al., 2010). In addition to better understand the etiology of this disease, my thesis project was also focused in having a better comprehension of MED17 role in gene regulation. Using two cells models, patients lymphoblastoid cells and transgenic mice embryonic fibroblasts, our preliminary results show that the mutation does not affect the mRNA expression of Mediator subunits, as well as the stability of those proteins. Interestingly, the composition of Mediator complex seems to be affected in mutated cells from mice but not from human. Moreover, MED17/L371P mutation disturbs the expression of Immediate Early Genes (IEGs) in response to serum mitogens by altering transactivation complex formation. It also impairs the response of genes activated by UV irradiation or nuclear hormone receptors.

## RESULTS & DISCUSSION

### *Two cells models*

We obtained lymphoblastoid cells (LCs) from a patient carrying the homozygous mutation p.L371P (c.1112 T>C) in *MED17* gene (Figure 12). In order to have a second cellular model, we also isolated MEFs (Mouse Embryonic Fibroblasts) from homozygous Med17/L371P mouse embryos which have been generated in collaboration with the Mouse Clinical Institute through GENCODYS project (see Mice models project). MEFs are easier to handle than lymphoblastoid cells and allow to have a control with the same genetic background, i.e. MEFs from the same litter of animals.

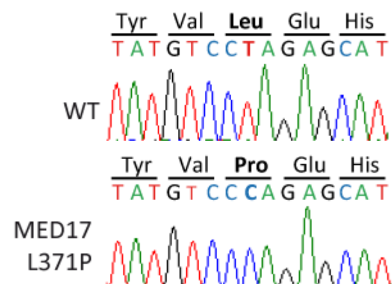



**Figure 12: MED17/L371P mutation.**
*Electrophoregrams showing the mutation c.1112 T>C in MED17 sequence of an individual with Infantile cerebral and cerebellar atrophy (bottom) and in comparison with the WT sequence (top).*

***MED17/L371P mutation does not change MED subunits gene expression and protein stability***

As previously done for MED12 mutations, I first evaluated the impact of MED17/L371P mutation on expression of the corresponding transcript and protein product in both cell models, LCs and MEFs. In normal (WT) and mutated (MED17/L371P) cells, RT-qPCR shows similar expression of *MED17* gene, as well as of subunits belonging either to the Kinase, Middle or Tail module (Figure 13A and B). Western-blot analysis performed on whole cell extracts reveals comparable amount of MED17 protein in WT and MED17/L371P cells (Figure 13C and D). A similar result is observed for other Mediator subunits from the Head (MED6), Tail (MED16 and MED23) and Kinase (MED12, CDK8 and CYCLIN C (CCNC)) modules. Consequently, the mutation does not affect the mRNA expression of Mediator subunits, as well as the stability of those proteins.

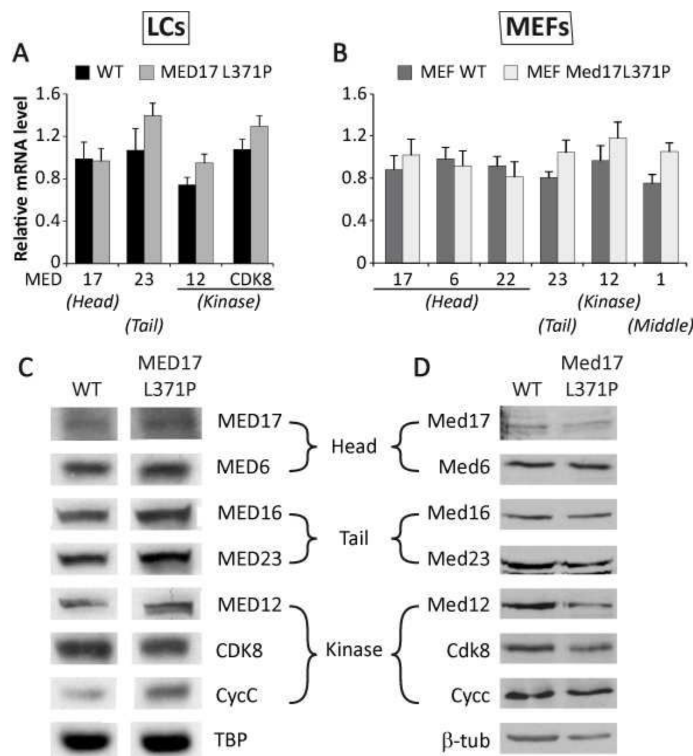



**Figure 13: Effect of Med17/L371P mutation on *MED17* expression and protein stability.**
***(A-B)*** *Expression of Med17, as well as other MED subunits belonging to the Head, Middle, Tail and Kinase modules, in control (WT) and mutant cells (MED17L371P). Values are normalized relatively to housekeeping gene (GAPDH gene for LCs and 18S gene for MEFs). Error bars represent the standard deviation of at least three independent experiments.*
***(C-D).*** *Immunoblot analysis of MED17 and six other Mediator subunits from the Head, Tail and Kinase modules in WT and mutated cells. The blots represent at least two independent experiments.*
*Experiments was performed both in Lymphoblastoids cells (LCs)* ***(A-C)*** *and in Mouse Embryonic Fibroblasts (MEFs)* ***(B-D)***

***Different effect of MED17/L371P on MED composition between human and mouse***

MED17 is a central component of Mediator architecture, playing a critical role in assembly of the Head module (Takagi et al., 2006; Imasaki et al., 2011; Tsai et al., 2014). Considering the well-defined Head module structure of yeast and the functional homology around the L371 residues between yeast and human, it seems that the α-helix of Med17,

containing disease mutation, interacts with Med11 and Med22 subunits (Figure 14A and B). Moreover, 2D structure prediction of the L371P substitution shows that the mutation inhibits helix formation (Figure 14C) (Petersen et al., 2009). Altogether, these data suggest that the mutation could disrupt module assembly.

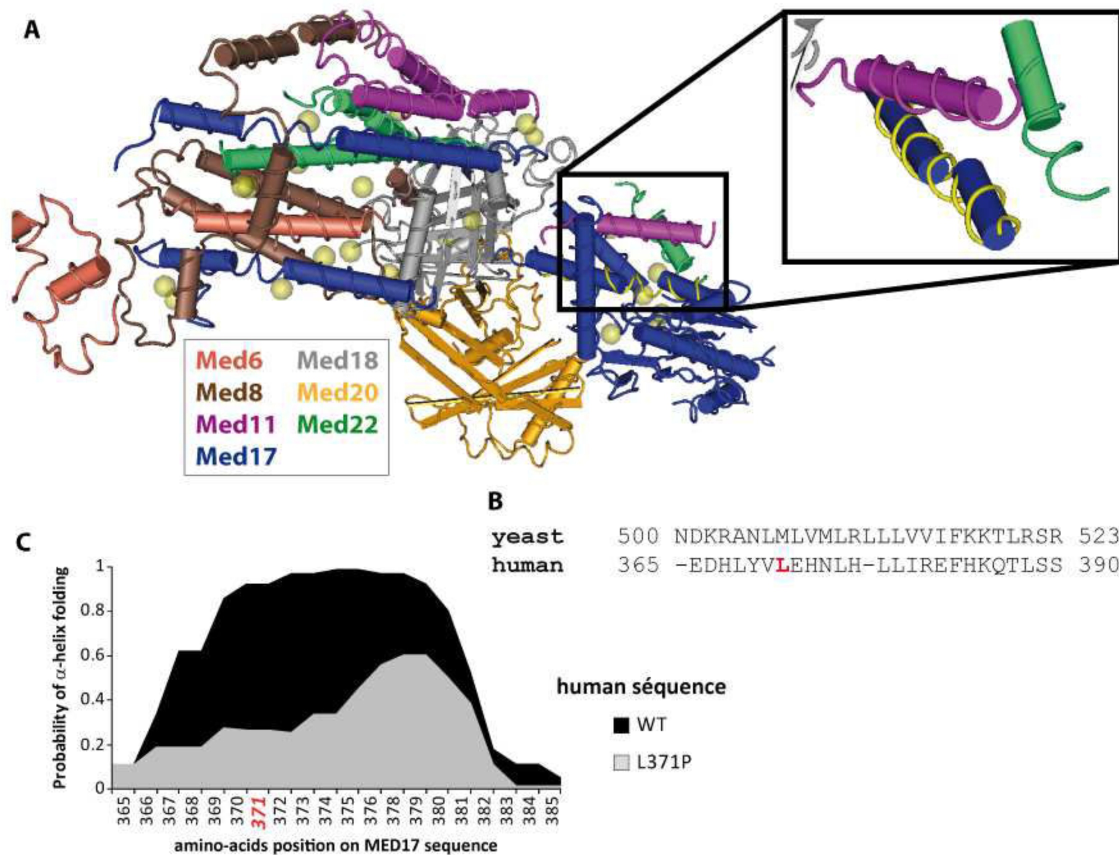



**Figure 14: Predicted effect of MED17/L371P mutation.**

***A.*** *Overall structure of the yeast Mediator head module (adapted from Imasaki et al., 2011). The α-helix of Med17 (residue 500 to 523) containing disease mutation are colored in yellow. Magnified window shows the atoms which are less than 10Å away from Med17 helix.*

