



HAL
open science

Clinico-biological refinement of BCL11B related disorder and identification of a genome-wide dna methylation signature

Quentin Sabbagh

► **To cite this version:**

Quentin Sabbagh. Clinico-biological refinement of BCL11B related disorder and identification of a genome-wide dna methylation signature. Human health and pathology. 2023. dumas-04550729

HAL Id: dumas-04550729

<https://dumas.ccsd.cnrs.fr/dumas-04550729v1>

Submitted on 18 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

UNIVERSITE DE MONTPELLIER
FACULTE DE MEDECINE MONTPELLIER-NIMES

THESE

Pour obtenir le titre de
DOCTEUR EN MEDECINE

Présentée et soutenue publiquement

Par

Quentin SABBAGH

le 27 Octobre 2023

**CLINICO-BIOLOGICAL REFINEMENT OF *BCL11B*-
RELATED DISORDER AND IDENTIFICATION OF A
GENOME-WIDE DNA METHYLATION SIGNATURE**

Directeur : Professeur David GENEVIEVE

JURY

Président : Professeur David GENEVIEVE

Assesseurs : Professeur Pascal PUJOL

Professeur Franck PELLESTOR

Membre invité : Docteur Marjolaine WILLEMS

UNIVERSITE DE MONTPELLIER
FACULTE DE MEDECINE MONTPELLIER-NIMES

THESE

Pour obtenir le titre de
DOCTEUR EN MEDECINE

Présentée et soutenue publiquement

Par

Quentin SABBAGH

le 27 Octobre 2023

**CLINICO-BIOLOGICAL REFINEMENT OF *BCL11B*-
RELATED DISORDER AND IDENTIFICATION OF A
GENOME-WIDE DNA METHYLATION SIGNATURE**

Directeur : Professeur David GENEVIEVE

JURY

Président : Professeur David GENEVIEVE

Assesseurs : Professeur Pascal PUJOL

Professeur Franck PELLESTOR

Membre invité : Docteur Marjolaine WILLEMS



LISTE DU CORPS ENSEIGNANT

ANNÉE UNIVERSITAIRE 2022-2023

Professeurs Honoraires

ALBAT Bernard	BRUNEL Michel	HERTAULT Jean	NAVARRO Maurice
ALLIEU Yves	CAMU William	HUMEAU Claude	NAVRATIL Henri
ALRIC Robert	CANAUD Bernard	JAFFIOL Claude	OTHONIEL Jacques
ARNAUD Bernard	CHAPTAL Paul-André	JANBON Charles	PAGES Michel
ASENCIO Gérard	CIURANA Albert-Jean	JANBON François	PEGURET Claude
ASTRUC Jacques	CLOT Jacques	JARRY Daniel	PELISSIER Jacques
AUSSILLOUX Charles	COSTA Pierre	JOURDAN Jacques	PETIT Pierre
AVEROUS Michel	COTTALORDA Jérôme	KLEIN Bernard	POUGET Régis
AYRAL Guy	D'ATHIS Françoise	LAFFARGUE François	PUJOL Henri
BAILLAT Xavier	DEMAILLE Jacques	LALLEMANT Jean Gabriel	RABISCHONG Pierre
BALDET Pierre	DESCOMPS Bernard	LAMARQUE Jean-Louis	RAMUZ Michel
BALDY-MOULINIER Michel	DIMEGLIO Alain	LAPEYRIE Henri	REBOUL Jean
BALMES Jean-Louis	DUBOIS Jean Bernard	LEROUX Jean-Louis	RIEU Daniel
BANSARD Nicole	DUJOLS Pierre	LESBROS Daniel	ROCHFORT Henri
BILLIARD Michel	DUMAZER Romain	LOPEZ François Michel	ROUANET DE VIGNE Jean Pierre
BLARD Jean-Marie	ECHENNE Bernard	LORIOT Jean	SAINT-AUBERT Bernard
BLAYAC Jean Pierre	FABRE Serge	LOUBATIERES Marie Madeleine	SANCHO-GARNIER Hélène
BLOTMAN Francis	FREREBEAU Philippe	MAGNAN DE BORNIER Bernard	SANY Jacques
BONNEL François	GALIFER René Benoît	MARTY ANE Charles	SEGNARBIEUX François
BOURGEOIS Jean-Marie	GODLEWSKI Guilhem	MATHIEU-DAUDE Pierre	SENAC Jean-Paul
BOUSQUET Jean	GRASSET Daniel	MEYNADIER Jean	SERRE Arlette
BRUEL Jean Michel	GUILHOU Jean-Jacques	MICHEL François-Bernard	TOUITOU Isabelle
BUREAU Jean-Paul	GUI TER Pierre	MION Charles	VIDAL Jacques
	HEDON Bernard	MIRO Luis	VISI ER Jean Pierre



Professeurs Emérites

ARTUS Jean-Claude	LANDAIS Paul
BASTIEN Patrick	LARREY Dominique
BLANC François	LE QUELLEC Alain
BONAFE Alain	MARES Pierre
BOULENGER Jean-Philippe	MAUDELONDE Thierry
BOURREL Gérard	MAURY Michèle
BRINGER Jacques	MESSNER Patrick
CLAUSTRES Mireille	MILLAT Bertrand
COMBE Bernard	MONNIER Louis
DAURES Jean-Pierre	MOURAD Georges
DAUZAT Michel	PREFAUT Christian
DAVY Jean-Marc	PUJOL Rémy
DEDET Jean-Pierre	RIBSTEIN Jean
DE LA COUSSAYE Jean-Emmanuel	SCHVED Jean-François
GROLLEAU RAOUX Robert	SULTAN Charles
GUERRIER Bernard	TOUCHON Jacques
GUILLOT Bernard	UZIEL Alain
HERISSON Christian	VOISIN Michel
JONQUET Olivier	ZANCA Michel

Docteurs Emérites

PRAT Dominique
PUJOL Joseph



Professeurs des Universités - Praticiens Hospitaliers

Classe exceptionnelle

AGUILAR MARTINEZ Patricia	Hématologie ; transfusion
ALRIC Pierre	Chirurgie vasculaire ; médecine vasculaire (option chirurgie vasculaire)
AVIGNON Antoine	Nutrition
AZRIA David	Cancérologie ; radiothérapie
BACCINO Eric	Médecine légale et droit de la santé
BEREGI Jean-Paul	Radiologie et imagerie médicale
BLAIN Hubert	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
BOULOT Pierre	Gynécologie-obstétrique ; gynécologie médicale
CAMBONIE Gilles	Pédiatrie
CANOVAS François	Anatomie
CAPDEVILA Xavier	Anesthésiologie-réanimation et médecine péri-opératoire
CHAMMAS Michel	Chirurgie orthopédique et traumatologique
COLSON Pascal	Anesthésiologie-réanimation et médecine péri-opératoire
COSTES Valérie	Anatomie et cytologie pathologiques
COUBES Philippe	Neurochirurgie
COURTET Philippe	Psychiatrie d'adultes ; addictologie
CRAMPETTE Louis	Oto-rhino-laryngologie
CRISTOL Jean-Paul	Biochimie et biologie moléculaire
CYTEVAL Catherine	Radiologie et imagerie médicale
DAUVILLIERS Yves	Physiologie
DE TAYRAC Renaud	Gynécologie-obstétrique ; gynécologie médicale
DE WAZIERES Benoît	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
DELAPORTE Eric	Maladies infectieuses ; maladies tropicales
DEMARIA Roland	Chirurgie thoracique et cardiovasculaire
DEMOLY Pascal	Pneumologie ; addictologie
DEREURE Olivier	Dermatologie – vénéréologie
DOMERGUE Jacques	Chirurgie viscérale et digestive
DUFFAU Hugues	Neurochirurgie
ELIAOU Jean-François	Immunologie
FABRE Jean-Michel	Chirurgie viscérale et digestive
HAMAMAH Samir	Biologie et médecine du développement et de la reproduction
JABER Samir	Anesthésiologie-réanimation et médecine péri-opératoire



JEANDEL Claire	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
JORGENSEN Christian	Thérapeutique ; médecine d'urgence ; addictologie
KOENING Michel	Génétique
KOTZKI Pierre-Olivier	Biophysique et médecine nucléaire
LABAUGE Pierre	Neurologie
LAFFONT Isabelle	Médecine physique et de réadaptation
LECLERCQ Florence	Cardiologie
LEFRANT Jean-Yves	Anesthésiologie-réanimation et médecine péri-opératoire
LEHMANN Sylvain	Biochimie et biologie moléculaire
LUMBROSO Serge	Biochimie et biologie moléculaire
MARIANO-GOULART Denis	Biophysique et médecine nucléaire
MERCIER Jacques	Physiologie
MEUNIER Laurent	Dermatologie – vénéréologie
MONDAIN Michel	Oto-rhino-laryngologie
MORIN Denis	Pédiatrie
PAGEAUX Georges-Philippe	Gastroentérologie ; hépatologie ; addictologie
PRUDHOMME Michel	Addictologie
PUJOL Pascal	Biologie cellulaire
QUERE Isabelle	Chirurgie vasculaire ; médecine vasculaire (option chirurgie vasculaire)
RENARD Eric	Endocrinologie, diabète et maladies métaboliques ; gynécologie médicale
REYNES Jacques	Maladies infectieuses ; maladies tropicales
RIPART Jacques	Anesthésiologie-réanimation et médecine péri-opératoire
ROUANET Philippe	Cancérologie ; radiothérapie
SOTTO Albert	Maladies infectieuses ; maladies tropicales
TAOUREL Patrice	Radiologie et imagerie médicale
TRAN Tu-Anh	Pédiatrie
VANDE PERRE Philippe	Bactériologie-virologie ; hygiène hospitalière
VERNHET Hélène	Radiologie et imagerie médicale
YCHOU Marc	Cancérologie ; radiothérapie

Première classe

ASSENAT Eric	Gastroentérologie ; hépatologie ; addictologie
BAGHDADLI Amaria	Pédopsychiatrie ; addictologie
BLANC Pierre	Gastroentérologie ; hépatologie ; addictologie
BORIE Frédéric	Chirurgie viscérale et digestive
BOURDIN Arnaud	Pneumologie ; addictologie



BOURGIER Céline	Cancérologie ; radiothérapie
CAPTIER Guillaume	Anatomie
CARTRON Guillaume	Hématologie ; transfusion
CAYLA Guillaume	Cardiologie
CHANQUES Gérard	Anesthésiologie-réanimation et médecine péri-opératoire
CORBEAU Pierre	Immunologie
COULET Bertrand	Chirurgie orthopédique et traumatologique
CUVILLON Philippe	Anesthésiologie-réanimation et médecine péri-opératoire
DADURE Christophe	Anesthésiologie-réanimation et médecine péri-opératoire
DE VOS John	Histologie, embryologie et cytogénétique
DROUPY Stéphane	Urologie
DUCROS Anne	Neurologie
DUPEYRON Arnaud	Médecine physique et de réadaptation
FESLER Pierre	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
GARREL Renaud	Oto-rhino-laryngologie
GENEVIEVE David	Génétique
GUILLAUME Sébastien	Psychiatrie d'adultes ; addictologie
GUILPAIN Philippe	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
GUIU Boris	Radiologie et imagerie médicale
HAYOT Maurice	Physiologie
HOUEDE Nadine	Cancérologie ; radiothérapie
JACOT William	Cancérologie ; radiothérapie
JUNG Boris	Médecine intensive-réanimation
KALFA Nicolas	Chirurgie infantile
KLOUCHE Kada	Médecine intensive-réanimation
KOUYOUMIDJIAN Pascal	Chirurgie orthopédique et traumatologique
LACHAUD Laurence	Parasitologie et mycologie
LAVABRE-BERTRAND Thierry	Histologie, embryologie et cytogénétique
LAVIGNE Jean-Philippe	Bactériologie-virologie ; hygiène hospitalière
LE MOING Vincent	Maladies infectieuses ; maladies tropicales
LETOUZEY Vincent	Gynécologie-obstétrique ; gynécologie médicale
MATECKI Stefan	Physiologie
MORANNE Olivier	Néphrologie
MOREL Jacques	Rhumatologie
NAGOT Nicolas	Biostatistiques et informatique médicale
NAVARRO Francis	Chirurgie viscérale et digestive
NOCCA David	Chirurgie viscérale et digestive



PANARO Fabrizio	Chirurgie viscérale et digestive
PASQUIE Jean-Luc	Cardiologie
PEREZ MARTIN Antonia	Chirurgie vasculaire ; médecine vasculaire (option médecine vasculaire)
PERNEY Pascal	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
PUJOL Jean-Louis	Pneumologie ; addictologie
PURPER OUAKIL Diane	Pédopsychiatrie ; addictologie
RIGAU Valérie	Anatomie et cytologie pathologiques
ROGER Pascal	Anatomie et cytologie pathologiques
ROUBILLE François	Cardiologie
SOLASSOL Jérôme	Biologie cellulaire
SULTAN Ariane	Nutrition
VINCENT Thierry	Immunologie

Deuxième classe

AYRIGNAC Xavier	Neurologie
BELZEAUX Raoul	Psychiatrie d'adultes ; addictologie
BERTRAND Martin	Anatomie
BOBBIA Xavier	Médecine d'urgence
CANAUD Ludovic	Chirurgie vasculaire ; médecine vasculaire (option chirurgie vasculaire)
CAPDEVIELLE Delphine	Psychiatrie d'adultes ; addictologie
CLARET Pierre-Géraud	Médecine d'urgence
COLOMBO Pierre-Emmanuel	Cancérologie ; radiothérapie
COSTALAT Vincent	Radiologie et imagerie médicale
DAIEN Vincent	Ophthalmologie
DE JONG Audrey	Anesthésiologie-réanimation et médecine péri-opératoire
DONNADIEU Hélène	Gastroentérologie ; hépatologie ; addictologie
DORANDEU Anne	Médecine légale et droit de la santé
FAILLIE Jean-Luc	Pharmacologie fondamentale ; pharmacologie clinique
FRANDON Julien	Radiologie et imagerie médicale
FUCHS Florent	Gynécologie-obstétrique ; gynécologie médicale
GABELLE DELOUSTAL Audrey	Neurologie
GANDET Thomas	Chirurgie thoracique et cardiovasculaire
GAUJOUX Viala Céline	Rhumatologie
GELIS Anthony	Médecine physique et de réadaptation
GODREUIL Sylvain	Bactériologie-virologie ; hygiène hospitalière
IMMEDIATO DAEIN Claire	Rhumatologie
JEZIORSKI Eric	Pédiatrie



LALLEMANT Benjamin	Oto-rhino-laryngologie
LATTUCA Benoît	Cardiologie
LE QUINTREC DONNETTE Moglie	Néphrologie
LONJON Nicolas	Neurochirurgie
LOPEZ CASTROMAN Jorge	Psychiatrie d'adultes ; addictologie
LUKAS Cédric	Rhumatologie
MAKINSON Alain	Maladies infectieuses ; maladies tropicales
MENJOT DE CHAMPFLEUR Nicolas	Radiologie et imagerie médicale
MILLET Ingrid	Radiologie et imagerie médicale
MURA Thibault	Biostatistiques et informatique médicale
OLIE Emilie	Psychiatrie d'adultes : addictologie
PARIS Françoise	Biologie et médecine du développement et de la reproduction
PELLESTOR Franck	Histologie, embryologie et cytogénétique
PERS Yves-Marie	Thérapeutique ; addictologie
POUDEROUX Philippe	Gastroentérologie ; hépatologie ; addictologie
RIVIER François	Pédiatrie
ROGER Claire	Anesthésiologie-réanimation et médecine péri-opératoire
ROSSI Jean-François	Hématologie ; transfusion
SEBANNE Mustapha	Médecine d'urgence
SIRVENT Nicolas	Pédiatrie
STERKERS Yvon	Parasitologie et mycologie
STOEBNER Pierre	Dermatologie – vénéréologie
THOUVENOT Eric	Neurologie
THURET Rodolphe	Urologie
TUAILLON Edouard	Bactériologie-virologie ; hygiène hospitalière
VENAIL Frédéric	Oto-rhino-laryngologie
VILLAIN Max	Ophtalmologie
VINCENT Denis	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale



Professeurs des Universités

Première classe

COLINGE Jacques	Cancérologie, signalisation cellulaire et systèmes complexes
LAOUDJ CHENIVASSE Dalila	Biochimie et biologie moléculaire
VISIER Laurent	Sociologie et démographie
AMOUYAL Michel	Médecine générale

Deuxième classe

CHAZAL Nathalie	Biologie cellulaire
MOREAUX Jérôme	Sciences biologiques fondamentales et cliniques

Professeurs Associés

BESSIS Didier	Dermatologie – vénéréologie
DEBIEN Blaise	Médecine d'urgence
MEUNIER Isabelle	Ophtalmologie
MULLER Laurent	Anesthésiologie-réanimation et médecine péri-opératoire
PERRIGAULT Pierre-François	Anesthésiologie-réanimation et médecine péri-opératoire
QANTIN Xavier	Pneumologie
ROUBERTIE Agathe	Pédiatrie
VIEL Eric	Soins palliatifs et traitement de la douleur
CLARY BERNARD	Médecine générale
GARCIA Marc	Médecine générale
MILLION Elodie	Médecine générale
REBOUL Marie-Catherine	Médecine générale



Maîtres de Conférence des Universités - Praticiens Hospitaliers

Hors classe - Echelon exceptionnel

MARTRILLE Laurent	Médecine légale et droit de la santé
RICHARD Bruno	Médecine palliative

Hors classe

BADIOU Stéphanie	Biochimie et biologie moléculaire
BOUDOUSQ Vincent	Biophysique et médecine nucléaire
BOULLE Nathalie	Biologie cellulaire
CACHEUX-RATABOUL Valère	Génétique
CARRIERE Christian	Bactériologie-virologie ; hygiène hospitalière
CHARACHON Sylvie	Bactériologie-virologie ; hygiène hospitalière
FABBRO-PERAY Pascale	Epidémiologie ; économie de la santé et prévention
GIANSILY-BLAIZOT Muriel	Hématologie ; transfusion

Première classe

BRET Caroline	Hématologie biologique
BROUILLET Sophie	Biologie et médecine du développement et de la reproduction
CHIRIAC Anca	Immunologie
COSSEE Mireille	Génétique
FITENI Frédéric	Cancérologie ; radiothérapie
GIRARDET BESSIS Anne	Biochimie et biologie moléculaire
HERRERO Astrid	Chirurgie viscérale et digestive
LAVIGNE Géraldine	Hématologie ; transfusion
LESAGE François-Xavier	Médecin et santé au travail
MATHIEU Olivier	Pharmacologie fondamentale ; pharmacologie clinique
MOUZAT Kévin	Biochimie et biologie moléculaire
PANABIERES Catherine	Biologie cellulaire
PANTEL Alix	Bactériologie-virologie ; hygiène hospitalière
RAVEL Christophe	Parasitologie et mycologie
ROUBILLE Camille	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
SCHUSTER BECK Iris	Physiologie
THEVENIN RENE Céline	Immunologie



Deuxième classe

BARATEAU Lucie	Physiologie
BERGOUIGNOUX Anne	Génétique
BOETTO Julien	Neurochirurgie
DESHAYES Emmanuel	Biophysique et médecine nucléaire
DUFLOS Claire	Biostatistiques et informatique médicale
DU THANH Aurélie	Dermatologie – vénéréologie
ENGELMANN Ilka	Bactériologie-virologie ; hygiène hospitalière
GATINOIS Vincent	Histologie, embryologie et cytogénétique
GOUZI Farès	Physiologie
HERBAUX Charles	Hématologie ; transfusion
HUBERLANT Stéphanie	Gynécologie-obstétrique ; gynécologie médicale
LEVEQUE Maude	Parasitologie et mycologie
MARIA Alexandre	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
RANISAVELJEVIC Noémie	Biologie et médecine du développement et de la reproduction
SOUICHE François-Régis	Chirurgie viscérale et digestive

Maîtres de Conférence des Universités

Hors classe

BECAMEL Carine	Neurosciences
DELABY Constance	Biochimie et biologie moléculaire
MIMOUN	Sciences physico-chimiques et ingénierie appliquée à la santé

Classe normale

CHAUMONT-DUBEL Séverine	Sciences du médicament et autres produits de santé
GUGLIELMI Laurence	Sciences biologiques fondamentales et cliniques
HENRY Laurent	Sciences biologiques fondamentales et cliniques
HERBET Guillaume	Neurosciences
LADRET Véronique	Mathématiques appliquées
LAINE Sébastien	Sciences du médicament et autres produits de santé
LE GALLIC Lionel	Sciences du médicament et autres produits de santé
LOZZA Catherine	Sciences physico-chimiques et ingénierie appliquée à la santé