***B*** *Sequence alignment of the human and yeast Med17 corresponding to the α-helix which contains the residue L371 in human.*

***C.*** *The probability of α-helix folding is predicted for human MED17 sequence from residue 365 to 385, as well as the consequence of the substitution of Leucine with Proline at position 371.*

I thus investigated Mediator architecture by performing Immuno-precipitation (IP) experiment using an antibody against one of the MED subunits. In MEFs, this analysis, carried out by means of Med6-specific antibodies, reveals less precipitation of Med22 subunit in Med17/L371P cells compared to WT cells (Figure 15A). In comparison, Med6, Med17 and Med22 subunits are precipitated with a similar fashion in both cells (Figure 15A). These results were observed two times. However, in LCs, a first experiment shows a similar pattern of

precipitation between control and mutant cells for MED17, as well as MED12, MED22 and MED23, by using antibodies against MED1 and MED6 (Figure 15B).

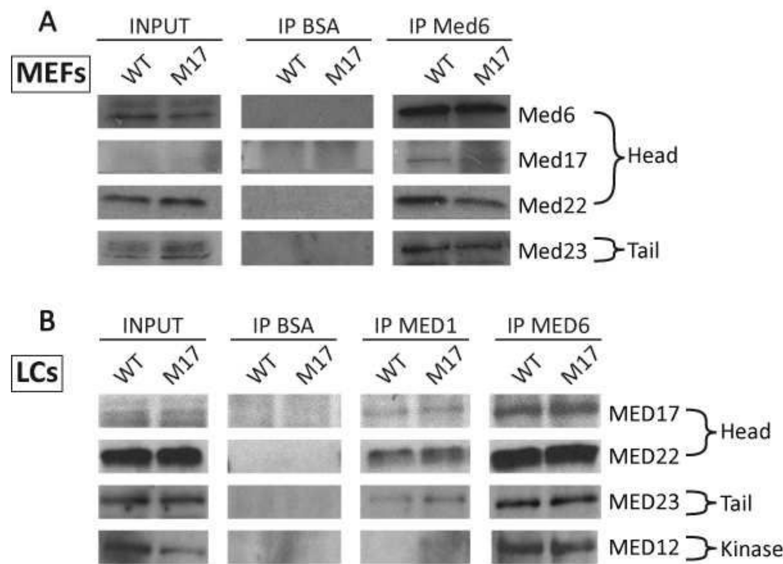



**Figure 15: Effect of Med17/L371P mutation on MED composition**
*MED complex organization was analyzed in both MEFs **(A)** and LCs **(B)** by immune-precipitation (IP) against MED6, MED1 or a control (BSA). The bound proteins were revealed by Western blot (WB) using antibodies against MED6, MED12, MED17, MED22 and MED23. Input, 20% of the lysate used for IP reactions.*

Although overall Mediator architecture seems to be conserved across eukaryotes, our data reveal that the mutation MED17/L371P affects MED composition in mice but not in human (Tsai et al., 2014). However, this difference could be attributed to the different cell lines used, i.e. lymphoblastoid cell in human and fibroblast in mice.

***MED17 mutated cells respond differently to external stimuli***

I proceeded as before and analyzed the effect of MED17/L371 mutations on gene expression induced by treatment of cells with serum, UV-irradiation or hormones. In LCs, we found that, upon serum induction, *JUN* and *EGR1* expression is down-regulated, while expression of *FOS* remains unchanged in MED17/L371P cells compared to control (Figure 16 A1, B1 and C1). We observed a parallel between the expression of IEGs and the recruitment of Pol II at their respective promoter (Figure 16 A2, B2 and C2). Indeed, *JUN* and *EGR1* down-regulation matches with a failure in Pol II recruitment (compare panels A1 with A2 and B1 with B3). By contrast, recruitment of Pol II at *FOS* promoter is similar in both WT and MED17/L371P cells (panel C2).

We observed that the recruitment of the transcription factor TCF4 at *JUN* promoter is impaired when Pol II binding is defective (Figure 16 A3). In mutated cells, ELK1 recruitment on *FOS* promoter is increased whereas the recruitment of its paralog ELK3 is reduced compared to

WT cells; suggesting that the mutation altered the constitution of what nonetheless remains a functional PIC.

Our results demonstrate that in LCs the mutation MED17/L371P disturbs the expression of Immediate Early Genes (*JUN*, *FOS* and *EGR1* genes) by altering pre-initiation complex formation at their promoters. This seems to confirm our previous hypothesis that intellectual disability observed in patients with mutations in Mediator or TFIIH could be the result of impaired fine-tuning of IEGs expression during development (Hashimoto et al., 2011).

Unexpectedly, in control and mutated MEFs, we observed similar expression of *Jun*, *Fos* and *Egr1* gene upon serum induction (Figure 16 H, I and J). However, a previous study showed that loss of MED23 severely affect *Egr1* expression in mouse Embryonic Stems cells, but only modestly in MEFs cells (Balamotis et al., 2009). ChIP experiments revealed that the difference between the two cells types is due to change in the relative amount of three related transcription factors at *Egr1* promoter. A similar process could explain our difference in LCs and MEFs for *Jun*, *Fos* and *Egr1* gene expression.

Knowing that some genes can be under the control of different stimuli, I next investigated the behavior of *JUN* and *FOS* genes, as well as *GADD45* gene, in response to UV irradiation in LCs. We observed a similar expression of *FOS* and *GADD45* genes in both cells (Figure 16 E and F). Interestingly, exposure to UV leads to an over-expression of *JUN* in MED17/L371P compared to control (panel D), rather than to its down-regulation, as it was expected. (panel D). Our data show that the mutation affects differently the expression of *JUN* gene depending on the nature of the stimulus (serum vs. UV).

The L371P mutation has no effect on *RAR*$\beta$ expression induced by tRA treatment in LCs (Figure 16 G). In contrast, following VitD induction in MEFs, *Cyp24* gene is up-regulated, whereas *Osteopontin* gene is down-regulated in mutated cells compared to control (Figure 16 K and L). Interestingly, we also observed up-regulation of *CYP24* gene in fibroblasts baring R1295H mutation in MED12 (see MED12 results), which suggest a similar effect on PIC formation of both MED12/R1295H and MED17/L371P mutations.

Altogether, the above data show a specific involvement of MED17 in the expression of genes regulated by different stimuli.

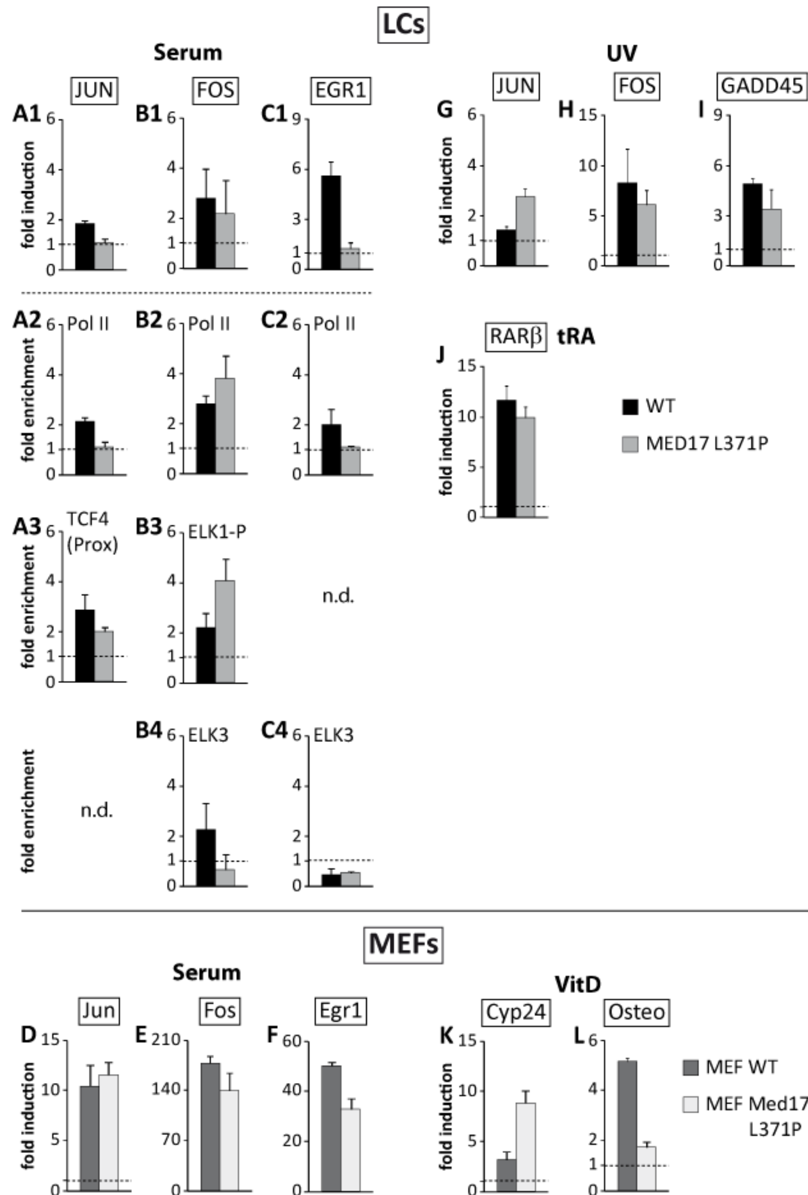



**Figure 16: Dysregulation of gene expression in MED17/L371P cells after various stimuli.**
*The mRNA expression of JUN **(A1-D)**, FOS **(B1-E)** and EGR1 **(C1-F)** genes 30min after serum addition to serum-starved cells is monitored, as well as of JUN **(G)**, FOS **(H)** and GADD45 **(I)** genes after 2h, 1h or 8h respectively of UV-irradiation (20 J/m²). The mRNA expression of RAR$\beta$ gene **(J)** 6h after tRA (10 µM) addition is also examined, as well as Cyp24 **(K)** and Osteopontine **(L)** genes after 8h of VitD (100 nM) treatment. The values from at least two different experiments are presented in fold induction, meaning the ratio of treated cells relative to non-treated cells after normalization against housekeeping gene (GAPDH gene for LCs and 18S gene for MEFs).*
*ChIP monitoring the serum-dependent recruitment of Pol II **(A2, B2, C2)**, TCF4 **(A3)**, ELK1-P **(B3)** and ELK3 **(B4, C4)** on the promoter of the three IEGs in LCs. Values are expressed as fold enrichment, which represent the ratio of the INPUT percentage between treated and non-treated cells. n.d.=no data*

**PERSPECTIVES**

Most of the results are still preliminary and should be confirmed. In addition, ChIP experiments will be pursued to study the composition of the transcription machinery at the

promoter of dysregulated genes in mutated cells. We will also analyze the histone post-translational modifications necessary for chromatin remodeling (methylation of H3K4 and H3K9 and acetylation of H3K9).

# NIPBL PROJECT:

# Cornelia de Lange syndrome and NIPBL mutation

## INTRODUCTION

Cohesin complex is a dynamic multiprotein complex, which works cooperatively with Mediator complex to facilitate the formation of specific chromatin structure (Kagey et al., 2010; Muto et al., 2014). The core of the cohesin complex is made up of four subunits: SMC1, SCM3, RAD21 and SA1/SA2. In addition, this core requires NIPBL (Nipped-B-like) and its partner MAU2 for its loading on chromatin. Mutations in cohesin complex are responsible for Cornelia de Lange Syndrome (CdLS), in which a number of characteristics are common with MED12-related disorders, such as intellectual disability. We thus decided to study the role of NIPBL during transcription process and thus gain a better comprehension of the etiology of CdLS. Using fibroblasts from patient carrying a mutation in NIPBL, I demonstrated an autoregulation of NIPBL and preliminary results suggest a role of this protein in the basal expression of *RAR$\beta$* gene.

## RESULTS & DISCUSSION

### *NIPBL regulates its own expression*

We obtained fibroblasts from an individual with Cornelia de Lange syndrome (CdLS). This patient has a heterozygous mutation c.6516-6517insA in *NIPBL* gene, mutation leading to the expression of a truncated protein (Figure 17A). The consequence of this mutation on *NIPBL* RNA expression was assessed by RT-qPCR. I observed a significant decrease of *NIPBL* mRNA levels in CdLS NIPBL cells compared to control cells, whereas the expression of its interactants (Mediator and cohesin subunits) remains unchanged (Figure 17B). Similar reductions of *NIPBL* mRNA have been reported in other cell lines derived from individuals with CdLS (Borck et al., 2006; Liu et al., 2009). I could not check whether transcript reduction results also in protein reduction, due to the absence of a good antibody against NIPBL.

In order to understand the decrease of *NIPBL* mRNA expression, I next analysed, by ChIP experiments, the recruitment of transcriptional factors on *NIPBL* promoter. Preliminary results reveal less occupancy of Pol II, NIPBL and SMC1 at *NIPBL* promoter in CdLS cells compared to WT, while TFIIB occupancy remains unchanged (Figure 17C). Altogether these data suggest that NIPBL regulates its own expression. This explanation is favoured by previous studies showing that both Nipbl$^{+/-}$ *Drosophila* and mice exhibit only a 25-30% drop in transcript levels, rather than an expected decrease of 50% (Rollins et al., 2004; Kawauchi et al., 2009).

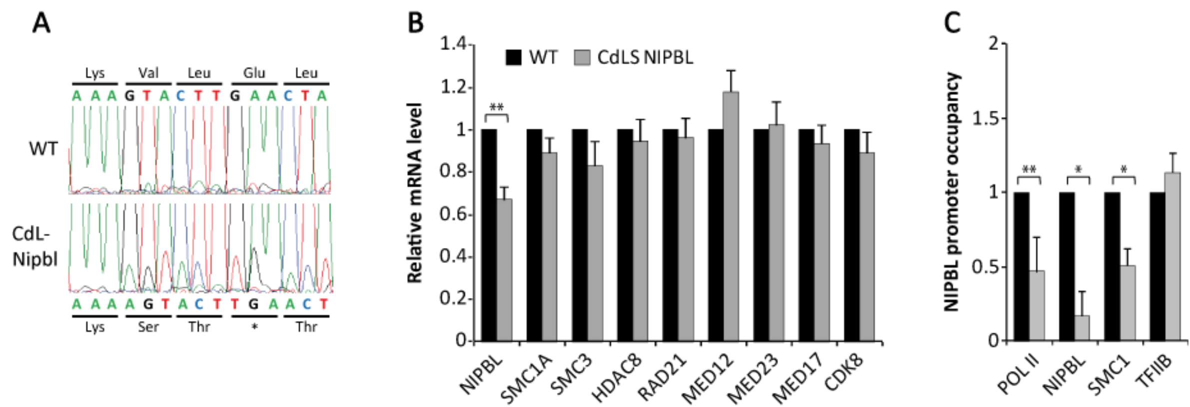



**Figure 17: Effect of *NIPBL* mutation on its gene expression.**
***A.*** *Electrophoregrams showing the mutation c.6516-6517insA in NIPBL sequence of an individual with Cornelia de Lange syndrome (bottom) and the WT sequence (top).*
***B.*** *Expression of NIPBL gene, as well as genes of cohesin and Mediator complex in WT and CdLS NIPBL cells. The values are normalized relative to GAPDH gene and expressed relative to mRNA levels in control cells, which are arbitrarily assigned a value of 1. In this and subsequent figures, error bars represent the standard deviation of at least three independent experiments.*
***C.*** *ChIP monitoring the occupancy of Pol II, NIPBL, SMC1 and TFIIB at NIPBL promoter. The level of occupancy for each protein is expressed relative to its level of occupancy in control cells, which are arbitrarily assigned a value of 1. Error bars represent the standard deviation of at least two independent experiments.*
*Asterisks denote statistically significant values relative to control (Student's t test, *p < 0.05, **p < 0.01).*

### *NIPBL mutation does not affect IEGs expression*

We previously hypothesized in our group that intellectual disabilities (ID) linked to mutations in Mediator and TFIIH could be associated with altered expression of the Immediate Early Genes (IEGs) *JUN*, *FOS* and *EGR1* ((Hashimoto et al., 2011). Considering that ID is one of the clinical features of CdLS and that cohesin complex interact with Mediator, I next evaluated by RT-qPCR the effect of *NIPBL* mutation on the expression of these three IEGs after serum addition to serum-starved cells. Contrary to what is expected, expression of *JUN*, *FOS* and *EGR1* is similar in WT and CdLS NIPBL cells (Figure 18).