MORITZ GASSER Sylvie	Neurosciences
MOUTOT Gilles	Philosophie
PASSERIEUX Emilie	Physiologie
RAMIEZ Jean-Marie	Histologie
RAYNAUD Fabrice	Sciences du médicament et autres produits de santé
TAULAN Magali	Biologie cellulaire
COSTA David	Médecine générale
OUDE ENGBERINK Agnès	Médecine générale
FOLCO-LOGNOS Béatrice	Médecine générale
CARBONNEL François	Médecine générale

Maîtres de Conférence Associés

BADIN Mélanie	Médecine générale
CAMPAGNAC Jérôme	Médecine générale
LOPEZ Antonio	Médecine générale
MINET Mathilde	Médecine générale
PAVAGEAU Sylvain	Médecine générale
SERAYET Philippe	Médecine générale

Praticiens Hospitaliers Universitaires

CAZAUBON Yoann	Pharmacologie fondamentale ; pharmacologie clinique
DAGNEAUX Louis	Chirurgie orthopédique et traumatologique
GAVOTTO Arthur	Pédiatrie
GOULABCHAND Radjiv	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
MIOT Stéphanie	Médecine interne ; gériatrie et biologie du vieillissement ; médecine générale
POULEN Gaëtan	Neurochirurgie
ROCH Benoît	Pneumologie ; addictologie
URSIC BEDOYA José	Gastroentérologie ; hépatologie ; addictologie



Praticiens Hospitaliers Chargés d'Enseignement

ABOUKRAT Patrick	BRISOT Dominique	FAURE Elsa	MARTIN Lucille
AGUILHON Sylvain	BRONER Jonathan	FILLERON Anne	MATTATIA Laurent
AKKARI Mohamed	CAIMMI Davide Paolo	FILLOLS Mélanie	MEROUEH Fadi
ALRIC Jérôme	CARR Julie	FOURNIER Philippe	MEYER Pierre
AMEDRO Pascal	CARTIER César	GAILLARD Nicolas	MICHEL Moïse
AMOUREUX Cyril	CASPER Thierry	GALMICHE Sophie	MILESI Christophe
ANTOINE Valéry	CASSINOTTO Christophe	GENY Christian	MORAU Estelle
ARQUIZAN Caroline	CATHALA Philippe	GERONIMI Laetitia	MORQUIN David
ATTALIN Vincent	CAZABAN Michel	GINIES Patrick	MOSER Camille
BADR Maliha	CHARBIT Jonathan	GRECO Frédéric	MOUSTY Eve
BAIS Céline	CHEVALLIER Thierry	GROSSIN Delphine	MOUTERDE Gaël
BARBAR Saber Davide	CHEVALLIER-MICHAUD Joseph	GUEDJ Anne Marie	PERNIN Vincent
BASSET Didier	COLIN Olivier	GUYON Gaël	PERRIGAULT Pierre-François
BATIFOL Dominique	CONEJERO Ismael	HEDON Christophe	PEYRON Pierre-Antoine
BATTISTELLA Pascal	CONSEIL Mathieu	HENRY Vincent	PICARD Eric
BAUCHET Luc	CORBEAU Catherine	JAMMET Patrick	PICOT Marie Christine
BELL Ariane	COROIAN Flavia-Oana	JEDRYKA François	POQUET Héléne
BENEZECH Jean-Pierre	COUDRAY Sarah	JREIGE Riad	PUJOL Sarah-Lise
BENNYS Karim	CUNTZ Danielle	KINNE Mélanie	PUPIER Florence
BERNARD Nathalie	DARDALHON Brigitte	LABARIAS Coralie	QUANTIN Xavier
BERTCHANSKY Ivan	DAVID Aurore	LABORDE Caroline	RAFFARD Laurence
BIBOULET Philippe	DE BOUTRAY Marie	LACAMBRE Mathieu	RAPIDO Francesca
BIRON-ANDREANI Christine	DE LA TRIBONNIÈRE Xavier	LANG Philippe	RIBRAULT Alice
BLANC Brigitte	DEBIEN Blaise	LAZERGES Cyril	RICHAUD-MOREL Brigitte
BLANCHARD Sylvie	DELPONT Marion	LE GUILLOU Cédric	RIDOLFO Jérôme
BLANCHET Catherine	DENIS Héléne	LEGLISE Marie Suzanne	RIPART Sylvie
BLATIERE Véronique	DEVILLE DE PERIERE Gilles	LE MOINE Marie-Christine	RONGIERES Michel
BOGE Gudrun	DI CASTRI Alberto	LETERTRE Simon	ROULET Agnès
BOURRAIN Jean Luc	DJANIKIAN Flora	LOPEZ Régis	RUBENOVITCH Josh
BOUYABRINE Hassan	FAIDHERBE Jacques	LUQUIENS Amandine	SANTONI Fannie
BRINGER DEUTSCH Sophie	FATTON Brigitte	MANZANERA Cyril	SASSO Milène
BRINGUIER BRANCHEREAU Sophie	FAUCHERRE Vincent	MARGUERITTE Emmanuel	SCHULDINER Sophie

REMERCIEMENTS

A mes Maîtres et membres du jury,

A Monsieur le Professeur David GENEVIEVE, Coordinateur du D.E.S. de Génétique Médicale de la subdivision Montpellier-Nîmes,

David, il était naturel pour moi de te placer tout en haut de ces remerciements. Ton savoir inépuisable, ton soin accordé à la formation des internes et ton sens aigu du compagnonnage font de toi un mentor d'exception. Ces quelques lignes seront malheureusement bien insuffisantes pour exprimer l'étendue de ce que je te dois déjà, je tâcherai donc d'être laconique. Merci de m'accompagner au quotidien, tant sur le plan médical, scientifique que personnel. Merci pour toutes ces discussions passionnées, pouvant aller de la « déformation diaphysaire en erlenmeyer » aux meilleurs vignobles de la vallée du Rhône, en passant par quelques anecdotes sur ton passé Neckerien et son lot de rencontres présidentielles. Tu sais pouvoir compter sur moi pour la suite, car oui, l'histoire ne fait que commencer ! Je suis profondément fier d'être ton élève.

A Monsieur le Professeur Pascal PUJOL, Responsable de l'Unité d'Oncogénétique du CHU de Montpellier,

Cher Pascal, c'est un très grand honneur que de pouvoir te compter parmi les membres de mon jury. L'humanité, la bienveillance et la culture dont tu fais preuve sont choses rares. Merci pour tes « propositions malhonnêtes » et toutes ces discussions enrichissantes aussi bien dans le service qu'en dehors. Bien que n'ayant pas emprunté la voie de l'oncogénétique, collaborer avec toi est toujours un véritable plaisir, et j'espère pouvoir continuer encore longtemps. ChatGPT ne nous aura pas ! Du moins pas tout de suite...

A Monsieur le Professeur Franck PELLESTOR, Responsable du Laboratoire de Cytogénomique Constitutionnelle du CHU de Montpellier,

Franck, je te remercie de l'immense honneur que tu me fais de participer à ce jury. Ta connaissance encyclopédique de la cytogénétique et des remaniements chromosomiques complexes forcent l'admiration. Merci de m'avoir initié à l'exercice du caryotype que tu as su, de toute évidence, élever au rang d'art Ô toi le Michel-Ange de la bande R ! Travailler sous ta responsabilité est un ravissement

quotidien dont je ne compte plus les bons moments. J'aspire à égaler ta connaissance de la génétique, aussi bien que ton impressionnante collection de calendriers et de goodies.

A Madame le Docteur Marjolaine WILLEMS, Responsable de l'Unité de Génétique Clinique du CHU de Montpellier,

Ma très chère Marjolaine, par où commencer ? Sans doute par te remercier de m'avoir tant appris. Côté un esprit si brillant au quotidien est une véritable chance. Ton intelligence rare, ta magnanimité et ton altruisme font de toi un modèle pour l'interne que je suis, tant sur le plan professionnel que personnel. Merci pour toutes ces discussions, tôt le matin ou tard le soir, où peuvent se mêler médecine, Nocturnes de Chopin, et contrepèteries toujours de très bon goût. Je suis fier d'être ton élève et de pouvoir te compter à mes côtés au club des chasseurs de la gallinette cendrée. Quand je serai grand, je pourrai dire que la meilleure généticienne de France était dans mon jury !

A mes Maîtres de l'Université de Nantes,

A Madame le Professeur Pascale JOLLIET, Doyen de la Faculté de Médecine de Nantes,

Chère Professeur JOLLIET, il me tenait à cœur de vous faire figurer dans cette liste de remerciements. Bien qu'ayant pris la décision de poursuivre mon cursus sur les bords de la Méditerranée, je n'ai pas oublié vos conseils avisés sur les carrières universitaires, ni la bienveillance dont vous avez su faire preuve à l'égard du jeune étudiant en médecine que j'étais. A ce titre, je vous prie de trouver en ces quelques lignes, l'expression de toute ma gratitude et l'assurance de mon meilleur souvenir.

A Madame le Docteur Julie GAVARD, Directrice de Recherche CNRS au Centre de Recherche en Cancérologie et Immunologie Intégrée Nantes-Angers (CRCI²NA),

Julie, l'histoire fut tellement belle que je ne sais par quel chapitre commencer ! Tu as eu le courage, ou peut-être la folie, de m'accueillir dans ton équipe il y a maintenant quelques années, alors que je n'étais encore qu'un jeune étudiant en médecine n'ayant jamais, ne serait-ce que tenu, une pipette de sa vie. Tu m'as fait découvrir, mais surtout aimer, le monde de la recherche fondamentale dans toute sa rigueur, sa complexité et son esthétisme. Avoir été l'élève d'une chercheuse

aussi brillante est un honneur dont peu d'étudiants peuvent se vanter. Alors, je te remercie pour tout ce que tu as pu m'apporter durant cette magnifique aventure, et j'espère que ce bon vieux Gustav trône toujours aussi fièrement derrière toi lors de tes visioconférences !

A Madame le Docteur Gwennan ANDRE-GREGOIRE, Chercheuse ICO au Centre de Recherche en Cancérologie et Immunologie Intégrée Nantes-Angers (CRCI²NA),

Gwennan, ou devrais-je dire, mon coup de cœur professionnel. Tu as été mon guide dans cette contrée inconnue qu'était pour moi la recherche fondamentale. Tu m'as tout appris, de la culture cellulaire, au pipetage dans la cuve de blot, en passant par les dilutions en cascade (en vain pour cette dernière...). Je ne te remercierai jamais assez de la patience et la générosité dont tu as su faire preuve à mon égard. Ton intelligence, ta pédagogie et ta bonne humeur à toute épreuve font de toi une encadrante exceptionnelle. Je suis fier de pouvoir te compter parmi mes amis désormais. Bonne retraite australienne Gwenny, mais reviens-nous vite !

A l'équipe de Génétique Médicale du CHU de Montpellier,

Je ne saurais assez vous remercier pour votre accueil et la bienveillance dont vous avez fait preuve à mon égard depuis mon tout premier jour d'internat. Merci de m'avoir fait une place parmi vous au sein de la belle et grande famille de la Génét(r)ique Montpelliéraine à laquelle je suis désormais fier d'appartenir. Apprendre à vos côtés est un plaisir de tous les instants et j'en mesure la chance au quotidien.

Aux membres de l'Unité de Génétique Clinique,

Christine, je te remercie affectueusement pour toutes nos discussions. Echanger avec toi est toujours un véritable plaisir, notamment quand il s'agit de nos goûts musicaux respectifs. Sans doute as-tu eu à subir les miens plus d'une fois... Je ne serais malheureusement pas des vôtres lors de ton départ, alors je profite de ces remerciements pour te souhaiter une douce et agréable retraite au moulin.

Amiral, merci pour ton humour, ton espièglerie et tous ces débats historiques, musicaux et philosophiques. Tes compétences hors pair de cytogénétique sont uniques et j'en suis profondément admiratif. Nos goûters au Relais H en compagnie

de Franck me manqueront. En parlant de nourriture, quand est-ce qu'on se le fait ce Deux Magots ?

Kévin, tu fais partie de ce cercle très fermé des personnes faisant l'unanimité. Je ne ferai pas exception à la règle, ne pouvant qu'être dithyrambique à ton sujet. Tu es un modèle d'intelligence, de modestie et de simplicité. Merci pour ces longues discussions à l'internat, parfois tard le soir, à échanger sur nos aspirations professionnelles et personnelles. Ne t'arrête jamais de créer, et si d'aventure tu as de nouveau besoin d'un bêta-testeur pour l'un de tes brillants projets, tu sais où me trouver ! Et quand je parle de projets, cela comprend également celui de finir la soirée à l'Antirouille...

Camille, bien que nous nous soyons davantage côtoyés au laboratoire, dans mon cœur tu demeures une clinicienne ! Merci de ton extrême gentillesse et de ta délicatesse. Tu m'as accompagné durant ce stage de génétique moléculaire que ton absence aurait rendu bien plus terne. Désormais nîmoise d'adoption, j'espère que tes nouveaux collègues continuent de penser à tes 2 croissants du matin (« un pour moi et un pour le bébé ! »). Bon, du coup maintenant ça n'en fait plus qu'un... Ne triche pas !

Un grand merci à Constance et Caroline de m'avoir fait découvrir le monde de la fœtopathologie et de la génétique prénatale. Merci à toutes les deux pour votre sympathie et votre grand professionnalisme.

Je tiens également à remercier Emmanuelle et Claire pour leur bonne humeur constante, leur altruisme et leur bienveillance lors de mes premières consultations de conseil génétique desquelles j'ai beaucoup appris. Je garderai un souvenir ému de mes premiers arbres généalogiques à vos côtés, ou devrais-je plutôt dire, de mes premières tentatives d'arbres généalogiques.

Merci à Nathalie et Natacha pour leurs grandes qualités humaines et professionnelles mises au service des patients. Je me fais toujours une joie de pouvoir échanger avec vous.

Nelly, Christiane et Nathalie, je vous remercie chaleureusement d'avoir pris soins de moi durant ces 2 années, aussi bien au sens propre qu'au figuré. Que de fous rires et de comptines fredonnées dans cette salle de soins. Merci pour tout,

notamment de m'avoir fait profiter de votre cachette magique. Promis, ça restera entre nous ! Votre capybara.

Enfin, un immense merci à Stéphanie, Gabrielle et Amélie, ou devrais-je dire les meilleures secrétaires du CHU de Montpellier, qui ont su me supporter au quotidien malgré mes demandes parfois atypiques. Les filles n'oubliez pas, il n'y a que 2 places dans ma valise, j'attends donc vos propositions pour faire mon choix... Stéphanie, j'espère que tu continueras à envoyer des rébus à ton rat de feu préféré, même après son départ. Nos chamailleries quotidiennes me manqueront. JE N'EN PEUX PLUS DE TOI !

Aux membres de l'Unité d'Oncogénétique,

Virginie, je ne sais comment te remercier pour tous ces souvenirs. Les bons moments passés à tes côtés sont si nombreux que je ne pourrais les lister ici au risque d'en oublier la plupart, tout comme tes qualités d'ailleurs. Tu incarnes toute la bonté du monde et j'ai pour toi la plus grande affection. Malgré l'éloignement, je sais que notre amitié demeurera intacte. Maman, ton fiston s'en va pour quelques temps mais ne t'en fais pas, il ne sera jamais très loin de toi ! Mais sache que ton sourire lui manquera beaucoup. Merci pour tout.

Merci à toi Carole de m'avoir initié au monde de l'oncogénétique. J'aurais aimé davantage de connaître. Je te souhaite le meilleur pour l'avenir.

Je tiens également à adresser un immense merci à l'ensemble des conseillères en génétique. Karen, je te remercie de ta gentillesse, ton humour et ton sens du raffinement. Tu resteras ma drama queen préférée. « Aujourd'hui c'est vendredi 13 donc ne venez pas me parler ! Et puis c'est quoi cette mode soudaine pour les spéculoos ? Un tiramisu, c'est au café ! ». Tatiana, merci de ta bonne humeur quotidienne. Après tout, Clément c'est quand même un joli prénom... Marianne, c'est toujours une joie que de pouvoir discuter avec toi, surtout quand tu te souviens de ma date d'anniversaire !

Enfin, je remercie chaleureusement l'équipe des secrétaires médicales pour leur accueil dans le service. Alexandra, promis, je continuerai de porter ma montre au poignet gauche. Mais si un jour je change de côté, tu sauras pourquoi... Surtout si j'arrive en avance ! Corine, merci pour ton humour et tes tenues vestimentaires toujours exactes.

Aux membres de la Plateforme d'Expertise Maladies Rares,

Chloé, il est des rencontres si inattendues qu'elles finissent par marquer une vie. Nos débuts difficiles n'auraient jamais laissé présager d'une amitié si sincère. Je ne liste plus nos fous rires, nos blagues parfois (très) limite, ni tous ces moments de complicité du quotidien pour lesquels je ne saurais assez te remercier. Saint-Exupéry disait de l'amitié qu'elle est une grande circulation de l'esprit au-dessus des détails vulgaires. A défaut d'être plus éloquent, je ne peux que m'approprier ses mots. Je m'apprête à vous quitter pour quelques temps, mais je sais d'ores et déjà que rien ne changera ! Merci pour tout Chloé, tu me manqueras beaucoup. Continue d'être libre.

Un grand merci à Clothilde pour son aide au quotidien. Je sais désormais que je peux facilement te corrompre avec du chocolat ! Promis je ne le dirai à personne. Merci à Anaïs dont j'ai eu l'honneur d'être le premier camarade de bureau. J'espère que tu utilises toujours ton Dyson Airwrap V2 et que tu continues à rédiger tes To do list. Enfin, un grand merci à Solène pour sa bonne humeur et son sens de l'humour dès 8h du matin.

Aux membres du Laboratoire de Génétique des Maladies Rares et Auto-Inflammatoires (UMAI),

Merci à Mouna, Guilaine et Mylène de m'avoir accueilli au sein de leur équipe. Vous avez su m'initier à l'exercice de la génétique moléculaire et vous avez pour cela toute ma reconnaissance. Guilaine, ton expertise de la génétique des maladies auto-inflammatoires force l'admiration. Merci de m'avoir encadré durant ce semestre, apprendre à tes côtés fut une grande chance.

Je souhaite également remercier chaleureusement l'ensemble des techniciennes du laboratoire, ou devrais-je dire mes rayons de soleil de l'UMAI ! Cécile, on se souviendra longtemps de ce stage... quel binôme ! Elodie, merci pour ta gentillesse et ta grande compétence (del in frame dans *DLG4* !). Ta pochette jaune hebdomadaire délicatement posée sur mon bureau me manquera. Aurélie, mes montées de lait n'auront plus vraiment la même saveur sans toi, je dois bien l'avouer ! Nathalie, mes différents envois clandestins de slips usagés outre-Atlantique ne m'auront finalement pas envoyé à Maguelone... Si le GIGN vient m'interpeller à l'heure du laitier, compte sur moi pour te dénoncer et te faire tomber avec moi !

Enfin, un grand merci à Elisabeth pour sa connaissance biblique de la bachata et de la kizomba (de toute façon, c'est pareil !).

Aux membres du Laboratoire de Cytogénomique Constitutionnelle,

Mon très cher Vincent, je te remercie infiniment de m'avoir accompagné tout au long de ce stage de cytogénétique qui, à l'heure où j'écris ces modestes remerciements, s'achèvera dans quelques jours. J'admire tes connaissances, ta finesse d'esprit et ton sens de la transmission. Tu es une personne brillante, tant sur le plan scientifique que relationnel, et tu resteras pour moi un modèle d'intelligence et de pédagogie. Merci de ton dévouement, et vive la Vosgienne !

Anouck, ce stage aurait été bien différent sans toi. Ce fut un véritable plaisir de travailler à tes côtés. Merci de m'avoir fait découvrir « la chtouille », « le lapin qui fait du rugby » et « l'ours qui fait du foot ». Je garderai un souvenir ému de nos premières séances microsat et des staffs CPDPN. Ne m'oublie pas trop vite ! D'ailleurs, cette phrase est-elle injonctive ou impérative ? Si tu trouves la bonne réponse, tu auras des crêpes (sans Nutella), et une compilation d'Axelle Red.

Thomas, merci d'avoir été mon sauveur informatique pendant ces 6 mois. Malheureusement, notre projet de voyage à Dubaï ne s'est pas concrétisé, mais je ne désespère pas. A quand un concert d'Ace of Base (calling) ensemble ?

Ma chère Magalie, quelle chance de t'avoir eue comme voisine de bureau durant ce semestre. Merci pour ta bonne humeur infaillible qui a su résister à mes goûts musicaux frôlant parfois la dépression mélancolique. Mais que veux-tu, la France n'a jamais rien fait de mieux depuis Serge Lama ! Nos discussions du quotidien et tes expressions du Sud me manqueront (on sent que ça réfléchi ici...).

Enfin, un immense merci à l'ensemble des techniciennes pour leur accueil (Camille j'attends toujours mon caillou personnalisé...), à Pierre-Olivier pour ses blagues toujours de très bon goût et à Marie-Laure avec qui je n'aurais malheureusement jamais eu la chance de prendre des cours de tango argentin (promis, je n'ai pas lu le mail !).

A mes co-internes de Génétique Médicale, de Montpellier et d'ailleurs,

Valentin, bien que tu sois désormais passé dans la cour des grands, ta place est ici pour moi. Premier camarade d'internat, tu es celui qui m'a véritablement

accompagné au quotidien dès mon arrivée dans l'équipe, et tu as pour cela toute ma gratitude. Merci de m'avoir initié aux différentes traditions locales et de m'avoir toujours suggéré les bons choix quant à ma vie d'interne. Tes qualités humaines et ton savoir-être professionnel feront de toi un excellent PH. Cependant, tu ne choisiras plus jamais ma bière au Ker Beer !

Flavien, merci d'avoir su apprécier à sa juste valeur l'avant-gardisme gastronomique de Cap'Chef. Lors du prochain congrès de la SFMPP, on tâchera de prendre un lit king size... Signé le barman spécialiste du white russian !

Charlotte, Lucile, Sarah, Jeanne, Vivien et Clément, merci de rendre ces semaines de cours et de DIU toujours plus agréables. Je ne compte plus les bons moments passés ensemble et j'espère que cette amitié durera encore longtemps. L'avenir de la génétique médicale française est entre de bonnes mains ! D'ailleurs, les Assises 2024 arrivant à grands pas, le prochain « cluster caveau » nous tends les bras...

Je tiens enfin à remercier l'ensemble des membres du Bureau 2023-2024 de la SIGF pour leur engagement quotidien au service des internes. Présider une si belle association m'oblige. Clément, de notre périple lisboète seront nées bien des projets de réforme, mais pas que... J'attends toujours le programme des sessions communes ! Rassure-toi je ne te demanderai pas le mode de fonctionnement des amphithéâtres de garnison avant 2010 par peur que tu enchaînes les insomnies...

Aux membres des Bureaux 2021-2022 et 2022-2023, ainsi qu'aux employés du SILR,

Alexis, merci de ton aide et de ton soutien durant ces deux mandats que tu as su mener d'une main de maître. Tes qualités de président ne sont plus à démontrer, tu peux désormais prendre ta retraite avec le sentiment du devoir accompli. Je suis ravi d'avoir été à tes côtés sur ce « chemin de l'espérance ».

Nejm, ton engagement syndical et associatif t'honore. Je suis heureux de pouvoir te considérer comme mon ami. Notre prochain sfendj/thé à la menthe du dimanche matin sera sur Bordeaux ! A défaut, on pourra toujours se rabattre sur un tournoi Tekken et du bissap maison.

Enfin, un immense merci à Sylvia, Marianne, Valérie et Jordan pour leur aide quotidienne et leur dévouement sans faille au service des internes. Vous nous êtes extrêmement précieux.

Aux membres de l'équipe SOAP,

Merci à Kilian et Clément, acolytes des pauses café mais surtout fidèles amis malgré la distance. Les gars, passez le bonjour à la NL et DJ Beb's de ma part. Mais surtout n'oubliez pas, les thésards, c'est comme les piscines...

Merci à toi Cloclo d'avoir supporté mes (rares) accès mégalomaniaques. J'espère que ton fond d'écran, et ton écran de veille, sont restés intacts.

Nicolas, merci pour ta bienveillance, tes conseils toujours pertinents, et ton sens de la rigueur scientifique. Nous parlons ici Messieurs-dames du seul chercheur de l'institut comptant à son actif 1 Nature, 1 Cell et 1 Science, excusez-moi du peu ! Tu es un esprit brillant que j'admire.

Ma chère Laëtitia, je te remercie profondément pour ton dévouement et le soutien que tu as su m'apporter dans mes choix de carrière. Je garde, encore à l'heure actuelle, un souvenir ému de tes multiples et vaines tentatives visant à m'expliquer le fonctionnement du NanoDrop. Rassure-toi, même après 2 ans de génétique, cet appareil demeure toujours aussi mystérieux pour moi...

Enfin, je remercie chaleureusement tous les autres membres de l'équipe avec qui j'ai pu partager un morceau de paillasse, un gel de blot, un concert des Ramoneurs de Menhirs ou une soirée au New Factory. D'ailleurs, c'est quand la prochaine ?

A mes amis,

Merci à Tristan, fidèle parmi les fidèles depuis la 5^{ème} 3. Mon frère, je me surprends parfois à me remémorer cette époque, et c'est non sans une pointe de nostalgie que je repense à tous nos souvenirs communs. Depuis tout ce temps, et malgré l'éloignement, rien n'a changé, sans doute est-ce à cela que l'on reconnaît une véritable amitié. On la finira ensemble cette histoire. L'avenir est à nous !

Merci à Clémence et Alex, élus « meilleurs critiqueurs des Sables d'Olonne » pour la 10^{ème} année consécutive. Papa, maman, regardez comme votre fils grandit ! Ça va faire plus 15 ans maintenant (bon un peu moins pour Alex et ses 3 ans de

réclusion criminelle à perpétuité) et pourtant tout est comme au premier jour. Merci pour tous ces moments d'amitié, de complicité, et j'espère qu'un jour nous aurons la chance de retrouver la Mouche.

Merci à mon ami Pap's, ou la valeur sûre du RCS, ou l'homme aux traits fins. Que de chemin parcouru depuis nos cours de physique-chimie de terminale, n'est-ce pas ? Merci pour ton soutien sans faille depuis toutes ces années. Et de toute façon, on l'a toujours dit, si je meurs avant 40 ans tu sauras très bien pourquoi !

Un immense merci à Simon et Alex pour la Vie Rapide. La plupart de nos aventures ne sont pas racontables ici, et vous savez desquelles je parle ! A vrai dire, il y en a tellement que je ne saurais par où commencer. Merci pour tout les gars, vous faites partie de la famille désormais. Quand je repense au Razzmatazz et à notre premier dîner chez maître Diouf...

Merci à mon grand frère Rob, toujours présent à mes côtés depuis maintenant 20 ans. Mon frère, tu sais déjà tout. Ensemble, on ira jusqu'au bout. MIA SAN MIA !

Enfin, merci à Guillaume et Gabriel, membres du « Labo des Légendes ». Nous avons tous les trois conscience de la valeur de notre amitié, ce qui la rend d'autant plus précieuse. Merci pour tous ces moments de franche camaraderie au Magellan ou au Café du Commerce, à ton grand désarroi Guillaume. Gabriel, cet affogato sous le ciel pluvieux de Montréal fera date dans l'Histoire. On l'aura notre institut ! Messieurs, je profite également de ces remerciements pour vous annoncer que nos prochaines retrouvailles se feront au Flore et nulle part ailleurs !

A mes parents,

Papa, Maman, merci de m'avoir tout donné depuis maintenant 27 ans. Vous m'avez transmis votre sens de l'effort et avez su me soutenir en toute circonstance, qu'importe mes choix professionnels et personnels. Maman, merci de ta douceur et de tes attentions quotidiennes. Papa, merci pour toutes ces valeurs que tu as su m'inculquer. Je n'aurais pas pu rêver de meilleurs parents pour me construire en tant qu'homme. J'espère vous rendre fiers et être à la hauteur de l'éducation que vous m'avez donnée. Vous êtes et resterez pour toujours ma priorité. Je vous aime d'un amour incommensurable.

TABLE DES MATIÈRES

INTRODUCTION	27
ORIGINAL ARTICLE	33
I. Preamble	34
<i>A. List of abbreviations</i>	34
<i>B. List of figures</i>	35
<i>C. List of tables</i>	35
<i>D. List of supplemental figures</i>	35
<i>E. List of supplemental tables</i>	35
II. Abstract	38
III. Introduction	39
IV. Material and methods	42
<i>A. Clinical and molecular data</i>	42
i. Recruitment of individuals and data collection	42
ii. DNA sample collection	43
iii. Statistical analyses	44
iv. Composite gelstat of <i>BCL11B</i> -RD	44
<i>B. Immunophenotyping</i>	44
<i>C. DNA methylation analysis</i>	44
<i>D. Matched controls selection and DNA methylation profiling</i>	45
<i>E. Construction of the binary classification model</i>	46
<i>F. Functional genomic assessment and methylome overlap</i>	47
V. Results	48
<i>A. Clinico-biological data</i>	48
<i>B. Episignature discovery</i>	52
i. Accurate episignature	53
ii. Subset specific episignature	55
<i>C. Functional genomic assessment and methylome overlap</i>	56
VI. Discussion	56
<i>A. Clinico-biological description</i>	56
<i>B. Epigenetic signature</i>	58
VII. References	61

VIII. Supplemental	65
<i>A. Supplemental figures</i>	65
<i>B. Supplemental tables</i>	70
IX. Additional information	71
<i>A. Data availability</i>	71
<i>B. Fundings and acknowledgments</i>	71
<i>C. Ethics declaration</i>	71
<i>D. Conflict of interest</i>	71
CONCLUSION	72
ANNEXES	78
SERMENT	82
RÉSUMÉ	83

INTRODUCTION

TROUBLE DU DÉVELOPPEMENT DE L'INTELLIGENCE : UN MODÈLE D'HÉTÉROGÉNÉITÉ CLINICO-GÉNÉTIQUE

Le trouble du développement de l'intelligence (TDI), anciennement appelé « déficience intellectuelle », est défini par l'Organisation Mondiale de la Santé (OMS) comme une capacité réduite de comprendre une information nouvelle ou complexe, d'apprendre et d'appliquer de nouvelles compétences, et de faire face à toute situation de manière indépendante¹. La prévalence du TDI dans la population occidentale est estimée **entre 1 et 3%** selon les différentes méta-analyses^{2,3}. En France, le TDI concernerait alors plus de **1.300.000** personnes et aux alentours de **120.000** pour la seule région Occitanie.

Selon le DSM-V, publié par l'*American Psychiatric Association (APA)* en 2015, le diagnostic de TDI doit reposer sur la présence de 3 critères⁴ :

1. Un **trouble de l'intelligence**, défini par un quotient intellectuel total (QIT) inférieur à 70, et objectivé grâce à une échelle de mesure du QIT standardisée et adaptée pour l'âge (échelles de Wechsler)⁵ ;
 - > *Echelle WPPSI-IV (2 sous-types)* : de 2 ans et 6 mois à 3 ans et 11 mois et de 4 ans à 7 ans et 7 mois ;
 - > *Echelle WISC-V* : de 6 ans à 16 ans et 11 mois ;
 - > *Echelle WAIS* : à partir de 16 ans.
2. Un **trouble du fonctionnement adaptatif** se manifestant par une incapacité à se conformer aux normes socioculturelles et à endosser sa responsabilité sociale dans la vie quotidienne ;
3. L'apparition de ces troubles dans **l'enfance**.

Avant 2015, la classification de la sévérité du TDI reposait uniquement sur la mesure du QIT (DSM-IV)⁶ :

¹ Salvador-Carulla L, Reed GM, Vaez-Azizi LM, Cooper SA, Martinez-Leal R, Bertelli M, Adnams C, Cooray S, Deb S, Akoury-Dirani L, Girimaji SC, Katz G, Kwok H, Luckasson R, Simeonsson R, Walsh C, Munir K, Saxena S. Intellectual developmental disorders: towards a new name, definition and framework for "mental retardation/intellectual disability" in ICD-11. *World Psychiatry*. 2011 Oct;10(3):175-80.

² Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders; Board on the Health of Select Populations; Board on Children, Youth, and Families; Institute of Medicine; Division of Behavioral and Social Sciences and Education; The National Academies of Sciences, Engineering, and Medicine; Boat TF, Wu JT, editors. *Mental Disorders and Disabilities Among Low-Income Children*. Washington (DC): National Academies Press (US); 2015 Oct 28. 15, Prevalence of Intellectual Disabilities.

³ McKenzie, K., Milton, M., Smith, G. et al. Systematic Review of the Prevalence and Incidence of Intellectual Disabilities: Current Trends and Issues. *Curr Dev Disord Rep* 3, 104–115 (2016).

⁴ Papazoglou A, Jacobson LA, McCabe M, Kaufmann W, Zabel TA. To ID or not to ID? Changes in classification rates of intellectual disability using DSM-5. *Intellect Dev Disabil*. 2014 Jun;52(3):165-74.

⁵ Boake C. From the Binet-Simon to the Wechsler-Bellevue: tracing the history of intelligence testing. *J Clin Exp Neuropsychol*. 2002 May;24(3):383-405.

⁶ Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders; Board on the Health of Select Populations; Board on Children, Youth, and Families; Institute of Medicine; Division of Behavioral and Social Sciences and Education; The National Academies of Sciences, Engineering, and Medicine; Boat TF, Wu JT, editors. *Mental Disorders and Disabilities Among Low-Income Children*. Washington (DC): National Academies Press (US); 2015 Oct 28. 9, Clinical Characteristics of Intellectual Disabilities.

- **TDI léger** : QIT entre 50 et 69 ;
- **TDI modéré** : QIT entre 36 et 49 ;
- **TDI sévère** : QIT entre 20 et 35 ;
- **TDI profond** : QIT inférieur à 20.

Depuis la publication du DSM-V, la sévérité du TDI est désormais évaluée à partir de l'autonomie du patient au quotidien basée sur ses capacités pratiques, sociales et conceptuelles⁴ :

- **TDI léger** : autonomie quotidienne avec assistance minimale ;
- **TDI modéré** : autonomie quotidienne nécessitant une assistance active (e.g. foyers de vie) ;
- **TDI sévère** : assistance nécessaire pour les tâches du quotidien (e.g. toilette, habillage, repas etc.) ;
- **TDI profond** : assistance 24h/24h nécessaire.

Selon les différentes études épidémiologiques, la distribution de la sévérité du TDI serait la suivante : 85% de forme légère, 10% de forme modérée, 3,5% de forme sévère et 1,5% de forme profonde^{2,6}.

Le spectre étiologique du TDI est extrêmement vaste, allant des expositions environnementales prénatales aux causes génétiques constitutionnelles.

La première étape de l'enquête étiologique d'un TDI est ainsi d'éliminer une cause acquise^{7,8,9} :

→ **Période prénatale**

- > *Causes médicamenteuses* : embryofœtopathie au valproate etc.
- > *Causes toxiques* : syndrome d'alcoolisation fœtale (SAF)
- > *Causes infectieuses (infections materno-fœtales)* : toxoplasmose congénitale, rubéole congénitale etc.
- > *Causes immunitaires* : alloimmunisation fœto-maternelle

→ **Période périnatale**

- > Grande prématurité
- > Encéphalopathie anoxo-ischémique néonatale

⁷ Reichenberg A, Cederlof M, McMillan A, Trzaskowski M, Kapra O, Fruchter E, et al. Discontinuity in the genetic and environmental causes of the intellectual disability spectrum. Proc Natl Acad Sci U S A. 2016;113(4):1098-103

⁸ de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BB, Brunner HG, Veltman JA, Vissers LE. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012 Nov 15;367(20):1921-9.

⁹ Leonard H, Wen X. The ep demiology of mental retardation: challenges and opportunities in the new millennium. Ment Retard Dev Disabil Res Rev. 2002;8(3):117-34.

→ **Période post-natale** : étiologies séquellaires

- › *Causes infectieuses* : méningites, syndrome hémolytique et urémique (SHU) etc.
- › *Causes traumatiques* : traumatisme crânien, *shaken baby syndrom* etc.

Une fois l'ensemble de ces étiologies acquises éliminées, des analyses génétiques constitutionnelles doivent être conduites à la recherche d'une variation chromosomique et/ou moléculaire, ces dernières étant identifiées chez plus de **60%** des patients atteints d'un TDI sans cause acquise retrouvée^{8,10}. Parmi les causes génétiques les plus fréquemment responsables de TDI, nous pouvons ainsi citer :

→ Les causes **chromosomiques** (10%), telles que :

- › *Les aneuploïdies* : trisomie 21, première cause de TDI d'origine génétique¹¹, trisomie 18 (syndrome d'Edwards) etc.
- › *Les syndromes microdélétionnels* : microdélétion 22q11.2 (syndrome vélo-cardio-facial), microdélétion 4pter (syndrome de Wolf-Hirschhorn), microdélétion 7q11.23 (syndrome de Williams-Beuren)¹² etc.
- › *Les syndromes microduplicationnels* : microduplication 7q11.23, microduplication Xq28 (syndrome de duplication du gène *MECP2*) etc.

→ Les causes **moléculaires** (50%), caractérisées par des variations ponctuelles (SNVs) et des petites insertions/délétions (indel) dans la séquence de gènes impliqués dans le développement du système nerveux central^{12,13}.

Le syndrome de l'X fragile, décrit en 1991 et secondaire à un nombre de répétitions du triplet CGG supérieur à 200 dans le gène *FMR1*, fut l'une des premières causes moléculaires identifiées comme responsable de TDI et représente aujourd'hui la première étiologie de TDI hérité^{14,15}. Depuis, le développement des nouvelles

¹⁰ Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BB, Kleefstra T, Brunner HG, Vissers LE, Veltman JA. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. 2014 Jul 17;511(7509):344-7.

¹¹ Sherman SL, Allen EG, Bean LH, Freeman SB. Epidemiology of Down syndrome. *Ment Retard Dev Disabil Res Rev*. 2007;13(3):221-7.

¹² Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet*. 2016 Jan;17(1):9-18.

¹³ Bowling KM, Thompson ML, Amaral MD, Finnila CR, Hiatt SM, Engel KL, Cochran JN, Brothers KB, East KM, Gray DE, Kelley WV, Lamb NE, Lose EJ, Rich CA, Simmons S, Whittle JS, Weaver BT, Nesmith AS, Myers RM, Barsh GS, Bebin EM, Cooper GM. Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med*. 2017 May 30;9(1):43. doi: 10.1186/s13073-017-0433-1.

¹⁴ Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP, et al. Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*. 1991 May 31;65(5):905-14.

¹⁵ Ciaccio C, Fontana L, Milani D, Tabano S, Miozzo M, Esposito S. Fragile X syndrome: a review of clinical and molecular diagnoses. *Ital J Pediatr*. 2017 Apr 19;43(1):39.

technologies telles que l'analyse chromosomique sur puce à ADN (ACPA) pour l'identification des microréarrangements chromosomiques (*i.e.* microdélétions et microduplications) et les examens pangénomiques basés sur le séquençage nouvelle génération (NGS), à l'instar du séquençage de l'exome ou du génome, ont entraîné une augmentation exponentielle des diagnostics génétiques chez les patients atteints d'un TDI (Figure A).

En effet, selon la dernière mise à jour du 03/08/2023 de la base de données SysNDD (<https://sysndd.dbmr.unibe.ch/>), **1615** gènes sont actuellement identifiés comme responsable de TDI isolé ou syndromique et **1368** sont recensés comme candidat.