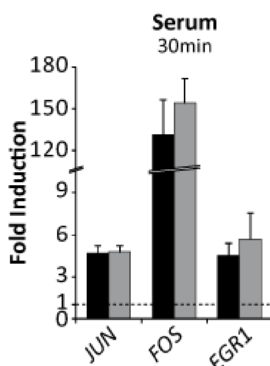



**Figure 18: Effect of *NIPBL* mutation on the expression of IEGs.**
*The fibroblasts were treated 30min with serum (20%) after serum starvation and relative mRNA expression of JUN, FOS and EGR1 genes were monitored by RT-qPCR. The values are presented in fold induction, meaning the ratio of treated cells relative to non-treated cells after normalization against GAPDH.*

However, about 40 cellular IEGs have been identified so far. It is possible that IEGs other than the three studied are deregulated. Gene expression analysis by microarray experiments after serum addition to serum-starved fibroblasts should be performed to identify the altered Immediate Early Genes.

***NIPBL mutation impairs only a subset of genes***

Considering the dysmorphic features of patients, we also analyzed the effect of *NIPBL* mutation in the expression of genes induced by hormones, such as Vitamin D (VitD) and all trans retinoic acid (tRA). Expression of *CYP24* and *OSTEOPONTIN* are unchanged after VitD treatment in both control and CdLS NIPBL cells (Figure 19).

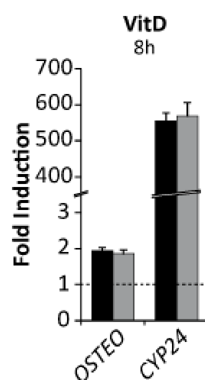



**Figure 19: Effect of *NIPBL* mutation on the expression of VitD response genes.**
*The fibroblasts were treated 8h with Vitamin D (100nM) and relative mRNA expression of CYP24 and OSTEOPONTIN genes were monitored by RT-qPCR. The values are presented in fold induction.*

In contrast, after tRA treatment, *RAR*$\beta$ gene is overexpressed 50 times more in mutated cells compared to controls (Figure 20A). Analysis of other tRA-target genes (*TMG2*, *RAR*$\alpha$, *NRIP1*, *SMAD3* and *PDK4*) shows no dysregulation as important as observed for *RAR*$\beta$. Indeed, *TMG2* and *RAR*$\alpha$ expression is up-regulated only 2 times more in CdLS NIPBL cells compared to WT cells (Figure 20B and C), while expression of *SMAD3* genes remains unchanged (Figure 20D). Furthermore, in response to tRA treatment, *NRIP1* is slightly down regulated while the induction of *PDK4* is delayed in time in mutated cells (Figure 20E and F). These results demonstrate a role of NIPBL in transcriptional regulation of only a subset of genes.

Interestingly, as observed in fibroblast carrying *MED12* mutation, the basal level of *RAR*$\beta$ expression is reduced several-fold in CdLS NIPBL cells, while it is not the case of other RA-target genes (Figure 20, time 0). ChIP experiments reveal less occupancy on *RAR*$\beta$ promoter of the transcription machinery (Pol II, RARα and MED1), as well as of the cohesin complex (NIPBL and SMC1) in CdLS cells compared to WT (Figure 21). We also observed a reduction of the histone

mark associated with active transcription, i.e. acetylation of histone H3 at lysine 9 (H3K9ac). These results reveal a role of NIPBL on *RARβ* basal expression.

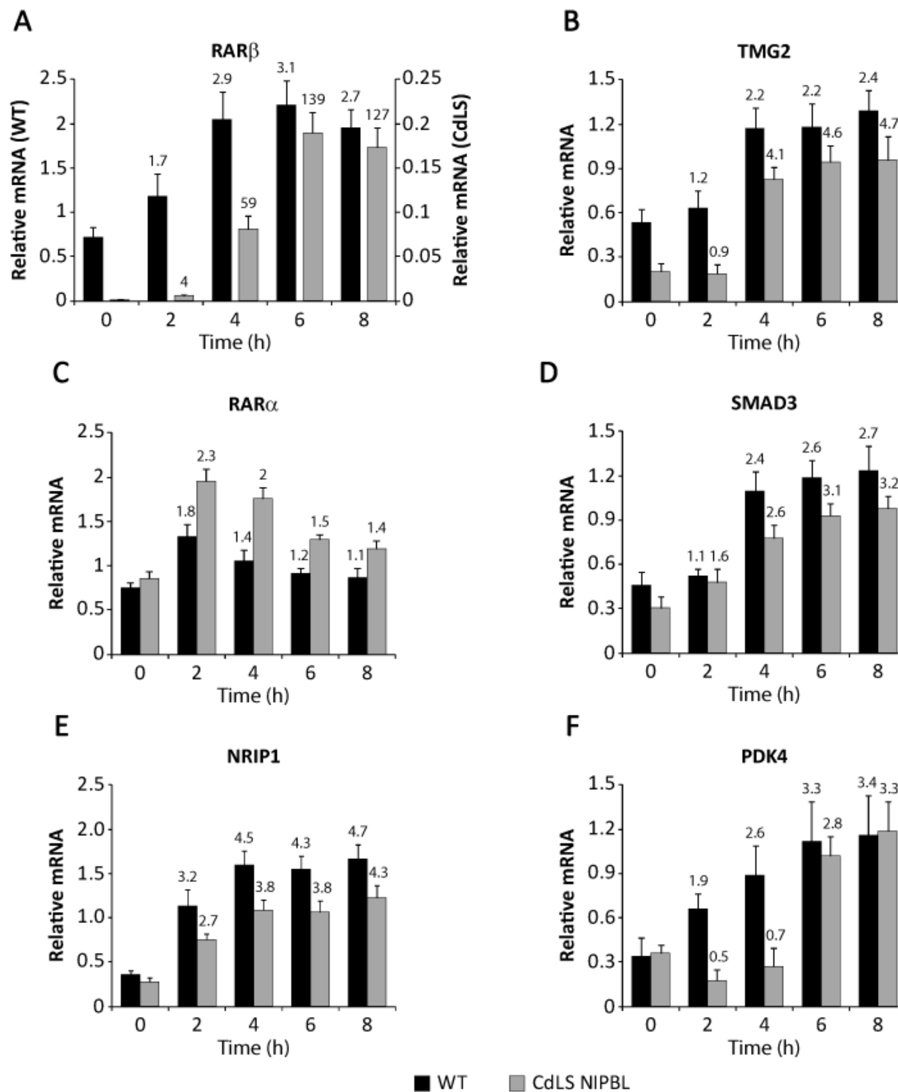



**Figure 20: Effect of *NIPBL* mutation on the expression of RA-target genes.**
*Time-course analysis of RARβ* ***(A)****, TMG2* ***(B)****, RARα* ***(C)****, SMAD3* ***(D)****, NRIP1* ***(E)*** *and PDK4* ***(F)*** *after all-trans retinoic acids treatment (10μM) to wild-type (WT) and CdLS NIPBL cells. The values are normalized to the housekeeping GAPDH gene. The fold induction values are indicated above each bar.*

Retinoic acid (RA) is involved in pleiotropic functions during vertebrate embryogenesis, such as brain or limb developement (Rhinn and Dolle, 2012). By consequence, dysregulation of RA-responsive gene and more particularly of *RARβ* gene observed in cells from patients could possibly contribute to some characteristics of Cornelia de Lange Syndrome, like intellectual disability and upper limb abnormalities.

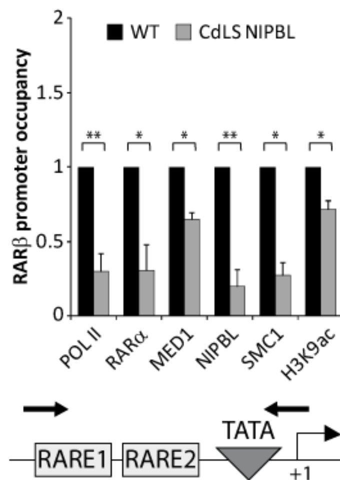



**Figure 21: Binding of transcriptional factors at *RAR$\beta$* promoter in physiological condition.**
*ChIP monitoring the occupancy of Pol II, RAR$\alpha$, MED1, NIPBL, SMC1 and H3K9ac on RAR$\beta$ promoter. The level of occupancy for each protein is expressed relative to its level of occupancy in control cells, which are arbitrarily assigned a value of 1. Error bars represent the standard deviation of at least two independent experiments.*
*Schematic representation of RAR$\beta$ gene with the indicated amplicons.*
*Asterisks denote statistically significant values relative to control (Student's t test, *$p < 0.05$, **$p < 0.01$).*

## PERSPECTIVES

Considering the above data, a microarrays experiment will be conducted to examine the effect of NIPBL mutation on genes expression after retinoic acid treatment. Moreover, our laboratory has shown that the formation of a DNA loop between the promoter and the terminator of *RAR$\beta$* gene is required for its optimal expression (Le May et al., 2012). In addition, NIPBL, together with Mediator, facilitates the formation of these chromatin loops (Kagey et al., 2010; Muto et al., 2014). Consequently, q3C (quantitative Chromatin Conformation Capture) experiment will be undertaken to analyze the spatial organization of *RAR$\beta$* gene in the cells from CdLS patients before and after tRA treatment. Later, Hi-C experiment could perhaps be considered in order to analyze the effect of NIPBL mutation on the three-dimensional architecture of whole genomes (Lieberman-Aiden et al., 2009).