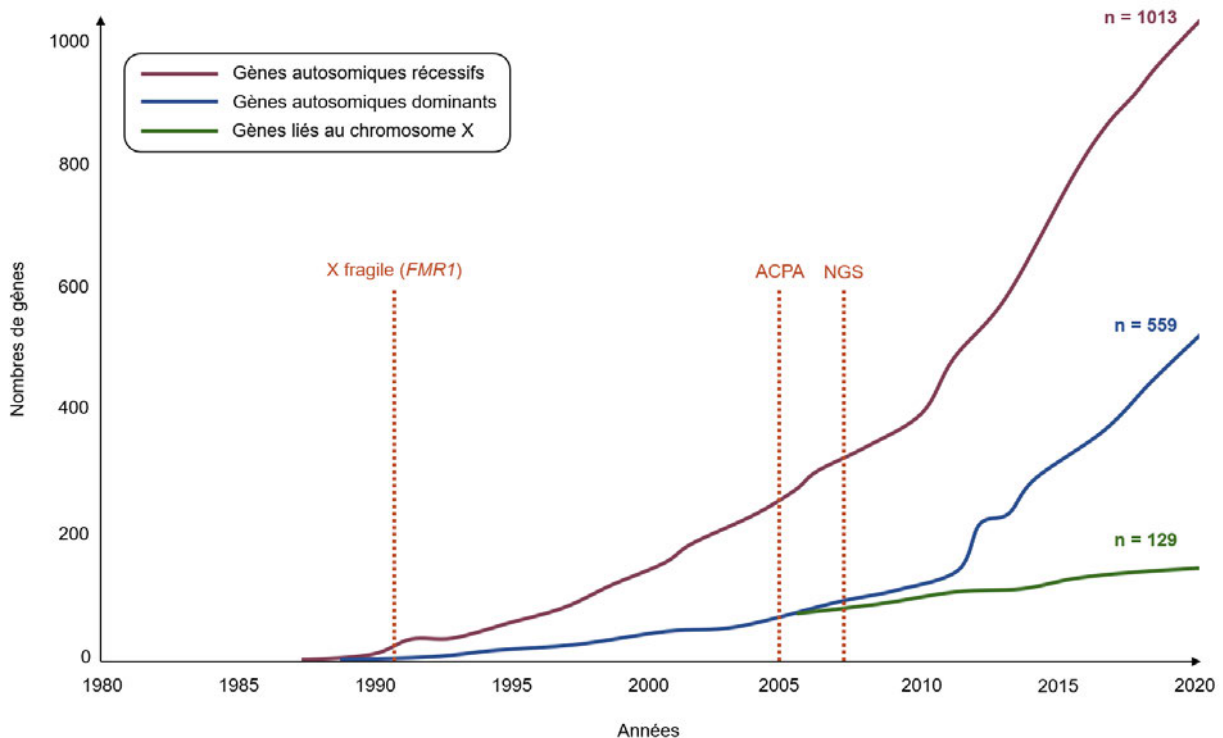


Figure A. Evolution du nombre de gènes responsables de trouble du développement de l'intelligence (TDI). A ce jour, 1615 gènes ont été décrits comme responsables de TDI, dont 1013 gènes autosomiques récessifs, 559 gènes autosomiques dominants (variations *de novo* pour 80% d'entre eux¹²), 129 gènes liés au chromosome X et 15 gènes du génome mitochondrial. Abréviations : ACPA (analyse chromosomique sur puce à ADN), NGS (séquençage nouvelle génération). Figure adaptée de [Visser *et al.*, *Nature Review Genetics*, 2016] à partir des données de la base SysNDD.

Cette hétérogénéité génétique majeure, associée à la variabilité de l'expression du phénotype présenté par les patients atteints d'un TDI (*i.e.* forme isolée ou syndromique) rendent ainsi la stratégie « *phenotype first* » et l'approche ciblée par séquençage d'un panel de gènes responsables de TDI désormais désuètes. Aussi, la stratégie diagnostique « *genotype first with reverse phenotyping* » demeure désormais la règle chez les patients atteints d'un TDI avec des explorations génétiques de première intention reposant sur une ACPA associée à un séquençage pangénomique de type exome ou génome, en fonction de l'urgence diagnostique et de la demande parentale^{16,17}.

L'article original utilisé comme base de cette thèse s'intéresse ainsi à l'un de ces nombreux gènes, dénommé *BCL11B*, et ayant récemment été identifié comme responsable d'un TDI syndromique de forme autosomique dominante¹⁸. Avec seulement **29 patients rapportés dans la littérature**, la maladie liée au gène *BCL11B* est caractérisée par un trouble du neurodéveloppement syndromique associant un TDI, des éléments morphologiques mineurs de la face et un déficit immunitaire¹⁷.

Basée sur une collaboration internationale, ce travail décrit les données cliniques, génotypiques et immunitaires de **20 nouveaux patients** porteurs d'une variation pathogène dans le gène *BCL11B* identifiée par examens pangénomiques. Il rapporte également l'existence d'une **signature épigénétique** spécifique de la maladie liée au gène *BCL11B* pouvant désormais être utilisée en pratique clinique quotidienne afin d'aider à la reclassification des variations de signification incertaine (VSI) de *BCL11B*, tout en permettant le diagnostic de cette maladie rare chez les patients dont les explorations génétiques se sont avérées non contributives.

¹⁶ 100,000 Genomes Project Pilot Investigators; Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, Cipriani V, Ellingford JM, Arno G, Tucci A, Vandrovicova J, Chan G, Williams HJ, Ratnaik T [...] Mumford AD, Wilson G, Beales P, Btner-Glindzicz M, Black G, Bradley JR, Brennan P, Burn J, Chinnery PF, Elliott P, Flinter F, Houlden H, Irving M, Newman W, Rahman S, Sayer JA, Taylor JC, Webster AR, Wilkie AOM, Ouwehand WH, Raymond FL, Chisholm J, Hill S, Bentley D, Scott RH, Fowler T, Rendon A, Caulfield M. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. *N Engl J Med.* 2021 Nov 11;385(20):1868-1880.

¹⁷ Turro E, Astle WJ, Megy K, Gräf S, Greene D, Shamardina O, Allen HL, Sanchis-Juan A, Frontini M, Thys C, Stephens J, Mapeta R, Burren OS, Downes K, Haimel M, Tuna S, Deevi SVV, Aitman TJ, Bennett DL, Calleja P, Carss K, Caulfield MJ, Chinnery PF, Dixon PH, Gale DP, James R, Koziell A, Laffan MA, Levine AP, Maher ER, Markus HS, Morales J, Morrell NW, Mumford AD, Ormondroyd E, Rankin S, Rendon A, Richardson S, Roberts I, Roy NBA, Saleem MA, Smith KGC, Stark H, Tan RYY, Themistocleous AC, Thrasher AJ, Watkins H, Webster AR, Wilkins MR, Williamson C, Whitworth J, Humphray S, Bentley DR; NIHR BioResource for the 100,000 Genomes Project; Kingston N, Walker N, Bradley JR, Ashford S, Penkett CJ, Freson K, Stirrups KE, Raymond FL, Ouwehand WH. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature.* 2020 Jul;583(7814):96-102.

¹⁸ Lessel D, Gehbauer C, Bramswig NC, Schluth-Bolard C, Venkataramanappa S, van Gassen KLI, Hempel M, Haack TB, Bares C A, Genetti CA, Funari MFA, Lessel I, Kuhlmann L, Simon R, Liu P, Denecke J, Kuechler A, de Kruijff I, Shoukier M, Lek M, Mullen T, Lüdecke HJ, Lerario AM, Kolbe R, Krieger T, Demeer B, Lebrun M, Keren B, Nava C, Buratti J, Afenjar A, Shinawi M, Guillen Sacoto MJ, Gauthier J, Hamdan FF, Laberge AM, Campeau PM, Louie RJ, Cathey SS, Prinz I, Jorge AAL, Terhal PA, Lenhard B, Wiczorek D, Strom TM, Agrawal PB, Britsch S, Tolosa E, Kubisch C. *BCL11B* mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain.* 2018 Aug 1;141(8):2299-2311.

ORIGINAL ARTICLE

I. Preamble

A. List of abbreviations

ACMG : American College of Medical Genetics

ASD : autism spectrum disorder

AUROC : area under ROC curve

BCL11B : B-cell lymphoma/leukemia 11B

CDK9 : cyclin-dependent kinase 9

CDS : coding sequence

CT : computed tomography

CTD : C-terminal domain

CTL : cytotoxic T cells

CGH : comparative genomic hybridization

DMPs : differentially methylated probes

DN : double negative

DNA : deoxyribonucleic acid

EKD : EpiSign Knowledge Database

ERN-ITHACA : European Reference Network on Rare Congenital Malformations and Rare Intellectual Disability

GADEVS : Gabriele-de Vries syndrome

H3 : histone 3

HEXIM1 : hexamethylene bis-acetamide inducible 1

HGVS : Human Genome Variation Society

ID : intellectual disability

IDDSFTA : intellectual developmental disorder with speech delay, dysmorphic facies and T cell abnormalities

IRB : institutional review board

K9 : lysine 9

K122 : lysine 122

KMT1A : lysine N-methyltransferase type 1A

LHSC : London Health Sciences Centre

LoF : loss-of-function

LOOCV : leave-one-out cross validation

LP : likely pathogenic

MAPK : mitogen-activated protein kinase

MDS : multidimensional scaling

MIM : mendelian inheritance in man

MRI : magnetic resonance imaging

MVP : methylation variant pathogenicity

NK : natural killer

P : pathogenic

P-TFEB : positive transcription elongation factor b

PBMC : peripheral blood mononuclear cells

PCA : principal component analysis

NuRD : nucleosome remodeling deacetylase

RD : related disorder

RNA : ribonucleic acid

RNAP II : RNA polymerase II

RTE : recent thymic emigrants

RTS : Rubinstein-Taybi syndrome

SD : standard deviation

SNP : single-nucleotide polymorphism

SP : single positive

SUMO : small-ubiquitin-like modifier

SVM : support vector machine

TCR : T cell receptor

TEM : effector memory CD8⁺ T cells

TEMRA : effector-memory expressing CD45RA CD8⁺ T cells

Th : T helper

Treg : regulatory T cell

TSS : transcription start site

VUS : variant of uncertain significance

ZnF : zinc finger

B. List of figures

Figure 1. Linear representation of BCL11B isoforms.

Figure 2. Control of BCL11B over transcriptional activity.

Figure 3. Involvement of BCL11B in T cell lymphopoiesis.

Figure 4. Front and lateral views of individuals with *BCL11B*-related disorder.

Figure 5. *BCL11B* genotypic spectrum in this series.

Figure 6. Assessment of the robustness of the episignature with higher sensitivity.

Figure 7. Assessment of the robustness of the episignature with higher specificity.

C. List of tables

Table 1. Clinical features of individuals with *BCL11B*-related disorder in this series and comparison with the literature.

Table 2. Immunophenotyping profiles of 4 individuals with *BCL11B*-related disorder associated with abnormal T cell compartment.

D. List of supplemental figures

Supplemental Figure 1. Flow chart diagram of the study.

Supplemental Figure 2. Facial features analysis of 4 individuals with *BCL11B*-related disorder (*BCL11B*-RD) using *Face2Gene* software.

Supplemental Figure 3. LOOCV for both episignatures.

Supplemental Figure 4. Comparison of the methylation pattern of *BCL11B*-RD and EpiSign v3 classifier disorders.

Supplemental Figure 5. Annotation of differentially methylated probes (DMPs).

E. List of supplemental tables

Supplemental Table 1. DNA samples annotation, molecular findings and detailed methylation variant pathogenicity (MVP) scores.

II. Abstract

Purpose: *BCL11B*-related disorder (*BCL11B*-RD) arises from rare genetic variants within the *BCL11B* gene, resulting in a distinctive clinical spectrum encompassing syndromic neurodevelopmental disorder, with or without intellectual disability, associated with facial features and impaired immune function. This study presents an in-depth clinico-biological analysis of 20 newly reported individuals with *BCL11B*-RD, coupled with a characterization of genome-wide DNA methylation patterns of this genetic condition.

Methods: Through an international collaboration, clinical and molecular data from 20 individuals were systematically gathered, and a comparative analysis was conducted between this series and existing literature. We further scrutinized peripheral blood DNA methylation profile of individuals with *BCL11B*-RD, contrasting them with healthy controls and other neurodevelopmental disorders marked by established epismature.

Results: Our findings unveil rarely documented clinical manifestations, notably including Rubinstein-Taybi-like facial features, craniosynostosis, and autoimmune disorders, all manifesting within the realm of *BCL11B*-RD. We refine the intricacies of T cell compartment alterations of *BCL11B*-RD, revealing decreased levels naïve CD4⁺ T cells and recent thymic emigrants (RTE) while concurrently observing an elevated proportion of effector-memory expressing CD45RA CD8⁺ T cells (TEMRA). Finally, a distinct DNA methylation epismature exclusive to *BCL11B*-RD is unveiled.

Conclusion: This study serves to enrich our comprehension of the clinico-biological landscape of *BCL11B*-RD, potentially furnishing a more precise framework for diagnosis and follow-up of individuals carrying pathogenic *BCL11B* variant. Moreover, the identification of a unique DNA methylation epismature offers a valuable diagnosis tool for *BCL11B*-RD, thereby facilitating routine clinical practice by empowering physicians to reevaluate variants of uncertain significance within the *BCL11B* gene.

Key words

BCL11B - Neurodevelopmental delay - Methylation - Epismature

III. Introduction

B-Cell Lymphoma/Leukemia 11B (*BCL11B*, MIM 606558) is a zinc finger transcription factor involved in the development of various tissues through a dual action (*i.e.* activation/repression) on epigenetic regulation (Figure 1) [1-2].

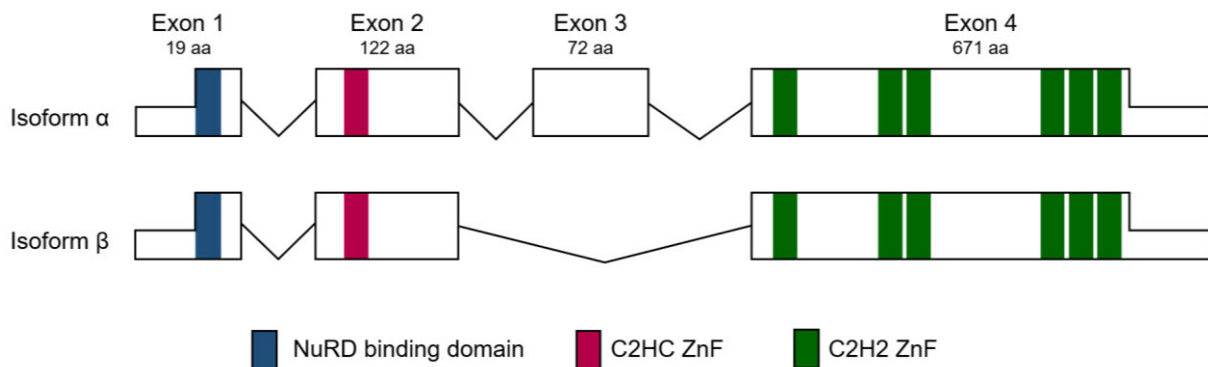


Figure 1. Linear representation of BCL11B isoforms. *BCL11B* gene is composed of 4 exons encoding 2 isoforms, α and β , consisting of 884 and 812 amino acids respectively. Isoforms α and β contain a NuRD interacting domain (blue), 1 C2HC (crimson) and 6 C2H2 (green) zinc finger domains [1-2].

MAPK-mediated SUMOylation of *BCL11B* results in recruitment of the transcriptional co-activator p300 with subsequent induction of transcriptional activity through histone 3 lysine 122 acetylation (H3K122ac) [3]. Oppositely, de-SUMOylated *BCL11B* interacts with the nucleosome remodelling and histone deacetylase (NuRD) complex and KMT1A (also known as SUV39H1), both leading to chromatin compaction by histone deacetylation and histone 3 lysine 9 trimethylation (H3K9me3), respectively (Figure 2, upper panel) [4][5].

BCL11B further represses RNA polymerase II (RNAP II) activity by the recruitment of HEXIM1 resulting in down-regulation of the positive transcription elongation factor b (P-TFEB) and subsequent dephosphorylation of RNAP II C-terminal domain (CTD) (Figure 2, bottom panel) [6].

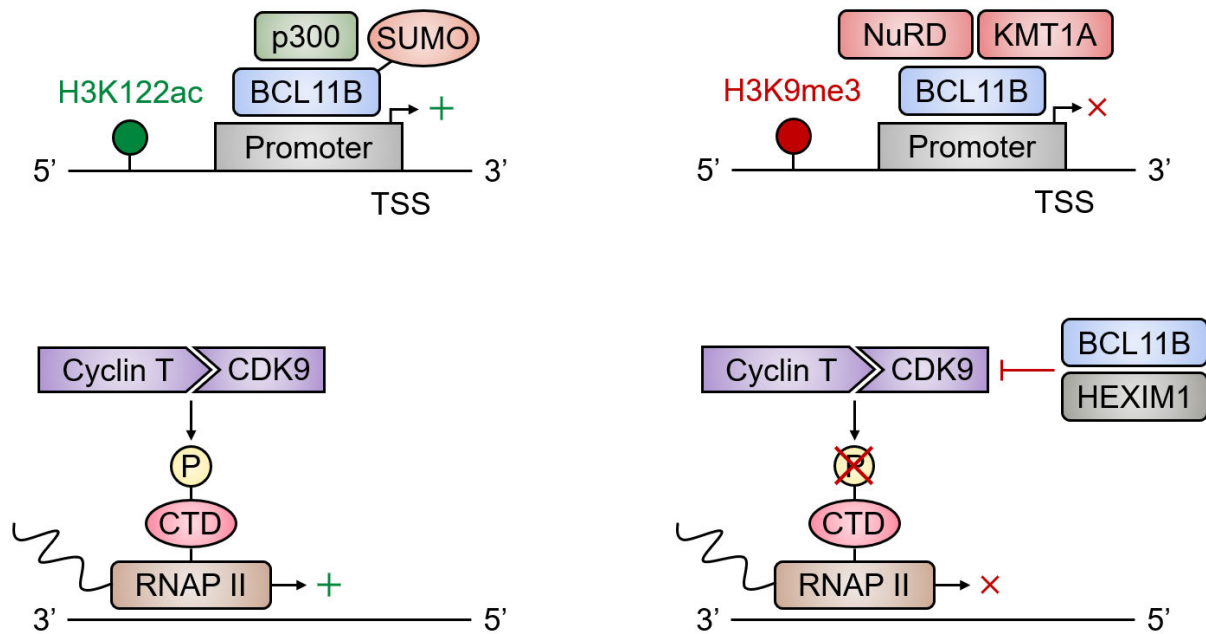


Figure 2. Control of BCL11B over transcriptional activity. BCL11B is a key regulator of transcriptional activity through its influence on chromatin dynamics and histone 3 (H3) post-translational modifications (*i.e* acetylation and methylation) (upper panel). On the other hand, the BCL11B/HEXIM1 heterodimer is a down-regulator of the positive transcription elongation factor b (P-TFEB), composed of cyclin T and cyclin-dependent kinase 9 (CDK9). This complex is involved in the phosphorylation of RNA polymerase II (RNAP II) C-terminal domain (CTD), an essential step for transcription activation [6-7] (bottom panel).

From a functional standpoint, BCL11B orchestrates various cellular and physiological functions including T cell lymphopoiesis, craniofacial bone development and postnatal neurogenesis [8-15]. In T cell lymphopoiesis, BCL11B is involved at multiple stages, from early T cell progenitors in the thymus to the differentiation of mature T cells in the peripheral blood (Figure 3). BCL11B plays a pivotal role in lineage commitment, governing the differentiation of hematopoietic stem cells into T cell precursors while inhibiting alternative cell fates, such as B cell development. Beyond its role in lineage specification, BCL11B further exerts control over mature T cell functions, including their activation, proliferation and cytokine production. Hence, BCL11B acts like a gatekeeper for effector and memory T cell differentiation, shaping T cell repertoire and ensuring T cell homeostasis [8].

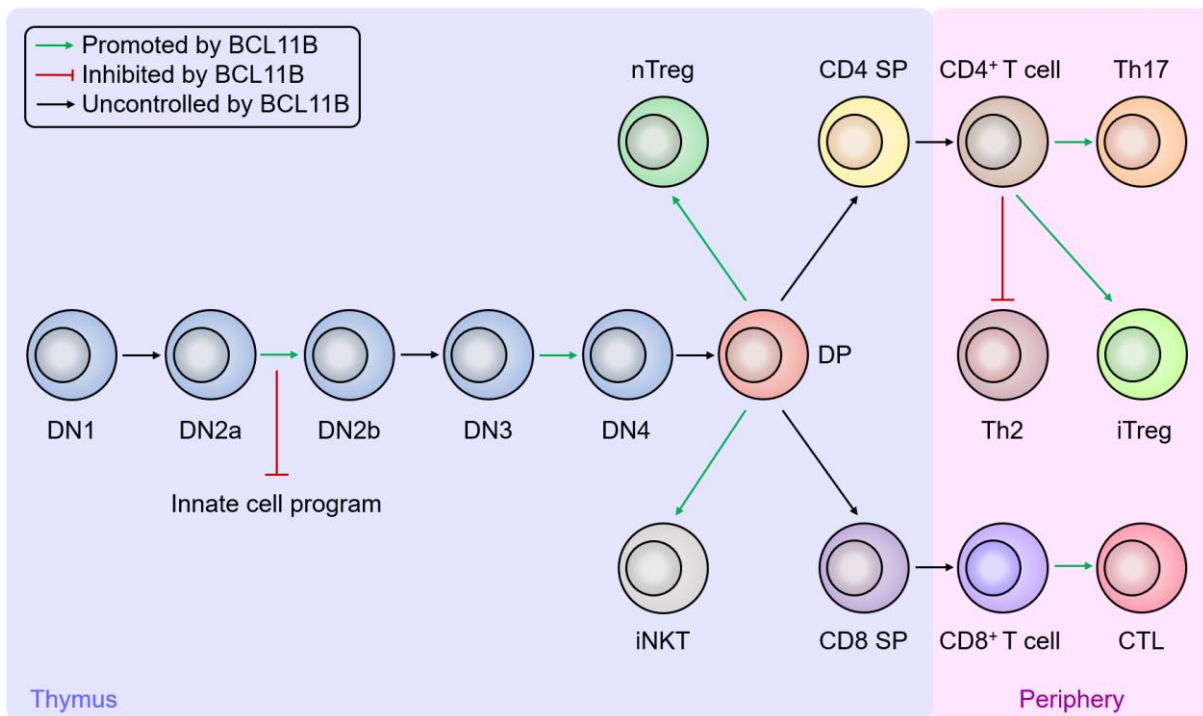


Figure 3. Involvement of BCL11B in T cell lymphopoiesis. BCL11B expression is initiated during T cell development at DN2 stage. BCL11B promotes DN3 and double positive (DP) thymocytes survival, governs TCR β -selection at DN3 stage and drives the differentiation of invariant natural killer T cells (iNKT) and regulatory T cells including natural (nTreg) and inducible (iTreg) subtypes. In mature T cells, BCL11B controls the expansion of cytotoxic T cells (CTL) and Th17 cells by inhibiting Th2 program [8-11].

BCL11B further assumes a pivotal role in craniofacial development, exhibiting high expression level in both osteogenic and sutural mesenchyme of the nascent craniofacial complex. Additionally, *Bcl11b* knock-out in mice results in a premature osteoblast proliferation associated with increased skull density and synostosis of facial and calvarial sutures [12-13]. Moreover, BCL11B activity extends into the realm of adult neurogenesis. *In vivo* experiments employing *Bcl11b* loss-of-function models identified its crucial role in hippocampal functionality, leading to reduced dentate gyrus size, blocking of granule cells differentiation, and impaired spatial learning abilities in adult mice [14-15].

BCL11B-related disease was firstly reported in 2018 by exome sequencing of individuals displaying neurodevelopmental delay associated with facial features

and immunodeficiency, revealing 13 individuals with a pathogenic variation in *BCL11B* [1]. Ever since, *BCL11B* pathogenic variants were found to be responsible for *BCL11B*-related disorder (*BCL11B*-RD), also known as intellectual developmental disorder with dysmorphic facies, speech delay and T cell abnormalities (IDDSFTA, MIM 618092).

Increasing number of genetic disorders including ones with genes related to DNA methylation and chromatin accessibility were demonstrated to exhibit sensitive and specific DNA methylation profiles in peripheral blood, also known as epesignature [16-17]. DNA methylation analysis has recently been implemented with the diagnosis of clinical genomic DNA methylation test EpiSign™ in individuals with rare genetic disorders, therefore proving strong evidence of clinical utility, including the ability to provide conclusive diagnostic findings in most tested individuals [18-19]. As abnormal transcription and chromatin compaction may have indirect consequences on DNA methylation, individuals with *BCL11B*-RD might display specific epesignature compared to other individuals.

Therefore, the purpose of this study is to give further insights into clinical manifestations and molecular spectrum of individuals with *BCL11B*-RD. To this aim, we describe 20 unreported individuals with *BCL11B*-RD and perform a literature review to summarize both clinical and molecular data of 29 previously published individuals. In addition to refine the clinico-biological spectrum of *BCL11B*-RD, we provide evidence of a sensitive and specific DNA methylation epesignature for individuals with pathogenic and likely pathogenic variants in *BCL11B*.

IV. Material and methods

A. Clinical and molecular data

i. Recruitment of individuals and data collection

Individuals with *BCL11B*-RD caused by rare *BCL11B* variants were identified by targeted genes panel, exome or genome sequencing after being referred to genetic centers for syndromic neurodevelopmental delay.

Data for 20 previously unreported individuals with heterozygous pathogenic or likely pathogenic variants in *BCL11B* were examined (Individual #1-20, Supplemental Table 1) and compared with a series of 29 individuals from the

literature [1] [20-29]. We assembled front and lateral view photographs for 15 and 12 individuals with *BCL11B*-RD respectively. Clinical phenotype was evaluated by analyzing health and medical data systematically collected by using a standardized form sent to collaborators. Compliance with Human Genome Variation Society (HGVS) nomenclature has been verified using batch validation tool of VariantValidator (<https://variantvalidator.org/service/validate/batch/>).

Clinical and molecular information for 7 unreported individuals with variant of uncertain significance (VUS) in *BCL11B* and neurodevelopmental delay were also gathered (VUS #1-7, Supplemental Table 1) but not included in the study. Detailed flow chart of data collection is available in Supplemental Figure 1.

Clinical and molecular data were collected using clinical networks including "Groupe Déficience Intellectuelle France, Anomalies du Développement et Déficience Intellectuelle de Cause Rare" (<http://anddi-rares.org/>) and "European Reference Network ITHACA Congenital Malformations and Rare Intellectual Disability" (<https://ern-ithaca.eu/>) or by directly contacting clinicians from centers in France, Italy and USA.

Previously reported individuals were retrieved using the terms "BCL11B", "BCL11B-related disorder" or "IDDSFTA" in articles indexed in PubMed between inception and April 1st 2023. The full texts of manuscripts were reviewed. The inclusion criteria for published individuals in this study were as followed: availability of clinical information; unambiguous description of the *BCL11B* variant; and no duplication in any previous report.

ii. DNA sample collection

For 16 individuals with pathogenic or likely pathogenic variant in *BCL11B* from this series (Individual #1-4, #6-7, #11-20, Supplemental Table 1) we obtained a DNA sample from previous genetic studies, which included a minimum sample of 0.5 µg (1.76×10^{-8} oz) of DNA from peripheral blood with information on the individual's gender and age at the time of sample collection. Two additional samples (Individual #21-22, Supplemental Table 1) were extracted from previously reported individuals whose clinical and molecular information were not included in the study [25]. Detailed flow chart of DNA sample collection is available in Supplemental Figure 1.

iii. Statistical analyses

For calculating frequencies of features, we excluded individuals with feature coded as “unknown” (*i.e.* when the presence or absence of a particular feature was not clearly documented or when a feature might not be applicable owing to the individual’s gender or age) in the clinical form.

Absolute and relative frequencies (expressed as n [%], n being the number of individuals with the feature, and the percentage was calculated as $n/\text{number of individuals with information on the clinical feature}$) were used for describing categorical variables. Median and interquartile range were used for describing continuous variables and calculated using Microsoft Excel software.

iv. Composite gestalt of *BCL11B*-RD

We used the Face2Gene online tool (<https://www.face2gene.com/>) to generate a computed gestalt of individuals with *BCL11B*-RD and to assess the Rubinstein-Taybi-like facial features of individuals from this series (Supplemental Figure 2).

B. Immunophenotyping

Immunophenotyping analyses were performed using flow cytometry of peripheral blood mononuclear cells (PBMCs). CD4⁺ and CD8⁺ T cells were sorted by antibody detection. T cells subpopulations were defined as follows: activated T cells (CD3⁺/HLA-DR⁺), recent thymic emigrants (RTE, CD4⁺/CD45RA⁺/CD31⁺), naïve CD4⁺ T cells (CD45RA⁺/CD4⁺); memory CD4⁺ T cells (CD45RO⁺/CD4⁺); regulatory T cells (Treg, CD25⁺/CD127⁻/CD4⁺); naïve CD8⁺ T cells (CD45RA⁺/CCR7⁺/CD8⁺); effector-memory expressing CD45RA⁻ CD8⁺ T cells (TEMRA, CD45RA⁺/CCR7⁻/CD8⁺), effector memory CD8⁺ T cells (TEM, CD45RA⁻/CCR7⁻/CD8⁺).

C. DNA methylation analysis

Following a similar approach to previously published results, DNA methylation analysis was conducted using Illumina Infinium MethylationEPIC BeadChip microarrays, which cover approximately 860,000 human genomic CpG sites (San Diego, CA) [30]. DNA extracted from peripheral blood samples was supplied to the microarrays after bisulfite conversion, following the manufacturer’s guidelines. The

detailed procedure for DNA methylation analysis and profiling has been described in previous publications [16-17][30]. The process can be summarized as follows: The resulting methylated and unmethylated signal intensities were analyzed using R version 4.1.2. The mini package version 1.40.0 was utilized for Illumina normalization with background correction [31]. Probes that contained SNPs near or at the CpG interrogation or single nucleotide extension locations, were located on the X or Y chromosomes, cross-reacted with chromosomal locations other than their target regions, or were suggested by Illumina to be excluded, were eliminated from the downstream data analysis. Additionally, arrays with high probe failure rates (> 5%) were removed. Principal component analysis (PCA) was used to visualize the overall structure of batches and to identify any outliers.

D. Matched controls selection and DNA methylation profiling

In this study, two epesignatures were identified for *BCL11B*-RD, one with higher sensitivity and the other with higher specificity. The initial cohort comprised 25 samples, consisting of 13 males and 12 females. Among these samples, 18 individuals had pathogenic (n = 15) or likely pathogenic variants (n = 3) in *BCL11B*. Seven individuals with VUS in *BCL11B* and neurodevelopmental delay (VUS #1-7, Supplemental Table 1) were also tested for epesignature but not included in the analysis (Supplemental Figure 1). To ensure the highest possible sensitivity, all 18 samples with pathogenic or likely pathogenic variants were included for methylation profiling. Using the MatchIt package version 4.3.3 [32], 54 control samples from the in-house database at LHSC, called the EpiSign Knowledge Database (EKD, <https://episign.lhsc.on.ca/index.html>), were selected. These control samples were matched to the case samples based on age, sex, and array type.

Methylation levels or β -values (ranging from 0 (no methylation) to 1 (full methylation)) were measured and converted *M*-values using logit transformation. The *M*-values were then used for linear regression modelling using the limma package version 3.50.0 [33]. This modelling approach created mean methylation differences between the case and control groups and *p*-values for each probe.

To account for confounding factors, Houseman's method was employed to estimate blood cell proportions [34]. These proportions were included in the model matrix

as confounding variables. Additionally, the batch information was added as another confounding variable to the model matrix. To moderate the resulting p -values, the eBayes function from the limma package was applied. Subsequently, multiple testing correction using the Benjamini-Hochberg (BH) method was performed to control for false discovery rate. A subset of probes was selected as the *BCL11B*-RD episignature based on the following criteria: The $i = 800$ probes with the highest product of mean methylation differences and the negative logarithm of p -values were retained. Among those probes, the $j = 200$ probes with the highest areas under the receiver operating characteristic curve (AUROC) were selected as the episignature probes. The robustness of the selected probes in distinguishing between the *BCL11B*-RD and control groups was assessed through the following methods: Hierarchical clustering: Ward's method was applied to calculate the Euclidean distance between samples, and hierarchical clustering was performed. Multidimensional scaling (MDS): Pairwise Euclidean distances between samples were scaled, and MDS was conducted.

To examine the reproducibility of the episignature, 18 iterations of leave-one-out cross-validation (LOOCV) were performed. At each iteration, one *BCL11B*-RD sample was excluded from the probe selection process. An MDS analysis was then conducted on the remaining selected probes, and the left-out sample was used for testing. To obtain a more specific episignature, the following six samples, which demonstrated an intermediate pattern, were selected to be removed from the probe selection set (Individuals #11-12-13-15-20-21, Supplemental Table 1) and above procedure was repeated.

E. Construction of the binary classification model

The e1071 package was utilized to construct a support vector machine (SVM) classifier, following the methodology described in previous publications [16-17][30]. The model was constructed based on the selected set of probes, by training the 18 samples with pathogenic *BCL11B* variants, against the matched control samples and 75% of other control samples and individuals with ~60 other genetic disorders with known episignatures from the EpiSign v3 classifier. The remaining 25% of the samples from the EpiSign v3 classifier database were used. The model generates methylation variant pathogenicity (MVP) scores that range from 0 (indicating the lowest similarity to the identified episignature) to 1

(indicating the highest similarity to the identified episignature). Applying a methodology akin to that outlined for the initial episignature, a second episignature was formulated and a classifier was established by utilizing 12 out of the 18 samples containing pathogenic variants for training. A sole additional criterion was applied to the episignature derived from these 12 samples, involving the exclusion of probes exhibiting pair-wise Pearson's correlation coefficients surpassing 0.9. This refinement led to a more optimal outcome.

F. Functional genomic assessment and methylome overlap

In accordance with the methodology outlined in the study of Levy and colleagues [35], the study aimed to determine the degree of overlap between the differentially methylated probes (DMPs) associated with *BCL11B*-RD and those associated with other disorders with previously published episignatures from the EpiSign v3 classifier within the EpiSign Knowledge Database (EKD) [36]. Control samples were selected from a pool of samples with no known disorder associated with an episignature. To select the DMPs, only probes that were common between the 450k and EPIC arrays were considered. The criteria for selection were as follows: mean methylation differences between the case and control groups had to be greater than 0.05, and the corrected *p*-values had to be less than 0.01. The investigation and visualization of the overlap between the DMPs of *BCL11B*-RD and other disorders' DMPs within the EKD were conducted using these methods, as detailed in the mentioned publications.

To annotate the DMPs, the annotation package version 1.20.0 was utilized [37]. AnnotationHub version 3.2.1, along with the annotations hg19_cpgs, hg19_basicgenes, hg19_genes_intergenic, and hg19_genes_intronexonboundaries, were employed for the annotation process. The TreeAndLeaf package version 1.6.1 was employed to visualize the similarity in methylation levels of differentially methylated probes (DMPs) among various disorders. This previously detailed approach facilitated the identification of potential relationships and clusters among the disorders through a tree-and-leaf diagram [35-36].

V. Results

A. Clinico-biological data

We have collected clinical data for 20 individuals with *BCL11B*-RD. Main clinical features are summarized in Table 1. Across this series, growth parameters were in the low normal range, with mean height of -0.2 SD, mean weight of -0.2 SD and mean head circumference of -0.8 SD. From a general development standpoint, all individuals displayed a global neurodevelopmental delay associated with an intellectual disability (ID) in 16 individuals out of 20. Motor skills were mostly preserved with mean age at walking independently of 18 months. Oppositely, first words were delayed with mean of 18.5 months.



Figure 4. Front and lateral views of individuals with *BCL11B*-related disorder. (A) Front view of 15 individuals with *BCL11B*-RD. Individuals #1, #2, #3 and #4 are shown at two different age period. **(B)** Lateral view of 12 individuals with *BCL11B*-RD.

Beyond the development of first words, communication could also be delayed as 6 individuals aged over 4 still do not use full sentences (*i.e.* single words or words association). Besides early walking milestones, 4 individuals displayed additional motor features including ataxic gait ($n = 2$) and pyramidal syndrome ($n = 2$). Psychiatric comorbidity was also frequent, including anxiety ($n = 8$) and auto/hetero-aggressive behaviour ($n = 7$). Autism spectrum disorder (ASD) was detected in 5 of 20 individuals and 9 individuals displayed sleep disturbance. Brain MRI was mostly normal when performed.

Table 1. Clinical features of individuals with *BCL11B*-related disorder in this series and comparison with the literature. ^a The number of individuals is detailed for every feature, and their frequencies/distributions were calculated according to that number. ^b Craniosynostosis included oxycephaly, scaphocephaly and plagiocephaly. ^c Dental anomalies included tooth agenesis, dental malposition and anterior overbite. ^d Anomaly of ear morphology included low-set ears and protruding ears. ^e Abnormal brain MRI included cerebellar tonsillar ectopia, vermian hypoplasia and delayed myelination. ^f Autoimmune disorder included Hashimoto thyroiditis, Graves' disease, idiopathic juvenile arthritis and crescentic glomerulonephritis.

Clinical features ^a	This series	Literature	Global
General information			
Gender (males/females, <i>n</i> [%]) (N = 20 + 29)	11 (55)/9 (45)	16 (55.2) /13 (44.8)	27 (55.1)/22 (44.9)
Age at last examination (years, m[IQR]) (N = 20 + 29)	12 [2.3;44]	10 [1;35]	10.5 [1;44]
LoF molecular variant, <i>n</i> (%) (N = 18 + 23)	16 (88.9)	14 (60.9)	30 (73.2)
<i>De novo</i> molecular variant, <i>n</i> (%) (N = 18 + 23)	17 (94.4)	16 (69.6)	33 (80.5)
Heterozygosity, <i>n</i> (%) (N = 20 + 29)	20 (100)	29 (100)	49 (100)
Growth parameters			
Height at last examination (SD, m[IQR]) (N = 20)	-0.2 [-3;3.5]	-	-
Weight at last examination (SD, m[IQR]) (N = 20)	-0.2 [-2;3.5]	-	-
Head circumference at last examination (SD, m[IQR]) (N = 20)	-0.8 [-3.6;3]	-	-
Developmental features			
Failure to thrive, <i>n</i> (%) (N = 20 + 9)	0 (0)	1 (11.1)	1 (3.4)
Neurodevelopmental delay, <i>n</i> (%) (N = 20 + 25)	20 (100)	25 (100)	55 (100)
Developmental regression, <i>n</i> (%) (N = 20 + 25)	1 (5)	2 (8)	3 (5.4)
Age at walking independently (months, m[IQR]) (N = 20)	18 [12;30]	-	-
Age at first words (months, m[IQR]) (N = 14)	18.5 [12;53]	-	-
Intellectual disability, <i>n</i> (%) (N = 20 + 25)	16 (80)	22 (88)	38 (69)
Ability to read, <i>n</i> (%) (N = 18)	9 (50)	-	-
Autism spectrum disorder, <i>n</i> (%) (N = 20 + 25)	5 (25)	6 (24)	11 (24.4)
Psychiatric features			
Attention deficit hyperactivity disorder, <i>n</i> (%) (N = 20)	2 (10)	-	-
Anxiety disorder, <i>n</i> (%) (N = 20 + 13)	8 (40)	1 (7.7)	9 (27.3)
Auto/heteroaggressivity, <i>n</i> (%) (N = 20)	7 (35)	-	-
Craniofacial features			
Craniosynostosis ^b , <i>n</i> (%) (N = 20 + 29)	4 (20)	6 (20.7)	10 (20.4)
Myopathic facial appearance, <i>n</i> (%) (N = 20 + 25)	3 (15)	9 (36)	12 (26.6)
Thin eyebrows, <i>n</i> (%) (N = 20 + 25)	13 (65)	14 (56)	27 (60)
High arched eyebrows, <i>n</i> (%) (N = 20 + 25)	6 (30)	3 (12)	9 (20)
Small palpebral fissures, <i>n</i> (%) (N = 20 + 25)	7 (35)	12 (48)	19 (42.2)
Semi-lunar shape of eyelids when smiling, <i>n</i> (%) (N = 20)	6 (30)	-	-
Hypertelorism, <i>n</i> (%) (N = 20 + 25)	3 (15)	11 (44)	14 (31.1)

Clinical features ^a	This series	Literature	Global
Broad nasal bridge, n (%) (N = 20 + 25)	4 (20)	3 (12)	7 (15.5)
Bulbous nose tip, n (%) (N = 20 + 25)	10 (50)	14 (56)	24 (53.3)
Beaked nose, n (%) (N = 20)	8 (40)	-	-
Long columella, n (%) (N = 20)	8 (40)	-	-
Long philtrum, n (%) (N = 20 + 25)	11 (55)	20 (80)	31 (68.8)
Smooth philtrum, n (%) (N = 20)	12 (60)	-	-
Thin upper lip vermillion, n (%) (N = 20 + 25)	16 (80)	16 (64)	32 (71.1)
Dental anomalies ^c , n (%) (N = 20 + 25)	5 (25)	11 (44)	16 (35.6)
Anomaly of ear morphology ^d , n (%) (N = 20 + 25)	3 (15)	3 (12)	6 (13.3)
Hands and feet features			
Broad/large thumb, n (%) (N = 20)	1 (5)	-	-
Laterally deviated thumb, n (%) (N = 20)	0 (0)	-	-
Broad/large hallux, n (%) (N = 20)	1 (5)	-	-
Laterally deviated hallux, n (%) (N = 20)	2 (10)	-	-
Flexion deformity of finger, n (%) (N = 20)	3 (15)	-	-
Neurological features			
Seizures, n (%) (N = 20 + 25)	0 (0)	2 (8)	2 (4.4)
Pyramidal syndrome, n (%) (N = 20)	2 (10)	-	-
Ataxic gait, n (%) (N = 20)	2 (10)	-	-
Sleep disturbance, n (%) (N = 20)	9 (45)	-	-
Abnormal brain MRI ^e , n (%) (N = 15 + 18)	3 (20)	3 (16.7)	6 (18.2)
Ophthalmological features			
Myopia, n (%) (N = 20 + 20)	2 (10)	2 (10)	4 (10)
Hyperopia, n (%) (N = 20 + 20)	4 (20)	4 (20)	8 (20)
Astigmatism, n (%) (N = 20 + 20)	1 (5)	2 (10)	3 (7.5)
Strabismus, n (%) (N = 20 + 25)	4 (20)	1 (4)	5 (11.1)
Immunological features			
Recurrent infections, n (%) (N = 20 + 29)	10 (50)	9 (31)	19 (38.8)
Asthma, n (%) (N = 20 + 29)	7 (35)	10 (34.5)	17 (34.7)
Autoimmune disorder ^f , n (%) (N = 20 + 25)	3 (15)	1 (4)	4 (8.9)

Craniofacial development was also affected as a craniosynostosis was found in 4 of 20 individuals. Most frequent facial features were thin upper lip vermillion, long and smooth philtrum, thin eyebrows, bulbous nose tip, broad nasal bridge and long columella (Figure 4). Refractive error and strabismus were observed in 7 and 4 individuals respectively.