# MICE PROJECT:

# Mice with

# MED12/R961W, MED17/L371P or MED23/R617Q

# mutation

## INTRODUCTION

As already mentioned in this manuscript, Opitz-Kaveggia syndrome, Infantile cerebral and cerebellar atrophy and non-syndromic mental retardation are caused by the mutation MED12/R961W, MED17/L371P and MED23/R617Q respectively (Risheg et al., 2007; Kaufmann et al., 2010; Hashimoto et al., 2011). Global alignment between *homo sapiens* and *mus musculus* sequences indicates that the protein MED12, MED17 and MED23 are highly conserved (data not shown). Moreover, the mutated residues are preserved across most orthologs, from *Xenopus* to human (Figure 22). We thus decided to generate, in collaboration with the Mouse Clinical Institute through GENCODYS project, homozygous mice carrying the MED12/R961W, MED17/L371P or MED23/R617Q mutation.

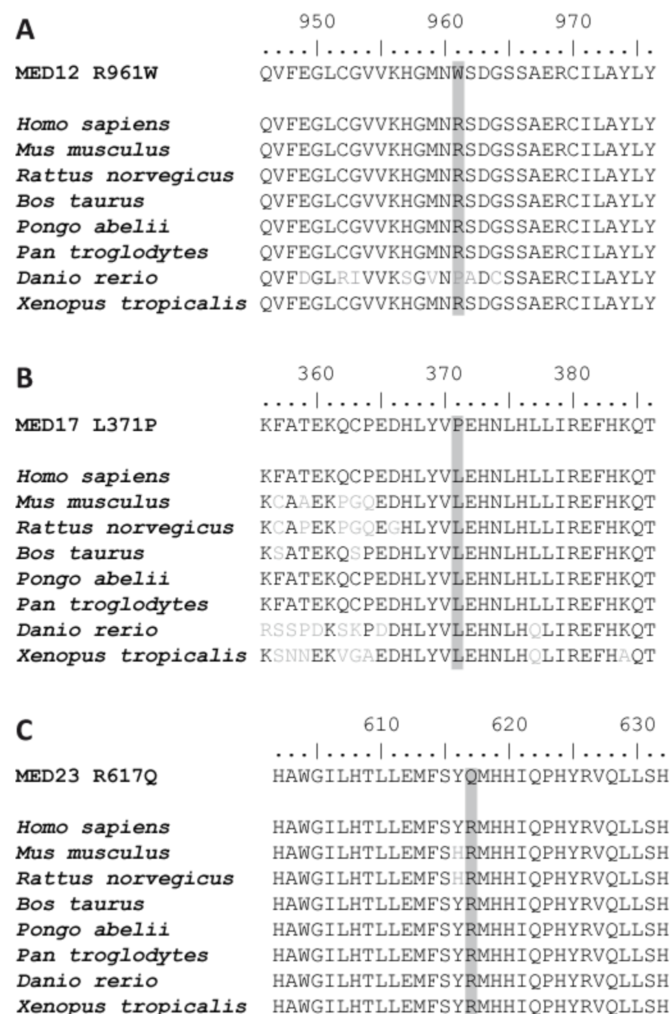



**Figure 22: Partial protein sequence alignment among multiple species around the residues MED12/R961, MED17/L371 and MED23/R617.**

## RESULTS & PERSPECTIVES

Previous studies showed that mice with an expression of Med12 drastically reduced or in a mosaic fashion, fail to develop beyond embryonic day 10 (Rocha et al., 2010a, 2010b). In our case, it was impossible to obtain chimera mice, i.e. mice mice developed from an embryo injected at blastocyst stage with embryonic stem cells carrying the mutation Med12/R961W. We thus tried to generate mice with a conditional knockout (cKO) of Med12 gene and once again with no success. All these data underline the essential role of Med12 during early mouse development.

Unlike patients, Med23/R617Q homozygous mice die during embryonic development (Figure 23). We thus created mice with a cKO of Med23 gene. These mice will be crossed with heterozygous Med23/R617Q mice to obtain $Med23^{cKO/R617Q}$ transgenic mice. This will allow us to study the effect of Med23 missense mutation in different tissues.

**Figure 23: embryos from a cross between Med23 heterozygous mice (+/R617Q).**
*Two different sizes of embryos are found from the same uterine horn after cross between Med23 heterozygous mice (+/R617Q).*

Unexpectedly, Med17/L371P homozygous mice die 6-8 weeks after birth with no reason known at this time. Monitoring of animal development (body temperature, weight, righting reflex ...etc) is currently underway. We will next analyze genes expression (Immediate Early Genes, Nuclear receptor target genes...etc) in different tissues by *in situ* hybridization, RT-qPCR, microarrays or RNAseq. Indeed, some MED subunits seem to specifically regulate certain developmental pathways (Yin and Wang, 2014). As a consequence, mutation in a MED subunit will probably not have the same effect on gene expression depending on the cells type.

We will also study the phenotype of heterozygous Med17/L371P and Med23/R206Q mice, taking into account the one developed by patients. Various behavioural tests (rotarod

test, different maze tests, fear conditioning...etc) will be performed to provide some information on the motor and learning ability of animals.

In parallel, we will also isolate mouse embryonic stem cells carrying the mutation to realize experiments of differentiation in various cell types (neurons, adipocytes, ... etc.) and thus analyze gene expression at different stages of differentiation (Martin, 1981).

# CONCLUSIONS

The ultimate goal of research on transcription is a complete understanding of gene expression regulation. By consequence, the structures, interaction networks and functions of all the proteins composing the transcription initiation complex have to be determined. Mediator (MED) was discovered two decades ago as one of many transcription factors of RNA polymerase II which played a particular role in transcription process (Flanagan et al., 1991). Since then our understanding of its importance only grew wider.

Few years ago, we decided to focus on this complex as it interacts with the general transcription and repair factor TFIIH, which has been studied in detail in our laboratory (Sakurai and Fukasawa, 1997; Akoulitchev et al., 2000; Esnault et al., 2008). We chose to work also on the cohesin loader, NIPBL, a partner of Mediator complex (Kagey et al., 2010; Muto et al., 2014). Historically, cohesin complex has mainly been studied for its role in chromosome segregation. However, during the last decade cohesin has been assigned with additional roles in DNA repair, as well as in regulation of gene expression.

This last year, an expanding list of human pathologies has been linked to genetic variations of Mediator subunits or its partners. Our comprehension of genotype-phenotype relations is still far to be complete. Nevertheless, it is important to understand the pathologies at the molecular and cellular level as it contributes to a better comprehension of biological processes, such as transcription.

The experiments detailed in this thesis gives an overview that how a subtle defect in MED subunit MED12 or MED17 or its partner NIPBL can lead to transcription dysregulation, further giving rise to diseases. It also underscores that each mutation is specific and thus give rise to different disorders although sharing some overlapping clinical features. Moreover, the use of different cellular models (fibroblasts or lymphoblastoid cells from human or mice) and comparisons between these models has provided informations which also contributed to our knowledge on gene expression. Indeed, it is becoming clear that Mediator activities are controlled differently depending on the tissues, the genes and the co-regulatory factors.

In conclusion, Mediator is a marvelous and fundamentally important protein complex. Its multifunctionality in transcription does not cease to astonish researchers throughout the last decade. And yet, new unexpected details of its cellular life are discovered and upgrade it on an even more important position among cellular machines. Despite all these new functional details, further studies are required to elucidate the etiology of phenotypes associated with mutations in Mediator complex and its partners, so that one day it will be possible to develop treatments to these disorders.

# REFERENCES

# A

Ahn SH, Kim M, Buratowski S (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3’ end processing. Mol Cell 13:67–76.

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

Al Tanoury Z, Piskunov A, Rochette-Egly C (2013) Vitamin A and retinoid signaling: genomic and nongenomic effects: Thematic Review Series: Fat-Soluble Vitamins: Vitamin A. J Lipid Res 54:1761–1775.