From an immunological standpoint, a significant and commonly encountered medical issue was recurrent non-severe infections, which were present in 10 of 20 individuals. Strikingly, autoimmune disorder was found in 3 individuals, including Hashimoto thyroiditis, Graves' disease and idiopathic juvenile arthritis in

association with crescentic glomerulonephritis. Seven individuals further presented allergic asthma.

Considering the crucial role of *BCL11B* in the production of T cell lineage, we collected immunophenotyping data from 10 individuals with *BCL11B*-RD [1][38]. Haemoglobin, haematocrit, platelets and leukocytes counts were normal in all individuals. Immune profiling of the T cell compartment was found abnormal for 4 individuals as they all displayed low levels of naïve CD4⁺ T cells and recent thymic emigrants (RTE) in association with increased percentage of CD4⁺ memory T cells and effector-memory expressing CD45RA CD8⁺ T cells (TEMRA) (Table 2). Additional features were identified including moderate T cell lymphocytosis with preserved CD4/CD8 ratio (n = 2), increased percentage of activated T cells (n = 2) and reduced levels of naïve CD8⁺ T cells (n = 3). Of note, T regulatory cells, NK cells and serum immunoglobulins were normal in all 10 individuals (data not shown).

Table 2. Immunophenotyping profiles of 4 individuals with *BCL11B*-related disorder associated with abnormal T cell compartment. Values above and below the normal range are marked in blue and orange respectively. Normal ranges are expressed according to individual's gender and age.

Cells	Individual #2	Individual #4	Individual #14	Individual #15
T cells (cells/mm ³)	3500 [1239-2611]	1187 [954-2332]	3200 [807-1844]	1028 [807-1844]
T cells (%)	77 [63-78]	80 [63-80]	86 [64-85]	80 [64-85]
CD3 ⁺ /CD4 ⁺ (cells/mm ³)	1702 [646-1515]	584 [610-1446]	2280 [460-1232]	466 [460-1232]
CD3 ⁺ /CD4 ⁺ (%)	38 [32-47]	40 [33-52]	62 [34-62]	36 [34-62]
CD3 ⁺ /CD8 ⁺ (cells/mm ³)	1273 [365-945]	461 [282-749]	1316 [187-844]	407 [187-844]
CD3 ⁺ /CD8 ⁺ (%)	30 [17-30]	30 [19-30]	36 [14-42]	33 [14-42]
CD4/CD8 ratio	1.34 [1-4.9]	1.27 [1-4.9]	1.73 [1-4.9]	1.13 [1-4.9]
CD3 ⁺ /HLA-DR ⁺ (%)	8 [2-7]	5 [2-9]	5 [2-12]	26 [1-8]
CD4 ⁺ /CD45RA ⁺ /CD31 ⁺ [RTE CD4 ⁺] (%)	31 [43-55]	23 [43-55]	34 [43-55]	7 [20-28]
CD45RA ⁺ /CD4 ⁺ [naïve T CD4 ⁺] (%)	51 [53-74]	32 [41-66]	33.5 [34.5-61]	20 [34.5-61]
CD45RO ⁺ /CD4 ⁺ [memory T CD4 ⁺] (%)	49 [22-37]	71 [25-52]	57 [16-34]	78 [16-34]
CD25 ⁺ /CD127 ⁻ /CD4 ⁺ [Treg] (%)	6 [4-10]	7 [4-10]	4 [4-10]	6 [4-10]
CD45RA ⁺ /CCR7 ⁻ /CD8 ⁺ [naive T CD8 ⁺] (%)	22 [52-68]	4 [52-68]	39.5 [52-68]	45 [29-57]
CD45RA ⁺ /CCR7 ⁻ /CD8 ⁺ [TEMRA CD8 ⁺] (%)	69 [16-28]	72 [16-28]	42 [16-28]	48 [11-26]
CD45RA ⁻ /CCR7 ⁻ /CD8 ⁺ [TEM CD8 ⁺] (%)	7 [11-20]	24 [11-20]	17 [11-20]	7 [20-41]

Four pathogenic chromosomal microdeletions were detected by array CGH analysis (Figure 5A). Fourteen heterozygous germline molecular variations, including 11 previously unreported pathogenic or likely pathogenic variants, were identified in this series using targeted genes panel, exome or genome sequencing (Figure 5B). Majority of genetic variations occurred *de novo*, with the exception of the p.(Lys825*) variant inherited from a healthy mother with germline mosaicism (Individual #20, Supplemental Table 1).

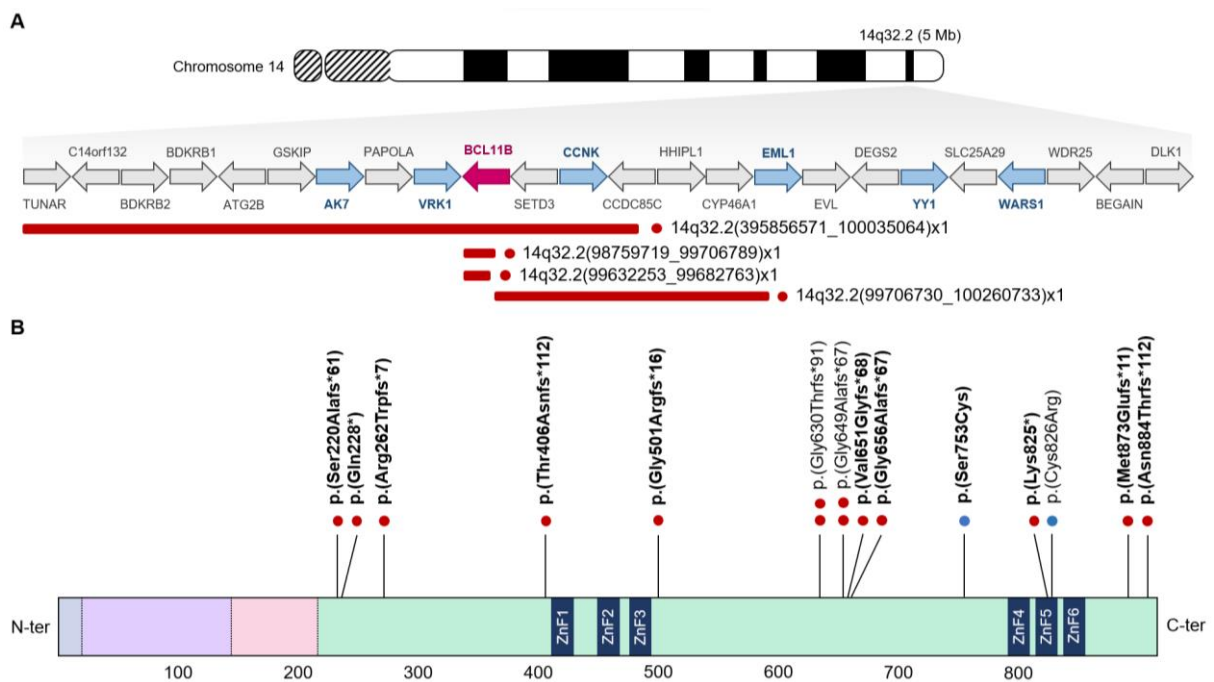


Figure 5. *BCL11B* genotypic spectrum in this series. (A) Array CGH analysis identified four heterozygous copy number variations responsible for complete or partial deletion of *BCL11B*. Additional OMIM genes are shown in blue. (B) Pathogenic and likely pathogenic single nucleotide variations in *BCL11B* (NM_138576.4) were identified by next generation sequencing including 12 loss-of-function (red) and 2 missense (blue) variants (GRCh37/hg19). Newly reported variants are shown in bold.

B. Episignature discovery

Two episignatures were identified for *BCL11B*-RD, one with higher sensitivity including samples with robust and milder methylation profiles (*i.e.* accurate

episignature) and one with higher specificity, including only samples with robust methylation profiles (*i.e.* subset specific episignature).

i. Accurate episignature

The identification and validation of the episignature for *BCL11B*-RD involved the use of 18 samples with pathogenic (n = 15) or likely pathogenic (n = 3) *BCL11B* variants and 54 matched control individuals. A total of 200 probes were selected as the identifying episignature for *BCL11B*-RD based on their ability to distinguish between case samples and controls. The robustness of these selected probes in differentiating case samples from controls was assessed using hierarchical clustering and multidimensional scaling (MDS), as depicted in Figures 6A and 6B, respectively. Furthermore, the reproducibility of the episignature was investigated through 18 iterations of LOOCV, as shown in Supplemental Figure 3A. This process involved excluding one *BCL11B*-RD sample at each iteration and using the remaining samples for probe selection. The excluded sample was then tested to observe if it clustered correctly with the remaining case samples, ensuring the consistency and reliability of the episignature. It is worth noting that out of the 18 *BCL11B*-RD samples, 6 samples (Individuals #11-12-13-15-20-21) exhibited an intermediate methylation pattern. These samples were excluded from the probe selection process in the discovery of the next episignature, likely to obtain a more specific episignature for *BCL11B*-RD. Based on the selected probe set, an SVM classifier was constructed. The SVM was trained using the 18 *BCL11B*-RD samples, matched controls, and 75% of other controls and EpiSign v3 classifier samples. The remaining 25% of the samples were used for testing the model. The resulting MVP scores were visualized in Figure 6C.

In Figure 6C, it can be observed that the majority of testing samples from the EpiSign v3 classifier received MVP scores near 0, indicating low similarity to the identified episignature for *BCL11B*-RD. This suggests that the model effectively distinguishes *BCL11B*-RD samples from other disorders with known episignatures. The model's specificity, calculated using a cohort of unresolved samples and a cut-off MVP score of 0.25, was found to be 99.96%. This indicates that the model has a high ability to correctly classify control samples and minimize false positive results. Among the 6 samples with an intermediate methylation pattern on the clustering plots, 5 of them also received lower MVP scores, as shown in Figure 6C.

This suggests that the model's scores align with the methylation patterns observed in these samples, providing additional support for its accuracy and reliability.

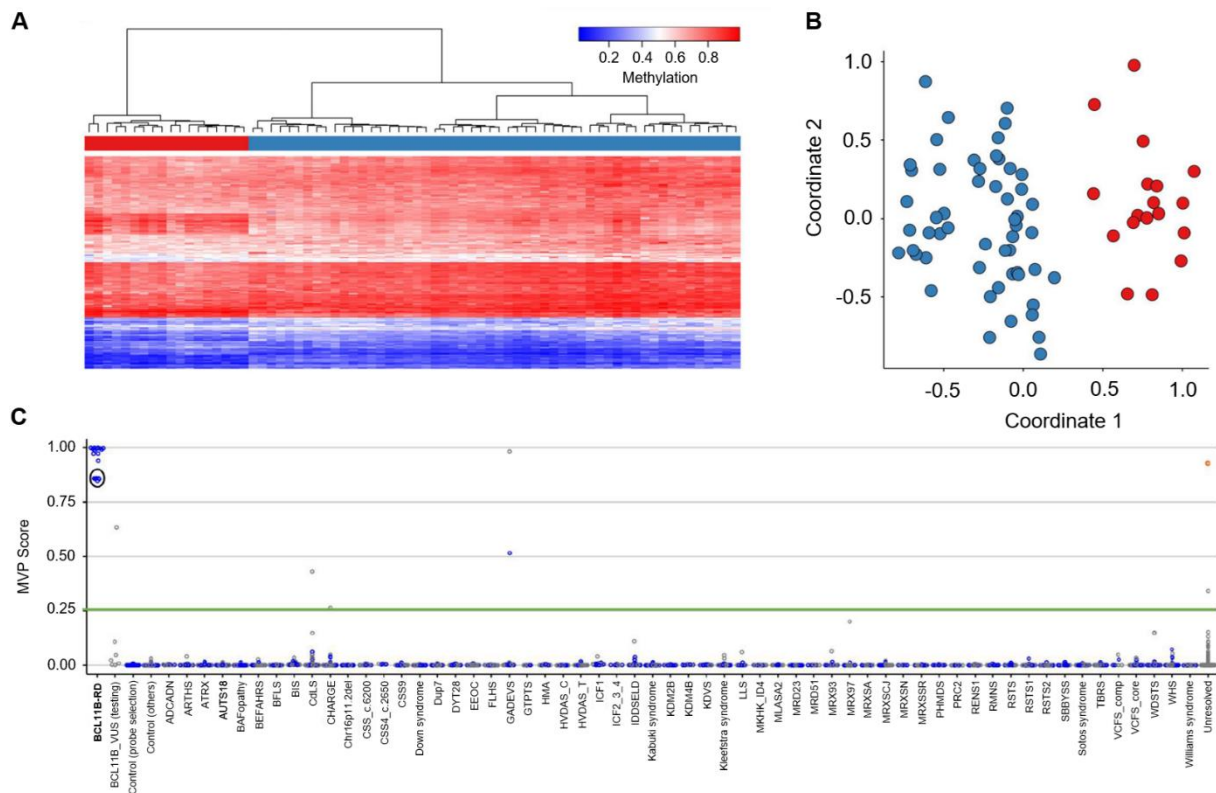


Figure 6. Assessment of the robustness of the epignature with higher sensitivity. **(A)** Hierarchical clustering was performed using the selected probes, with rows and columns representing probes and samples, respectively. On the heatmap pane, red and blue represent *BCL11B*-RD and matched control samples, respectively. The colour scale illustrates methylation levels ranging from blue (0, no methylation) to red (1, full methylation). **(B)** MDS was conducted, with red and blue circles representing *BCL11B*-RD and control samples, respectively. Both the heatmap and the MDS plots indicate a clear separation between the case and control groups, confirming the robustness of the selected probe set. **(C)** The MVP scores generated by the classification model were visualized, with blue and grey circles representing the training and testing samples, respectively. The orange circle represents individual #8 which received an MVP score of 0.93. One GADEVIS sample also received an MVP score close to 1. Among the *BCL11B*-RD training samples, majority of samples with an intermediate LOOCV methylation pattern (5 out of 6) also received a lower MVP score, indicated by a circular boundary around them. All MVP scores are fully described in Supplemental Table 1.

ii. Subset specific episignature

After removing the 6 samples that did not cluster completely with *BCL11B*-RD cases in the previous analysis, probe selection was performed using the remaining 12 *BCL11B*-RD samples and 60 matched control individuals. The robustness and reproducibility of the 130 identified probes were confirmed through hierarchical clustering, MDS, and 12 iterations of LOOCV, as shown in Figures 7A and 7B.

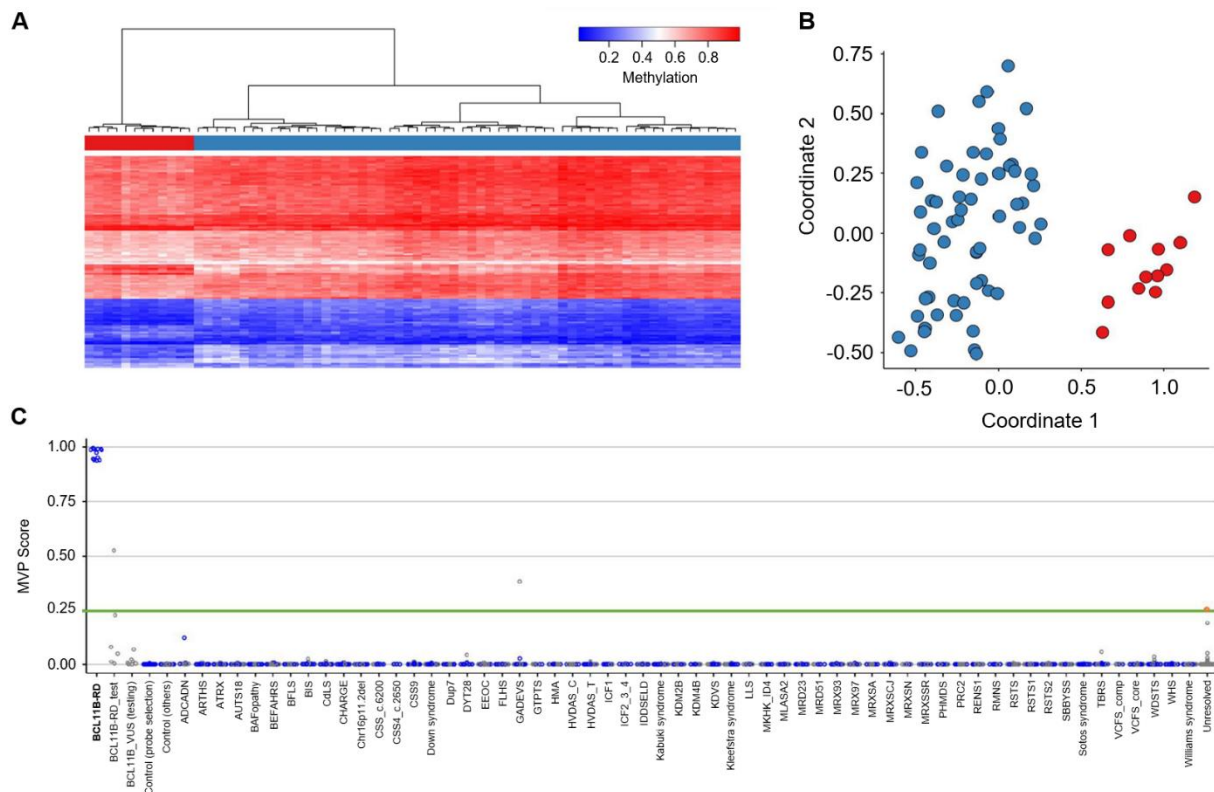


Figure 7. Assessment of the robustness of the episignature with higher specificity. (A-B) Hierarchical clustering and MDS were performed, with case and control samples represented by red and blue, respectively. Both plots demonstrate separate clusters for the case and control samples, thus verifying the robustness of the episignature. **(C)** MVP scores measured using the SVM classifier were plotted, with blue, grey, and orange circles representing the training and testing samples, as well as individual #8, respectively. Individual #8 and the GADEVS subject received MVP scores of 0.25 and 0.38. The group labelled as "*BCL11B*-RD test" represents the 6 samples with an intermediate methylation pattern in the higher sensitivity episignature. All MVP scores are fully described in Supplemental Table 1.