Anderson DE (2002) Condensin and cohesin display different arm conformations with characteristic hinge angles. J Cell Biol 156:419–424.

Ansari A, Hampsey M (2005) A role for the CPF 3’-end processing machinery in RNAP II-dependent gene looping. Genes Dev 19:2969–2978.

Ansari SA, Morse RH (2013) Mechanisms of Mediator complex action in transcriptional activation. Cell Mol Life Sci 70:2743–2756.

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

Armache K-J, Mitterweger S, Meinhart A, Cramer P (2005) Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J Biol Chem 280:7131–7134.

# B

Balamotis MA, Pennella MA, Stevens JL, Wasylyk B, Belmont AS, Berk AJ (2009) Complexity in transcription control at the activation domain-mediator interface. Sci Signal 2:ra20.

Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. Cell Res 21:381–395.

Barisic I, Tokic V, Loane M, Bianchi F, Calzolari E, Garne E, Wellesley D, Dolk H, EUROCAT Working Group (2008) Descriptive epidemiology of Cornelia de Lange syndrome in Europe. Am J Med Genet A 146A:51–59.

Bastien J, Adam-Stitah S, Riedl T, Egly JM, Chambon P, Rochette-Egly C (2000) TFIIH interacts with the retinoic acid receptor gamma and phosphorylates its AF-1-activating domain through cdk7. J Biol Chem 275:21896–21904.

Bengal E, Flores O, Krauskopf A, Reinberg D, Aloni Y (1991) Role of the mammalian transcription factors IIF, IIS, and IIX during elongation by RNA polymerase II. Mol Cell Biol 11:1195–1206.

Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ (1998) Coupling termination of transcription to messenger RNA maturation in yeast. Science 280:298–301.

Borck G, Zarhrate M, Cluzeau C, Bal E, Bonnefont J-P, Munnich A, Cormier-Daire V, Colleaux L (2006) Father-to-daughter transmission of Cornelia de Lange syndrome caused by a mutation in the 5’ untranslated region of the NIPBL Gene. Hum Mutat 27:731–735.

Bourbon H-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.

Bruck N, Vitoux D, Ferry C, Duong V, Bauer A, de Thé H, Rochette-Egly C (2009) A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. EMBO J 28:34–47.

Buckley MS, Kwak H, Zipfel WR, Lis JT (2014) Kinetics of promoter Pol II on Hsp70 reveal stable pausing and key insights into its regulation. Genes Dev 28:14–19.

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

Burke TW, Kadonaga JT (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 TW, Kadonaga JT (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.

Bushnell DA, Kornberg RD (2003) Complete, 12-subunit RNA polymerase II at 4.1-A resolution: Implications for the initiation of transcription. Proc Natl Acad Sci 100:6969–6973.

# C

Carninci P et al. (2006) Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet 38:626–635.

Chalkley GE, Verrijzer CP (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.

Cheng B, Price DH (2007) Properties of RNA Polymerase II Elongation Complexes Before and After the P-TEFb-mediated Transition into Productive Elongation. J Biol Chem 282:21901–21912.

Clements JB, Watson RJ, Wilkie NM (1977) Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. Cell 12:275–285.

Coin F, Bergmann E, Tremeau-Bravard A, Egly JM (1999) Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. EMBO J 18:1357–1366.

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

Conaway RC, Conaway JW (2013) The Mediator complex and transcription elongation. Biochim Biophys Acta BBA - Gene Regul Mech 1829:69–75.

Cook PR (2003) Nongenic transcription, gene regulation and action at a distance. J Cell Sci 116:4483–4491.

Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM (2006) Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Res 16:1–10.

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

Cramer P (2000) Architecture of RNA Polymerase II and Implications for the Transcription Mechanism. Science 288:640–649.

Cramer P (2001) Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. Science 292:1863–1876.

# D

Dahmus ME (1995) Phosphorylation of the C-terminal domain of RNA polymerase II. Biochim Biophys Acta BBA - Gene Struct Expr 1261:171–182.

Deardorff MA et al. (2012) RAD21 Mutations Cause a Human Cohesinopathy. Am J Hum Genet 90:1014–1027.

Deng W, Malecová B, Oelgeschläger T, Roberts SGE (2009) TFIIB recognition elements control the TFIIA-NC2 axis in transcriptional regulation. Mol Cell Biol 29:1389–1400.

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

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

Dilworth FJ, Chambon P (2001) Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. Oncogene 20:3047–3054.

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

Duong V, Rochette-Egly C (2011) The molecular physiology of nuclear retinoic acid receptors. From health to disease. Biochim Biophys Acta BBA - Mol Basis Dis 1812:1023–1031.

Dvir A, Conaway RC, Conaway JW (1997) A role for TFIIH in controlling the activity of early RNA polymerase II elongation complexes. Proc Natl Acad Sci U S A 94:9006–9010.

# E

El Kaderi B, Medler S, Raghunayakula S, Ansari A (2009) Gene Looping Is Conferred by Activator-dependent Interaction of Transcription Initiation and Termination Machineries. J Biol Chem 284:25015–25025.

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

Evans R (2001) Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. Genes Dev 15:2945–2949.

# F

Fay A, Misulovin Z, Li J, Schaaf CA, Gause M, Gilmour DS, Dorsett D (2011) Cohesin Selectively Binds and Regulates Genes with Paused RNA Polymerase. Curr Biol 21:1624–1634.

Flanagan PM, Kelleher RJ, Sayre MH, Tschochner H, Kornberg RD (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 ZF, 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 JD, Ge H, Roeder RG (1996) Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. Proc Natl Acad Sci U S A 93:8329–8333.

# G

Gaillard E, Bruck N, Brelivet Y, Bour G, Lalevée S, Bauer A, Poch O, Moras D, Rochette-Egly C (2006) Phosphorylation by PKA potentiates retinoic acid receptor alpha activity by means of increasing interaction with and phosphorylation by cyclin H/cdk7. Proc Natl Acad Sci U S A 103:9548–9553.

Gard S, Light W, Xiong B, Bose T, McNairn AJ, Harris B, Fleharty B, Seidel C, Brickner JH, Gerton JL (2009) Cohesinopathy mutations disrupt the subnuclear organization of chromatin. J Cell Biol 187:455–462.

Ge H, Roeder RG (1994) The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with the TATA-binding protein. J Biol Chem 269:17136–17140.

Geng F, Wenzel S, Tansey WP (2012) Ubiquitin and Proteasomes in Transcription. Annu Rev Biochem 81:177–201.

Gershenzon NI, Ioshikhes IP (2005) Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinforma Oxf Engl 21:1295–1300.

Ghazy MA, Brodie SA, Ammerman ML, Ziegler LM, Ponticelli AS (2004) Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. Mol Cell Biol 24:10975–10985.

Gillis L, Mccallum J, Kaur M, Descipio C, Yaeger D, Mariani A, Kline A, Li H, Devoto M, Jackson L (2004) NIPBL Mutational Analysis in 120 Individuals with Cornelia de Lange Syndrome and Evaluation of Genotype-Phenotype Correlations. Am J Hum Genet 75:610–623.

Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science 292:1876–1882.

Goldman SR, Ebright RH, Nickels BE (2009) Direct Detection of Abortive RNA Transcripts in Vivo. Science 324:927–928.

Goldsborough AS, Kornberg TB (1996) Reduction of transcription by homologue asynapsis in Drosophila imaginal discs. Nature 381:807–810.

Goltzman D, Hendy GN, White JH (2014) Vitamin D and its receptor during late development. Biochim Biophys Acta.

Gonzalez-Parra E, Rojas-Rivera J, Tuñón J, Praga M, Ortiz A, Egido J (2012) Vitamin D receptor activation and cardiovascular disease. Nephrol Dial Transplant Off Publ Eur Dial Transpl Assoc - Eur Ren Assoc 27 Suppl 4:iv17–iv21.

Goodrich JA, Tjian R (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell 77:145–156.

Gruber S, Haering CH, Nasmyth K (2003) Chromosomal Cohesin Forms a Ring. Cell 112:765–777.

Gullerova M, Proudfoot NJ (2008) Cohesin Complex Promotes Transcriptional Termination between Convergent Genes in S. pombe. Cell 132:983–995.

# H

Haering CH, Löwe J, Hochwagen A, Nasmyth K (2002) Molecular Architecture of SMC Proteins and the Yeast Cohesin Complex. Mol Cell 9:773–788.

Ha I, Roberts S, Maldonado E, Sun X, Kim LU, Green M, 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.

Hampsey M (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiol Mol Biol Rev MMBR 62:465–503.

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

Henriques T, Gilchrist DA, Nechaev S, Bern M, Muse GW, Burkholder A, Fargo DC, Adelman K (2013) Stable Pausing by RNA Polymerase II Provides an Opportunity to Target and Integrate Regulatory Signals. Mol Cell 52:517–528.

Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. Science 308:118–120.

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

Holstege FC., Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA (1998) Dissecting the Regulatory Circuitry of a Eukaryotic Genome. Cell 95:717–728.

Holstege FC, van der Vliet PC, Timmers HT (1996) Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. EMBO J 15:1666–1677.

# I

Imasaki T, Calero G, Cai G, Tsai K-L, Yamada K, Cardelli F, Erdjument-Bromage H, Tempst P, Berger I, Kornberg GL, Asturias FJ, Kornberg RD, Takagi Y (2011) Architecture of the Mediator head module. Nature 475:240–243.

# J

Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST (1994) DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol 14:116–127.

Juven-Gershon T, Kadonaga JT (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol 339:225–229.

# K

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA (2010) Mediator and cohesin connect gene expression and chromatin architecture. Nature 467:430–435.

Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJM (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet 37:761–765.

Kaufmann R, Straussberg R, Mandel H, Fattal-Valevski A, Ben-Zeev B, Naamati A, Shaag A, Zenvirt S, Konen O, Mimouni-Bloch A, Dobyns WB, Edvardson S, Pines O, Elpeleg O (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.

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, Nussenzweig A, Kitzes L, Yokomori K, Hallgrimsson B, Lander AD (2009) Multiple Organ System Defects and Transcriptional Dysregulation in the Nipbl+/− Mouse, a Model of Cornelia de Lange Syndrome Bickmore WA, ed. PLoS Genet 5:e1000650.

Killeen M, Coulombe B, 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 TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B (2005) A high-resolution map of active promoters in the human genome. Nature 436:876–880.

Kim YJ, Björklund S, Li Y, Sayre MH, Kornberg RD (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell 77:599–608.

Komarnitsky P (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev 14:2452–2460.

Kumar KP, Akoulitchev S, Reinberg D (1998) Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event. Proc Natl Acad Sci U S A 95:9767–9772.

Kutach AK, Kadonaga JT (2000) The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. Mol Cell Biol 20:4754–4764.

# L

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

Lavigne AC, Mengus G, Gangloff YG, Wurtz JM, 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.

Lee D-H, Gershenzon N, Gupta M, Ioshikhes IP, Reinberg D, Lewis BA (2005) Functional Characterization of Core Promoter Elements: the Downstream Core Element Is Recognized by TAF1. Mol Cell Biol 25:9674–9686.

Lee DK, DeJong J, Hashimoto S, Horikoshi M, Roeder RG (1992) TFIIA induces conformational changes in TFIID via interactions with the basic repeat. Mol Cell Biol 12:5189–5196.

Leid M, Kastner P, Chambon P (1992) Multiplicity generates diversity in the retinoic acid signalling pathways. Trends Biochem Sci 17:427–433.

Lei L, Ren D, Burton ZF (1999) The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. Mol Cell Biol 19:8372–8382.

Le May N, Fradin D, Iltis I, Bougnères P, 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 SA, Ropers H-H, Kalscheuer VM, Des Portes V, Labalme A, Sanlaville D, Edery P, Raynaud M, Lespinasse J (2013) Clinical and neurocognitive characterization of a family with a novel *MED12* gene frameshift mutation. Am J Med Genet A 161:3063–3071.

Li B, Carey M, Workman JL (2007) The Role of Chromatin during Transcription. Cell 128:707–719.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 326:289–293.

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

Liu J, Krantz ID (2009) Cornelia de Lange syndrome, cohesin, and beyond. Clin Genet 76:303–314.

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID (2009) Transcriptional Dysregulation in NIPBL and Cohesin Mutant Human Cells Hastie N, ed. PLoS Biol 7:e1000119.

Li Y, Flanagan PM, Tschochner H, Kornberg RD (1994) RNA polymerase II initiation factor interactions and transcription start site selection. Science 263:805–807.

Lu H, Zawel L, Fisher L, Egly J-M, Reinberg D (1992) Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature 358:641–645.

# M

MacLaughlin JA, Anderson RR, Holick MF (1982) Spectral character of sunlight modulates photosynthesis of previtamin D3 and its photoisomers in human skin. Science 216:1001–1003.

Maldonado E, Ha I, Cortes P, Weis L, 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.

Maldonado E, Shiekhattar R, Sheldon M, Cho H, Drapkin R, Rickert P, Lees E, Anderson CW, Linn S, Reinberg D (1996) A human RNA polymerase II complex associated with SRB and DNA-repair proteins. Nature 381:86–89.

Malik S, Gu W, Wu W, Qin J, Roeder RG (2000) The USA-Derived Transcriptional Coactivator PC2 Is a Submodule of TRAP/SMCC and Acts Synergistically with Other PCs. Mol Cell 5:753–760.

Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM (1995) The nuclear receptor superfamily: The second decade. Cell 83:835–839.

Marshall NF, Peng J, Xie Z, Price DH (1996) Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem 271:27176–27183.

Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 78:7634–7638.

Matsui T, Segall J, Weil PA, Roeder RG (1980) Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. J Biol Chem 255:11992–11996.

Maxon ME, Goodrich JA, 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.

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature 385:357–361.

Merino A, Madden KR, Lane WS, Champoux JJ, Reinberg D (1993) DNA topoisomerase I is involved in both repression and activation of transcription. Nature 365:227–232.

Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K (2001) Bacterial RNA polymerase subunit and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci 98:892–897.

Misulovin Z, Schwartz YB, Li X-Y, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D (2008) Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. Chromosoma 117:89–102.

Mizzen CA, Yang X-J, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, Nakatani Y, Allis CD (1996) The TAFII250 Subunit of TFIID Has Histone Acetyltransferase Activity. Cell 87:1261–1270.

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

Muto A, Ikeda S, Lopez-Burks ME, Kikuchi Y, Calof AL, Lander AD, Schilling TF (2014) Nipbl and mediator cooperatively regulate gene expression to control limb development. PLoS Genet 10:e1004671.

# N

Näär AM, Beaurang PA, Zhou S, Abraham S, Solomon W, Tjian R (1999) Composite co-activator ARC mediates chromatin-directed transcriptional activation. Nature 398:828–832.

Nag A, Narsinh K, Martinson HG (2007) The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat Struct 38 Mol Biol 14:662–669.

Nonet ML, Young RA (1989) Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of Saccharomyces cerevisiae RNA polymerase II. Genetics 123:715–724.

# O

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

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

Olefsky JM (2001) Nuclear Receptor Minireview Series. J Biol Chem 276:36863–36864.

Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS (2005) Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell 120:613–622.

Ossipow V, Tassan JP, Nigg EA, Schibler U (1995) A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. Cell 83:137–146.

O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, Mellor J, Proudfoot NJ (2004) Gene loops juxtapose promoters and terminators in yeast. Nat Genet 36:1014–1018.

# P

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

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

Petersen B, Petersen T, Andersen P, Nielsen M, Lundegaard C (2009) A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Struct Biol 9:51.

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

Poss ZC, Ebmeier CC, Taatjes DJ (2013) The Mediator complex and transcription regulation. Crit Rev Biochem Mol Biol 48:575–608.

# R

Rachez C, Suldan Z, Ward J, Chang CP, Burakov D, Erdjument-Bromage H, Tempst P, Freedman LP (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.

Rajesh D, Dickerson SJ, Yu J, Brown ME, Thomson JA, Seay NJ (2011) Human lymphoblastoid B-cell lines reprogrammed to EBV-free induced pluripotent stem cells. Blood 118:1797–1800.

Rhinn M, Dolle P (2012) Retinoic acid signalling during development. Development 139:843–858.

Risheg H, Graham JM, Clark RD, Rogers RC, Opitz JM, Moeschler JB, Peiffer AP, May M, Joseph SM, Jones JR, Stevenson RE, Schwartz CE, Friez MJ (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. Nat Genet 39:451–453.

Rocha PP, Bleiss W, Schrewe H (2010a) Mosaic expression of Med12 in female mice leads to exencephaly, spina bifida, and craniorachischisis. Birt Defects Res A Clin Mol Teratol 88:626–632.

Rocha PP, Scholze M, Bleiss W, Schrewe H (2010b) Med12 is essential for early mouse development and for canonical Wnt and Wnt/PCP signaling. Development 137:2723–2731.

Rochette-Egly C, Adam S, Rossignol M, Egly J-M, Chambon P (1997) Stimulation of RARα Activation Function AF-1 through Binding to the General Transcription Factor TFIIH and Phosphorylation by CDK7. Cell 90:97–107.