Furthermore, an SVM classifier was constructed using the selected probes set. The SVM was trained using the 12 *BCL11B*-RD samples, matched controls, and 75% of control and other syndromes individuals, with the remaining 25% used for testing the classifier. The performance of the classifier was assessed on the *BCL11B*-RD samples that were not used for probe selection and model construction, including samples with *BCL11B* VUS (Figure 7C, Supplemental Table 1).

The full specificity (100%) of the model suggests that it correctly classifies the control samples and samples from other rare genetic disorders without producing false positives.

C. Functional genomic assessment and methylome overlap

Using the criteria described in the Materials and Methods section, we identified 3983 DMPs between 18 *BCL11B*-RD samples and 54 matched control samples. We compared these DMPs to those found in other disorders classified by the EpiSign v3 classifier, examining the degree of overlap and visualizing similarities in methylation levels in Supplemental Figures 4A and 4B. Additionally, in Supplemental Figure 4C, we demonstrated the global hyper/hypomethylation pattern across all cohorts, including the *BCL11B*-RD cohort, by calculating the average mean methylation differences between the case and control groups for all DMPs within each cohort. Notably, *BCL11B*-RD exhibited an overall hyper-methylated profile. We also annotated the genomic locations of the DMPs in relation to CpG islands and genes (Supplemental Figure 5).

VI. Discussion

A. Clinico-biological description

The majority of clinical manifestations exhibited a congruent distribution between this series and existing literature (Table 1). It is also pertinent to note that no discernible phenotypic differences were discerned between individuals harbouring large deletion of encompassing *BCL11B* and those with *BCL11B* molecular variant.

Growth parameters generally fell within the lower bounds of the normal range, while psychomotor development was delayed for all individuals. Neurodevelopmental disorder, notably intellectual disability (ID) and autism

spectrum disorder (ASD), manifested in 69% and 24.4% of individuals respectively, upon amalgamation of data from this series and literature. Consequently, individuals carrying a pathogenic *BCL11B* variant necessitate focused attention with regard to early developmental milestones, ideally with the involvement of physiotherapists, psychometricians and speech-language pathologists, when deemed necessary.

Anxiety disorders, sleep disturbance, as well as auto/hetero-aggressive behaviours, were seldom documented in this existing literature but exhibited a higher frequency within our series. These interconnected comorbidities often form a self-perpetuating cycle, warranting in-depth investigation through supplementary studies to identify optimal therapeutic strategies.

Refractive errors were also prevalent in this series, implying that meticulous ophthalmological evaluation with regular assessments is crucial for individuals with *BCL11B*-RD.

Within this series, only one individual with craniosynostosis displayed a missense variant (p.(Ser753Cys)), whereas three other individuals carried loss-of-function variant, including two frameshift variations (p.(649Alafs*67); p.(Gly656Alafs*67)) and a single chromosomal microdeletion. Hence, it appears that no substantiated genotype-genotype correlation exists between missense variants in *BCL11B* and the occurrence of craniosynostosis. This leads us to suggest consistent monitoring of craniofacial development, including routine head circumference assessments and the consideration of low-dose CT scans in cases of abnormal skull morphology.

Two novel facial features were identified in our investigation: a beaked nose and semi-lunar eyelid shape upon smiling, present in approximately 40% and 30% of individuals from this series, respectively. Together, these morphological features could inadvertently lead to a misdiagnosis of Rubinstein-Taybi syndrome (RTS). This phenotypic overlap is further corroborated by the misclassification of 4 individuals from this series as RTS by the Face2Gene machine learning software, displaying a high-confidence index (Supplemental Figure 2). Nevertheless, it is important to highlight that significant morphological features of RTS, such as a broad and laterally deviated thumb/hallux, are conspicuously absent in *BCL11B*-RD, as confirmed by the pooled data from this series and existing literature. Therefore, this study underscores *BCL11B*-RD as a differential clinical diagnosis of

RTS, emphasizing the importance of meticulous assessment of hands and feet morphology to discriminate these two distinct genetic conditions. From a broader standpoint, it is reiterated that facial recognition software underpinned by artificial intelligence should be employed with caution, ideally by clinicians well-versed in dysmorphology.

Although recent research has indicated *BCL11B*'s role in preventing autoimmunity by promoting the Treg cell program, instances of autoimmune manifestations in this context of *BCL11B*-RD are sparsely documented within available literature [20][39]. In this study, 3 novel individuals were documented wherein autoimmune disorders including Hashimoto's thyroiditis, Grave's disease, and idiopathic juvenile arthritis associated with crescentic glomerulonephritis, were observed. This outlines the importance of systematically exploring autoimmune symptoms at the time of diagnosis and instituting rigorous monitoring during individual follow-up.

Notably, the immunophenotyping profile of 4 out of 10 individuals displayed abnormalities, characterized by diminished levels of naïve CD4⁺ T cells and RTE associated with increased percentage of TEMRA. These findings align with recent research and underscore the pivotal role of *BCL11B* in regulating the T cell lineage [1][36][40-41]. Hence, we posit that immunophenotyping of peripheral blood lymphocytes serve as a valuable diagnostic tool for *BCL11B*-RD. This tool can be employed either prior to molecular testing in individuals with suggestive morphological features of *BCL11B*-RD or as part of a reverse phenotyping approach when interpreting *BCL11B* variants identified through pangenomic analyses.

B. Epigenetic signature

DNA methylation epesignatures are being utilized as stable and dependable biomarkers for diagnosing rare genetic disorders in clinical labs [19]. For *BCL11B*-RD, 2 distinct epesignatures have been identified, each specifically designed to prioritize a different aspect of diagnostic performance.

The first epesignature was developed to maximize sensitivity, ensuring a high rate of correctly identifying samples with pathogenic or likely pathogenic variants in *BCL11B*. On the other hand, the second epesignature was designed to prioritize specificity, aiming to minimize false positive results. This epesignature focused on selecting probes that exhibited high specificity in distinguishing *BCL11B*-RD cases

from controls, while reducing the inclusion of samples with an intermediate pattern or overlapping characteristics with other disorders.

By having 2 episignatures with distinct emphasis on sensitivity and specificity, researchers can apply the appropriate episignature based on the specific diagnostic needs and goals of a given scenario.

In our study, we initially detected an episignature for a cohort consisting of 18 individuals with pathogenic or likely pathogenic *BCL11B* variants. Using this episignature, we constructed a highly sensitive and specific classifier with a remarkable specificity of 99.96%.

Further analysis was conducted on a cohort of unresolved samples from our database, and Individual #8 received an MVP score of 0.93. Through comprehensive investigation, including the creation of hierarchical clustering and MDS plots, this individual was classified as positive for the *BCL11B*-RD episignature. Subsequent communication with the corresponding physician, who submitted the sample previously for EpiSign analysis and had been stored in the EpiSign Knowledge Database, confirmed that this individual indeed carries a *BCL11B* pathogenic variant, further highlighting the robustness of *BCL11B*-RD episignature.

Another intriguing case in our database received an MVP score close to 1, indicating a strong association with the episignature (Figure 6C). This sample was identified as a Gabriele-de Vries syndrome (GADEVS, MIM 617557) sample and exhibited an 18 megabase deletion at 14q32.2 locus overlapping with both *YY1* and *BCL11B*.

To assess the reproducibility of the episignature, we performed LOOCV. Among the tested samples, 6 displayed an intermediate methylation pattern (Individuals #11-12-13-15-20-21, Supplemental Figure 3A, Supplemental Table 1). Interestingly, 5 of these samples also received lower MVP scores. These samples are marked with a circular boundary in Figure 6C. After excluding the 6 samples with an intermediate methylation pattern from the episignature discovery cohort, we constructed a classifier with full specificity. It is worth noting that individual #8 and the GADEVS individual obtained MVP scores of 0.25 and 0.38, respectively, indicating a slightly lower sensitivity compared to the previous model. However, despite this slight decrease in sensitivity, the classifier still correctly classified these individuals.

Newly discovered *BCL11B*-RD epismature was finally challenged by analyzing the 7 subjects with neurodevelopmental delay and VUS in *BCL11B* (VUS #1-7, Figure 6C, Figure 7C). VUS #6 individual carrying *de novo* methionine-1 variant shown an intermediate pattern with MVP scores of 0.63 of 0.07. However, methylation profiles of VUS #1-5 individuals sharing the familial p.(Glu203Serfs*15) variant did not match any *BCL11B*-RD epismatures. This mismatch could result from an alternative co-segregating variant within a non-OMIM morbid gene, capable of contributing to the familial phenotype or exerting an influence on the epismature signal. Similarly, VUS #7 individual with ID and p.(Pro725His) variant obtained MVP scores of 0. Although occurred *de novo*, this missense variant is not localized in a functional domain of *BCL11B* and further considered as "uncertain" by *in silico* prediction scores (Revel: 0.48; MetaLR: 0.1). Thus, the diagnosis of *BCL11B*-RD can reasonably be excluded for this individual.

In conclusion, the detection of sensitive and specific DNA methylation signatures for *BCL11B*-RD facilitates a more accurate diagnosis of individuals with clinical features evocative of *BCL11B*-RD and VUS in *BCL11B*. Moreover, functional genomic assessment provides further insights into pathophysiology of *BCL11B*-RD and potential molecular pathways shared between *BCL11B*-RD and other genetic disorders.

VII. References

- [1] Lessel D, Gehbauer C, Bramswig NC, Schluth-Bolard C, Venkataramanappa S, van Gassen KLI, et al. BCL11B mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain*. 2018;141(8):2299-311.
- [2] Daher MT, Bausero P, Agbulut O, Li Z, Parlakian A. Bcl11b/Ctip2 in Skin, Tooth and Craniofacial System. *Frontiers in Cell and Developmental Biology*. 2020; 8:581674.
- [3] Zhang L, Juan, Vogel WK, Liu X, Topark-Ngarm A, Arbogast BL, Maier CS, et al. Coordinated regulation of transcription factor Bcl11b activity in thymocytes by the mitogen-activated protein kinase (MAPK) pathways and protein sumoylation. *J Biol Chem*. 2012;287(32):26971-88.
- [4] Cismasiu VB, Adamo K, Gecewicz J, Duque J, Lin Q, Avram D. BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene*. 2005;24(45):6753-64.
- [5] Cherrier T, Suzanne S, Redel L, Calao M, Marban C, Samah B, et al. p21WAF1 gene promoter is epigenetically silenced by CTIP2 and SUV39H1. *Oncogene*. 2009;28(38):3380-9.
- [6] Cherrier T, Le Douce V, Eilebrecht S, Riclet R, Marban C, Dequiedt F, et al. CTIP2 is a negative regulator of P-TEFb. *Proceedings of the National Academy of Sciences*. 2013;110(31):12655-60.
- [7] Phatnani HP, Greenleaf AL. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*. 1 nov 2006;20(21):2922-36.
- [8] Avram D, Califano D. The multifaceted roles of Bcl11b in thymic and peripheral T cells - impact on immune diseases. *J Immunol*. 1 sept 2014;193(5):2059-65.
- [9] Li P, Burke S, Wang J, Chen X, Ortiz M, Lee SC, et al. Reprogramming of T Cells to Natural Killer-Like Cells upon Bcl11b Deletion. *Science*. 2 juill 2010;329(5987):85-9.
- [10] Zhang S, Rozell M, Verma RK, Albu DI, Califano D, VanValkenburgh J, et al. Antigen-specific clonal expansion and cytolytic effector function of CD8+ T lymphocytes depend on the transcription factor Bcl11b. *J Exp Med*. 2 août 2010;207(8):1687-99.

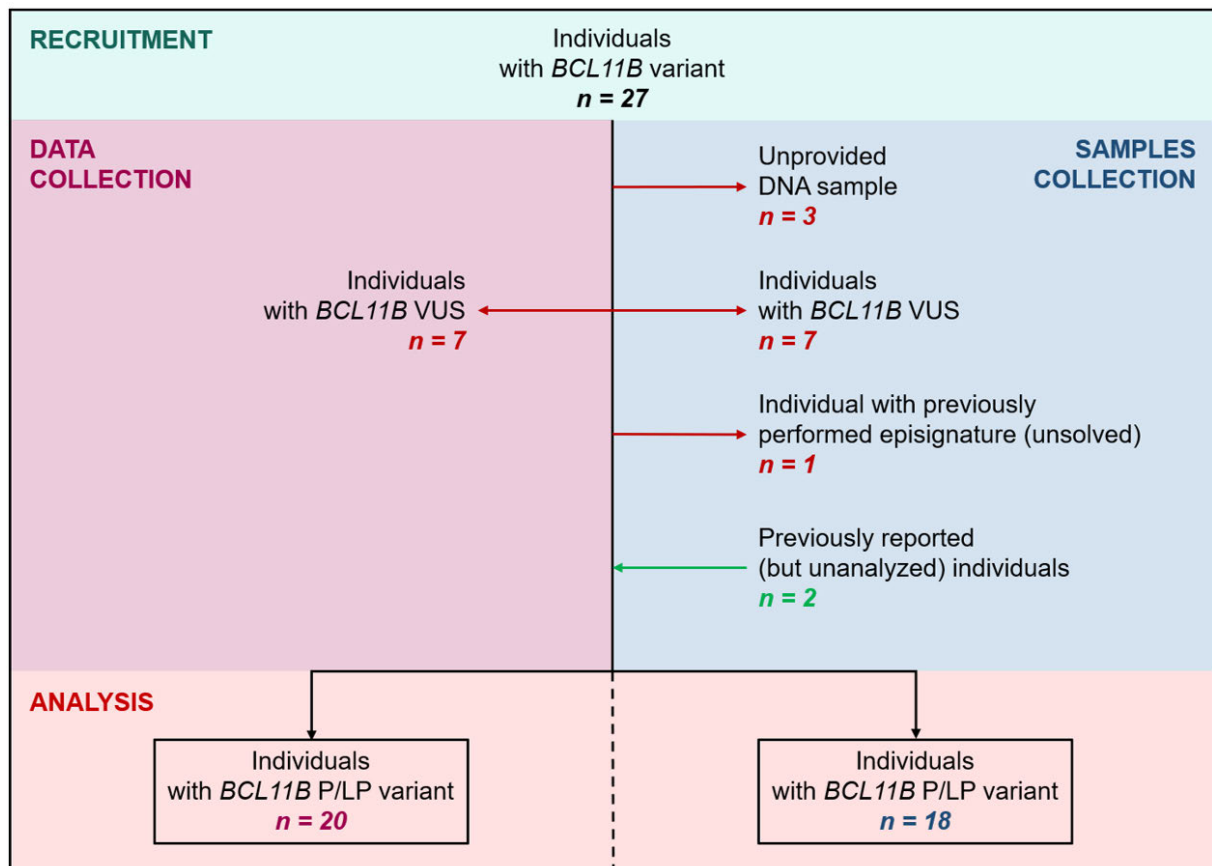
- [11] Kastner P, Chan S, Vogel WK, Zhang LJ, Topark-Ngarm A, Golonzhka O, et al. Bcl11b represses a mature T-cell gene expression program in immature CD4+CD8+ thymocytes. *Eur J Immunol.* août 2010;40(8):2143-54.
- [12] Holmes G, van Bakel H, Zhou X, Losic B, Jabs EW. BCL11B expression in intramembranous osteogenesis during murine craniofacial suture development. *Gene Expr Patterns.* janv 2015;17(1):16-25.
- [13] Kyrylkova K, Iwaniec UT, Philbrick KA, Leid M. BCL11B regulates sutural patency in the mouse craniofacial skeleton. *Dev Biol.* 15 juill 2016;415(2):251-60.
- [14] Simon R, Baumann L, Fischer J, Seigfried FA, De Bruyckere E, Liu P, et al. Structure-function integrity of the adult hippocampus depends on the transcription factor Bcl11b/Ctip2. *Genes Brain Behav.* avr 2016;15(4):405-19.
- [15] Simon R, Brylka H, Schwegler H, Venkataramanappa S, Andratschke J, Wiegrefe C, et al. A dual function of Bcl11b/Ctip2 in hippocampal neurogenesis. *EMBO J.* 29 juin 2012;31(13):2922-36.
- [16] Aref-Eshghi E, Rodenhiser DI, Schenkel LC, Lin H, Skinner C, Ainsworth P, et al. Genomic DNA Methylation Signatures Enable Concurrent Diagnosis and Clinical Genetic Variant Classification in Neurodevelopmental Syndromes. *Am J Hum Genet.* 2018;102(1):156-74.
- [17] Aref-Eshghi E, Bend EG, Colaiacovo S, Caudle M, Chakrabarti R, Napier M, et al. Diagnostic Utility of Genome-wide DNA Methylation Testing in Genetically Unsolved Individuals with Suspected Hereditary Conditions. *Am J Hum Genet.* 2019;104(4):685-700.
- [18] Sadikovic B, Levy MA, Aref-Eshghi E. Functional annotation of genomic variation: DNA methylation epigenatures in neurodevelopmental Mendelian disorders. *Hum Mol Genet.* 2020;29(R1):R27-32.
- [19] Sadikovic B, Levy MA, Kerkhof J, Aref-Eshghi E, Schenkel L, Stuart A, et al. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet Med.* 2021;23(6):1065-74.
- [20] Roa-Bautista A, López-Duarte M, Paz-Gandiaga N, San Segundo Arribas D, Ocejó-Vinyals JG. Deletion in the BCL11B Gene and Intellectual Developmental Disorder with Speech Delay, Dysmorphic Facies, and T-cell Abnormalities - a Case Report. *EJIFCC.* déc 2022;33(4):325-33.
- [21] Zhao X, Wu B, Chen H, Zhang P, Qian Y, Peng X, et al. Case report: A novel truncating variant of BCL11B associated with rare feature of craniosynostosis and global developmental delay. *Front Pediatr.* 2022;10:982361.

- [22] Che F, Tie X, Lei H, Zhang X, Duan M, Zhang L, et al. Identification of two novel variants of the BCL11B gene in two Chinese pedigrees associated with neurodevelopmental disorders. *Front Mol Neurosci*. 2022;15:927357.
- [23] Gaillard L, Goverde A, van den Bosch QCC, Jehee FS, Brosens E, Veenma D, et al. Case Report and Review of the Literature: Congenital Diaphragmatic Hernia and Craniosynostosis, a Coincidence or Common Cause? *Front Pediatr*. 2021;9:772800.
- [24] Lu HY, Sertori R, Contreras AV, Hamer M, Messing M, Del Bel KL, et al. A Novel Germline Heterozygous BCL11B Variant Causing Severe Atopic Disease and Immune Dysregulation. *Front Immunol*. 2021;12:788278.
- [25] Prasad M, Balci TB, Prasad C, Andrews JD, Lee R, Jurkiewicz MT, et al. BCL11B-related disorder in two canadian children: Expanding the clinical phenotype. *Eur J Med Genet*. sept 2020;63(9):104007.
- [26] Yang S, Kang Q, Hou Y, Wang L, Li L, Liu S, et al. Mutant BCL11B in a Patient With a Neurodevelopmental Disorder and T-Cell Abnormalities. *Front Pediatr*. 2020;8:544894.
- [27] Goos JAC, Vogel WK, Mlcochova H, Millard CJ, Esfandiari E, Selman WH, et al. A de novo substitution in BCL11B leads to loss of interaction with transcriptional complexes and craniosynostosis. *Hum Mol Genet*. 1 août 2019;28(15):2501-13.
- [28] Qiao F, Wang C, Luo C, Wang Y, Shao B, Tan J, et al. A De Novo heterozygous frameshift mutation identified in BCL11B causes neurodevelopmental disorder by whole exome sequencing. *Mol Genet Genomic Med*. sept 2019;7(9):e897.
- [29] Punwani D, Zhang Y, Yu J, Cowan MJ, Rana S, Kwan A, et al. Multisystem Anomalies in Severe Combined Immunodeficiency with Mutant BCL11B. *N Engl J Med*. 1 déc 2016;375(22):2165-76.
- [30] Aref-Eshghi E, Kerkhof J, Pedro VP, France GD, Barat-Houari M, Ruiz-Pallares N, et al. Evaluation of DNA Methylation Episignatures for Diagnosis and Phenotype Correlations in 42 Mendelian Neurodevelopmental Disorders. *Am J Hum Genet*. 3 juin 2021;108(6):1161-3.
- [31] Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 15 mai 2014;30(10):1363-9.