Roeder RG, Rutter WJ (1969) Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms. Nature 224:234–237.

Rollins RA, Korom M, Aulner N, Martens A, Dorsett D (2004) Drosophila Nipped-B Protein Supports Sister Chromatid Cohesion and Opposes the Stromalin/Scc3 Cohesion Factor To Facilitate Long-Range Activation of the cut Gene. Mol Cell Biol 24:3100–3111.

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

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

# S

Saiz L, Rubi JM, Vilar JMG (2005) Inferring the in vivo looping properties of DNA. Proc Natl Acad Sci 102:17642–17645.

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

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

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

Schaaf CA, Kwak H, Koenig A, Misulovin Z, Gohara DW, Watson A, Zhou Y, Lis JT, Dorsett D (2013) Genome-Wide Control of RNA Polymerase II Activity by Cohesin Ren B, ed. PLoS Genet 9:e1003382.

Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JH, Chambon P, Egly JM (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science 260:58–63.

Schwartz CE 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.

Seitan VC, Banks P, Laval S, Majid NA, Dorsett D, Rana A, Smith J, Bateman A, Krpic S, Hostert A, Rollins RA, Erdjument-Bromage H, Tempst P, Benard CY, Hekimi S, Newbury SF, Strachan T (2006) Metazoan Scc4 homologs link sister chromatid cohesion to cell and axon migration guidance. PLoS Biol 4:e242.

Shandilya J, Roberts SGE (2012) The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. Biochim Biophys Acta BBA - Gene Regul Mech 1819:391–400.

Shykind BM, Kim J, Sharp PA (1995) Activation of the TFIID-TFIIA complex with HMG-2. Genes Dev 9:1354–1365.

Singh BN, Hampsey M (2007) A transcription-independent role for TFIIB in gene looping. Mol Cell 27:806–816.

Sommer A (2008) Vitamin a deficiency and clinical disease: an historical overview. J Nutr 138:1835–1839.

Spaeth JM, Kim NH, Boyer TG (2011) Mediator and human disease. Semin Cell Dev Biol 22:776–787.

Sporn MB, Dunlop NM, Newton DL, Smith JM (1976) Prevention of chemical carcinogenesis by vitamin A and its synthetic analogs (retinoids). Fed Proc 35:1332–1338.

Steinmetz EJ, Conrad NK, Brow DA, Corden JL (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3’-end formation of RNA polymerase II transcripts. Nature 413:327–331.

Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell 24:735–746.

Sun X, Zhang Y, Cho H, Rickert P, Lees E, Lane W, Reinberg D (1998) NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. Mol Cell 2:213–222.

# T

Takagi Y, Calero G, Komori H, Brown JA, Ehrensberger AH, Hudmon A, Asturias F, Kornberg RD (2006) Head Module Control of Mediator Interactions. Mol Cell 23:355–364.

Tan Q, Linask KL, Ebright RH, Woychik NA (2000) Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. Genes Dev 14:339–348.

Thomas MC, Chiang C-M (2006) The General Transcription Machinery and General Cofactors. Crit Rev Biochem Mol Biol 41:105–178.

Tirode F, Busso D, Coin F, Egly JM (1999) Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. Mol Cell 3:87–95.

Tokusumi Y, Ma Y, Song X, Jacobson RH, 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.

Tsai K-L, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Asturias FJ (2014) Subunit Architecture and Functional Modular Rearrangements of the Transcriptional Mediator Complex. Cell 157:1430–1444.

# V

Vulto-van Silfhout AT, de Vries BBA, van Bon BWM, Hoischen A, Ruiterkamp-Versteeg M, Gilissen C, Gao F, van Zwam M, Harteveld CL, van Essen AJ, Hamel BCJ, Kleefstra T, Willemsen MAAP, Yntema HG, van Bokhoven H, Brunner HG, Boyer TG, de Brouwer APM (2013) Mutations in MED12 Cause X-Linked Ohdo Syndrome. Am J Hum Genet 92:401–406.

# W

Wang Y, Fairley JA, Roberts SGE (2010) Phosphorylation of TFIIB links transcription initiation and termination. Curr Biol CB 20:548–553.

Watanabe T, Hayashi K, Tanaka A, Furumoto T, Hanaoka F, 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.

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters J-M (2006) Human Scc4 Is Required for Cohesin Binding to Chromatin, Sister-Chromatid Cohesion, and Mitotic Progression. Curr Biol 16:863–874.

Weil PA, Luse DS, Segall J, Roeder RG (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.

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell 153:307–319.

Woychik NA, Young RA (1990) RNA polymerase II: subunit structure and function. Trends Biochem Sci 15:347–351.

# Y

Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell 97:41–51.

Yan Q, Moreland RJ, Conaway JW, Conaway RC (1999) Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. J Biol Chem 274:35668–35675.

Yin J -w., Wang G (2014) The Mediator complex: a master coordinator of transcription and cell lineage development. Development 141:977–987.

Young RA (1991) RNA Polymerase II. Annu Rev Biochem 60:689–715.

Yudkovsky N, Ranish JA, Hahn S (2000) A transcription reinitiation intermediate that is stabilized by activator. Nature 408:225–229.

# Z

Zehring WA, Lee JM, Weeks JR, Jokerst RS, Greenleaf AL (1988) The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro. Proc Natl Acad Sci U S A 85:3698–3702.

Zhang C, Burton ZF (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.

Zhou Q, Li T, Price DH (2012) RNA polymerase II elongation control. Annu Rev Biochem 81:119–143.

Zhu X, Zhang Y, Bjornsdottir G, Liu Z, Quan A, Costanzo M, Davila Lopez M, Westholm JO, Ronne H, Boone C, Gustafsson CM, Myers LC (2011) Histone modifications influence mediator interactions with chromatin. Nucleic Acids Res 39:8342–8354.

Zuin J, Franke V, van IJcken WFJ, van der Sloot A, Krantz ID, van der Reijden MIJA, Nakato R, Lenhard B, Wendt KS (2014) A Cohesin-Independent Role for NIPBL at Promoters Provides Insights in CdLS Barsh GS, ed. PLoS Genet 10:e1004153.

UNIVERSITÉ DE STRASBOURG

igbmc institut de génétique et de biologie moléculaire et cellulaire

Lise-Marie DONNIO

# Étude transcriptionnelle des mutations dans le Médiateur ou dans son partenaire NIPBL à l'origine de maladies génétiques

## Résumé en Français

Le Médiateur (MED) est un complexe multi-protéique dont le principal rôle est de transmettre les différents signaux fournis par les facteurs fixés sur des séquences d'ADN spécifique à la machinerie transcriptionnelle de base, permettant ainsi une régulation fine de l'expression des gènes. Des mutations dans le MED ou ses partenaires, comme NIPBL, sont à l'origine de diverses maladies telles que des malformations congénitales, des troubles neurodéveloppementaux ou des cancers. A partir de cellules provenant de patients portant différentes mutations dans les sous-unités MED12 ou MED17 du MED ou dans NIPBL, nous avons observé une altération du niveau d'expression de certains gènes qui dépend de la localisation de la mutation et de la nature de leur activation. Ces variations de l'expression des gènes sont la conséquence d'un défaut dans la formation du complexe de transcription et du remodelage de la chromatine (modifications post-traductionnelles des histones). Outre une meilleure appréhension du rôle des sous-unités MED12 et MED17 du MED ainsi que NIPBL, sur la transcription des gènes, ma thèse a permis de mieux comprendre l'étiologies des maladies associées à une mutation dans ces protéines.

**Mots-clé :** le complexe Médiateur, Maladies génétiques, Transcription, Déficiences intellectuelles lié à l'X, MED12, MED17, NIPBL.

## English abstract

Mediator (MED) is a multi-protein complex whose main role is to convey the different signals from factors bound at specific DNA sequences to basal transcriptional machinery, allowing thus a fine regulation of gene expression. Mutations in MED or its partners, like NIPBL, cause various diseases, such as congenital malformations, neurodevelopmental disorders or cancers. Using cells from patients carrying different mutations in the MED subunits MED12 or MED17 or in NIPBL, we observed an alteration of the expression of studied genes which depend on the position of the mutation and on the nature of the activation. These variations of gene expression are the consequence of a defect in transcription complex formation, as well as in chromatin remodeling (histones post-translational modifications). In addition to better comprehend the role of the MED subunits MED12 and MED17, and of NIPBL on gene transcription, my thesis helped to better understand the ethiology of the disorders associated with mutations in these proteins.

**Key words:** Mediator complex, genetic disorders, Transcription, X-linked intellectual disabilities, MED12, MED17, NIPBL.