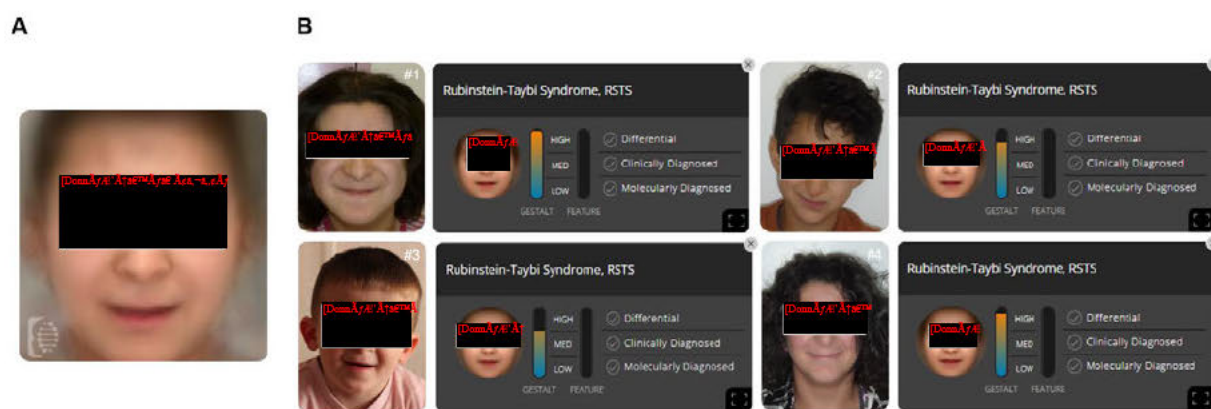
- [32] Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal.* 2007;15(3):199–236.
- [33] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 20 avr 2015;43(7):e47.
- [34] Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 8 mai 2012;13:86.
- [35] Levy MA, Relator R, McConkey H, Pranckeviciene E, Kerkhof J, Barat-Houari M, et al. Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders. *Hum Mutat.* nov 2022;43(11):1609-28.
- [36] Levy MA, McConkey H, Kerkhof J, Barat-Houari M, Bargiacchi S, Biamino E, et al. Novel diagnostic DNA methylation epesignatures expand and refine the epigenetic landscapes of Mendelian disorders. *HGG Adv.* 13 janv 2022;3(1):100075.
- [37] Cao B, Luo H, Luo T, Li N, Shao K, Wu K, et al. The performance of whole genome bisulfite sequencing on DNBSEQ-Tx platform examined by different library preparation strategies. *Heliyon.* 22 mai 2023;9(6):e16571.
- [38] Ikawa T, Hirose S, Masuda K, Kakugawa K, Satoh R, Shibano-Satoh A, et al. An essential developmental checkpoint for production of the T cell lineage. *Science.* 2 juill 2010;329(5987):93-6.
- [39] Drashansky TT, Helm E, Huo Z, Curkovic N, Kumar P, Luo X, et al. Bcl11b prevents fatal autoimmunity by promoting Treg cell program and constraining innate lineages in Treg cells. *Sci Adv.* août 2019;5(8):eaaw0480.
- [40] Sidwell T, Rothenberg EV. Epigenetic Dynamics in the Function of T-Lineage Regulatory Factor Bcl11b. *Front Immunol.* 2021;12:669498.
- [41] Li L, Leid M, Rothenberg EV. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science.* 2 juill 2010;329(5987):89-93.

VIII. Supplemental

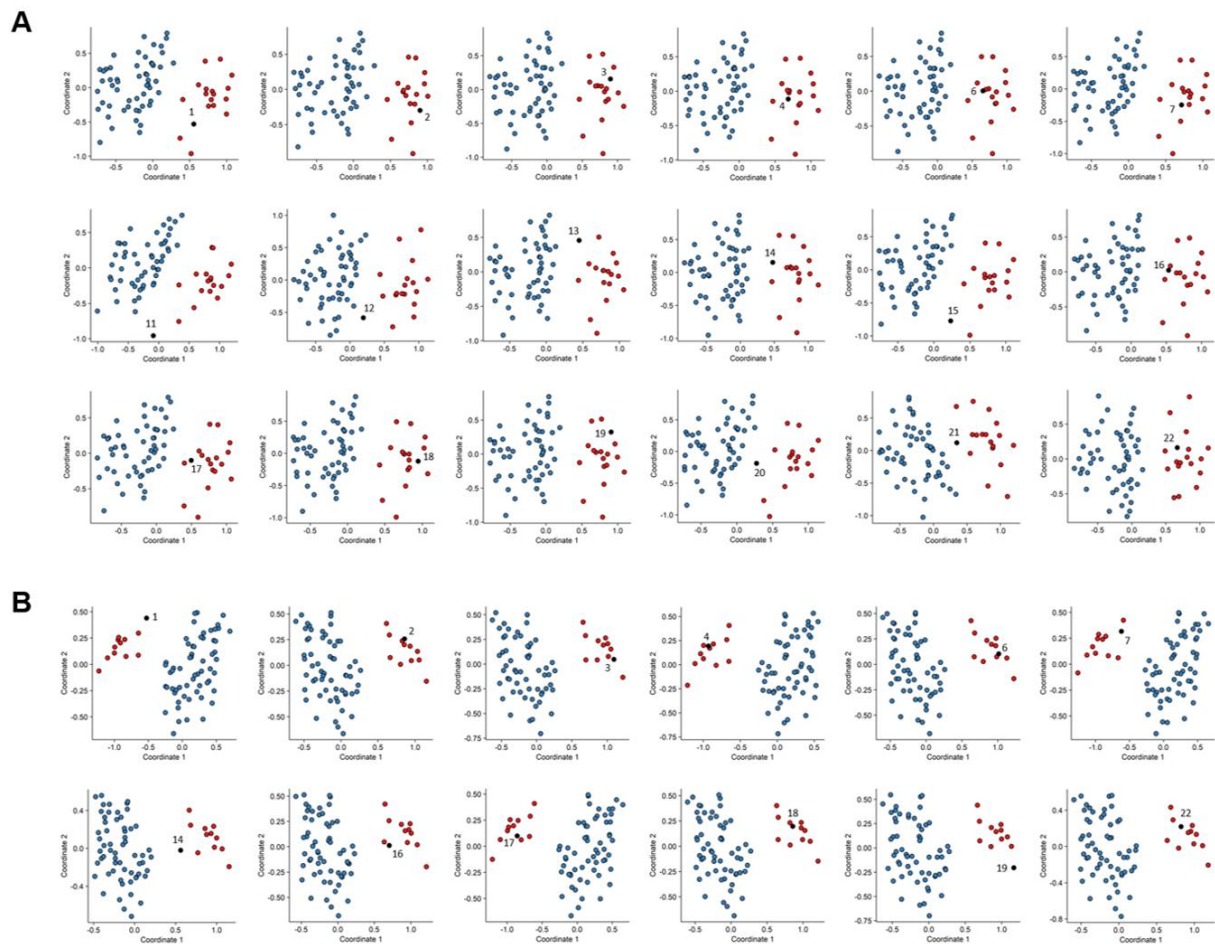
A. Supplemental figures



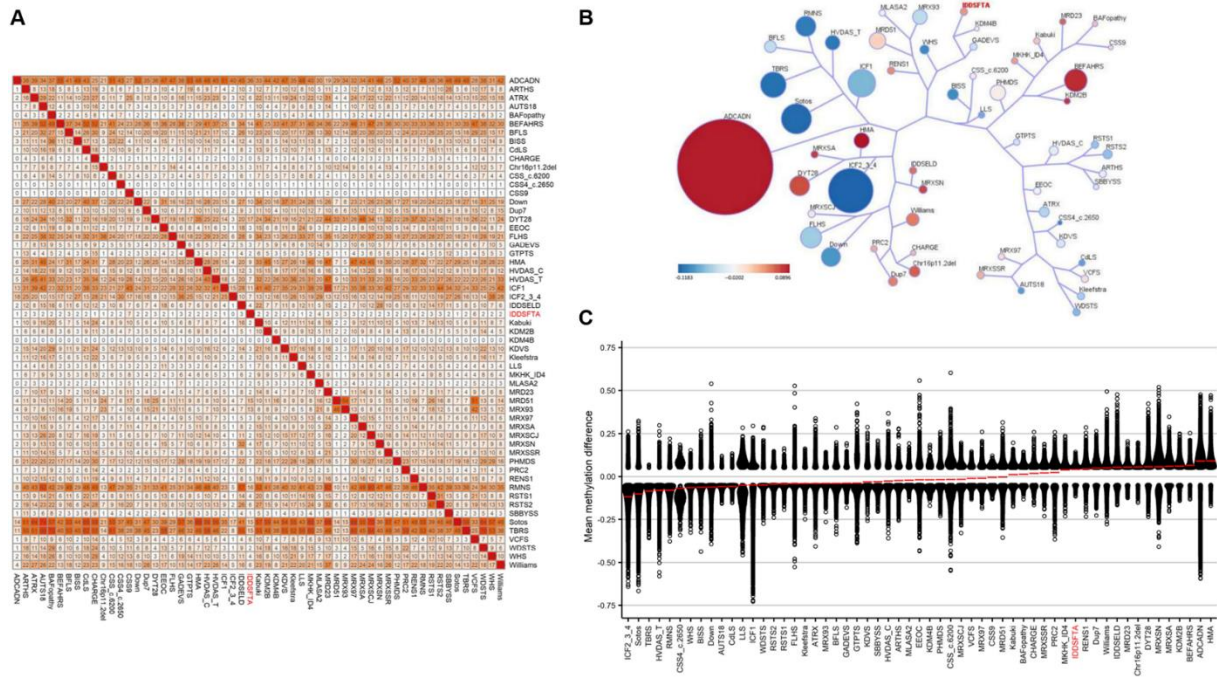
Supplemental Figure 1. Flow chart diagram of the study. Twenty-seven unreported individuals with *BCL11B* variant and neurodevelopmental delay were initially recruited. To ensure accurate clinico-biological refinement of *BCL11D*-RD, 7 individuals were excluded from the clinical and molecular analysis as their variant were classified as VUS (VUS #1-7). On the other hand, the epigenetic discovery was based on DNA methylation analysis of 18 samples from individuals with pathogenic or likely pathogenic *BCL11B* variant: DNA sample was unprovided for 3 individuals (Individual #5-9-10); 7 individuals with *BCL11B* VUS were excluded (VUS#1-7); and 1 individual had previously performed DNA methylation profile and was part of the unresolved cohort (Individual #8). Two additional samples, extracted from previously reported individuals with pathogenic variant in *BCL11B* (Individual #21-22), were included in DNA methylation analysis [25]. Abbreviations: VUS (variant of uncertain significance), P (pathogenic), LP (likely pathogenic).



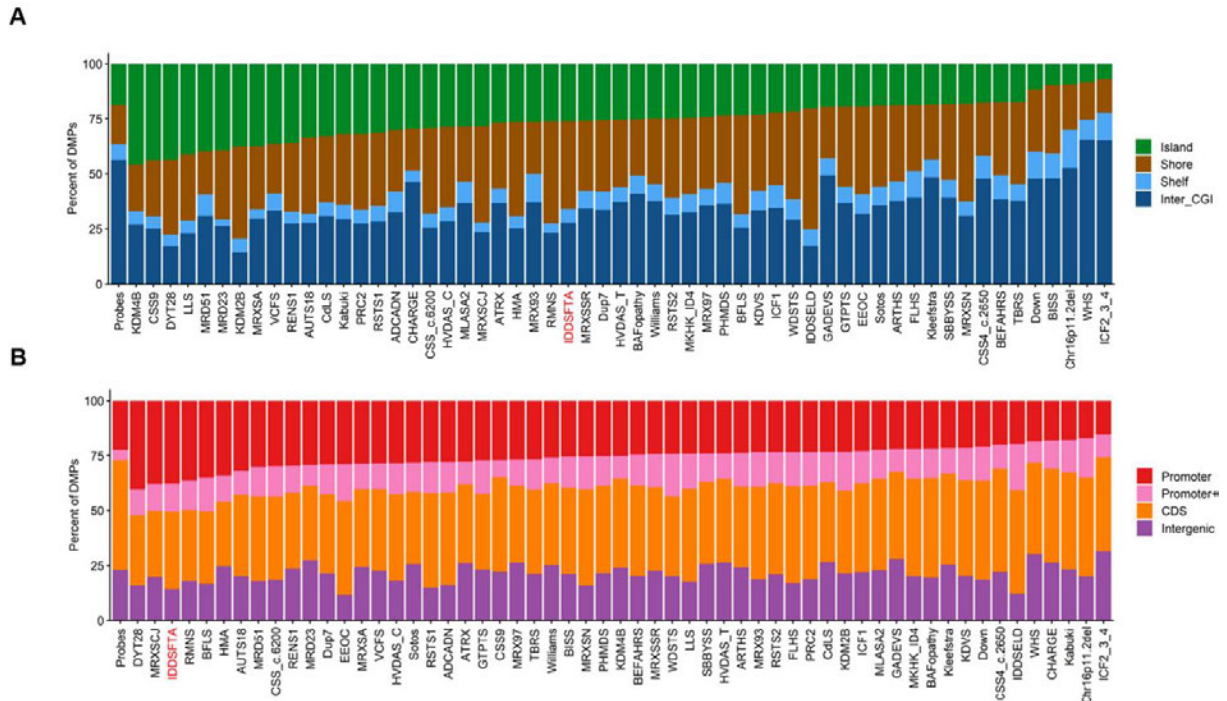
Supplemental Figure 2. Facial features analysis of 4 individuals with *BCL11B*-related disorder (*BCL11B*-RD) using Face2Gene software. (A) Composite gestalt for *BCL11B*-RD. (B) Rubinstein-Taybi syndrome as the top-ranking suggested diagnosis for 4 individuals.



Supplemental Figure 3. LOOCV for both ep signatures. (A) LOOCV for the more sensitive ep signature was performed. In 12 out of 18 iterations, the sample that was left out from probe selection (depicted as a black circle) clustered with the remaining *BCL11B*-RD samples (represented by red circles), separate from the control samples (shown as blue circles). Additionally, in 6 rounds, the testing sample clustered between the case and control groups, indicating that those samples exhibit an intermediate methylation pattern. **(B)** LOOCV for the more specific ep signature was conducted. In all iterations, the left-out sample clustered with the *BCL11B*-RD samples used for probe selection.



Supplemental Figure 4. Comparison of the methylation pattern of *BCL11B*-RD and EpiSign v3 classifier disorders. (A) The color scale, ranging from white to red, represents the percentage of DMPs shared between each pair of disorders, ranging from 0% to 100%. Each entry demonstrates the percentage of DMPs present in the syndrome below that are also present in the DMPs of the syndrome to the right. **(B)** A tree-and-leaf diagram was created to demonstrate the degree of similarity in methylation levels between different disorders. Each node represents a cohort, with the node size being proportional to the number of DMPs in the corresponding cohort divided by the total number of probes. The color of a node illustrates the global mean methylation difference between the case and control groups for that particular cohort. Cohorts located closer to each other on the tree indicate a higher similarity in their methylation levels. **(C)** An overall hypermethylation pattern was observed in the *BCL11B*-RD cohort.



Supplemental Figure 5. Annotation of differentially methylated probes (DMPs). The "probes" column represents the "default" distribution of all probes that remain after the initial probe-filtering process. **(A)** Annotation in the context of CpG islands included the following regions: CpG islands, shores (within 2 kb of a CpG island boundary), shelves (within 2-4 kb of a CpG island boundary), and inter_CGI regions (elsewhere on the genome). The distribution of DMPs was as follows: shores (40%), inter_CGI regions (27%), shelf regions (7%), and islands (6%). **(B)** The DMPs were annotated with respect to genes based on the following regions: promoter (within 1 kb upstream of the transcription start site) and promoter+ (1-5 kb upstream of the transcription start site), coding sequence (CDS), and intergenic regions (other genomic regions). The majority of the DMPs were found in promoter and CDS regions.

B. Supplemental tables

Supplemental Table 1. DNA samples annotation, molecular findings and detailed methylation variant pathogenicity (MVP) scores.

Individual	DNA sample	Variant	ACMG class	MVP score 1	MVP score 2	Reported individual?
#1	MS10504	p.(Val651Glyfs*68)	5	0.99	0.94	No
#2	MS10499	p.(Gly630Thrfs*91)	5	1	0.99	No
#3	MS10508	p.(Gly656Alafs*67)	5	1	1	No
#4	MS10498	p.(Gly630Thrfs*91)	5	1	0.99	No
#5	No sample	p.(Met873Glufs*11)	5			No
#6	MS10509	p.(Gly649Alafs*67)	5	1	0.99	No
#7	MS10510	p.(Cys826Arg)	5	0.97	0.95	No
#8	MS07238	p.(Ser753Cys)	5	0.93	0.25	No
#9	No sample	p.(Ser220Alafs*61)	5			No
#10	No sample	14q32.2 deletion	4			No
#11	MS10502	p.(Gly501Argfs*16)	5	0.86	0.01	No
#12	MS10503	p.(Arg262Trpfs*7)	5	0.86	0.23	No
#13	MS10507	p.(Thr406Asnfs*112)	5	0.97	0.53	No
#14	MS10506	p.(Gln228*)	5	0.86	0.94	No
#15	MS10500	14q32.2 deletion	5	0.86	0.08	No
#16	MS13162	14q32.2 deletion	4	0.94	0.94	No
#17	MS10505	p.(Asn884Thrfs*112)	5	1	0.99	No
#18	MS13170	14q32.2 deletion	4	1	0.95	No
#19	MS13169	p.(Gly649Alafs*67)	5	1	0.99	No
#20	MS13160	p.(Lys825*)	4	0.85	0.05	No
#21	MS12648	p.(Asn807Lys)	5	0.86	0	Yes [25]
#22	MS13074	p.(Gly649Alafs*67)	5	0.99	0.98	Yes [25]
VUS #1	MS13164	p.(Glu203Serfs*15)	3	0.11	0.01	No
VUS #2	MS13166	p.(Glu203Serfs*15)	3	0.02	0.02	No
VUS #3	MS13165	p.(Glu203Serfs*15)	3	0.05	0	No
VUS #4	MS13163	p.(Glu203Serfs*15)	3	0	0	No
VUS #5	MS13167	p.(Glu203Serfs*15)	3	0	0.03	No
VUS #6	MS13161	p.(Met1?)	3	0.63	0.07	No
VUS #7	MS13168	p.(Pro725His)	3	0	0	No

IX. Additional information

A. Data availability

The authors confirm that the data supporting the findings of this study are available within the article and in Figures 1-5 and Supplemental Table 1.

B. Fundings and acknowledgments

Funding for this study was provided, in part, by the London Health Sciences Molecular Diagnostics Innovation and Development Fund and Genome Canada Genomic Applications Partnership Program Grant (Beyond Genomics: Assessing the Improvement in Diagnosis of Rare Diseases using Clinical Epigenomics in Canada, EpiSign-CAN) awarded to B.S. One individual from this series further belongs to the REDIA study, which was supported by a grant from the French Ministry of Health (PHRC-I 18-38). We thank the "European Reference Network ITHACA Congenital Malformations and Rare Intellectual Disability" (<https://ern-ithaca.eu>) and the "Anomalies du Développement Déficience Intellectuelle de causes Rares" network (<https://anddi-rares.org>) for the diffusion of the partnership proposal.

C. Ethics declaration

This research protocol was validated by the institutional review board of the University of Montpellier (September 17th 2020, IRB-MTP_2020_09_202000584); ClinicalTrials.gov Identifier: NCT04541927. Written consent was obtained from all individuals, or their legal guardians, for participation in the study and publication of photographs.

D. Conflict of interest

The authors declare no conflicts of interest.

CONCLUSION

MALADIES RARES :

LE CHALLENGE DE L'INNOVATION DIAGNOSTIQUE

Les maladies rares sont définies par une prévalence inférieure à **1/2000** et concernent **5%** de la population générale dans le monde occidental, soit plus de **3.000.000** de personnes sur le territoire français¹⁹ et **290.000** dans la région Occitanie²⁰. **Huit-mille** maladies rares ont été identifiées à ce jour, avec une origine génétique établie pour plus de **80%** d'entre elles²¹.

Suite à la loi n°2004-806 du 9 août 2004 établissant les maladies rares comme l'une des 5 priorités de la politique de santé publique, la France fut le premier pays européen à se doter d'un **Plan National Maladies Rares** (PNMR)²². Dès lors, 3 PNMR se succédèrent avec comme objectifs l'amélioration de l'accès au diagnostic, l'optimisation du parcours de soins proposé aux patients et la promotion de la recherche dans le domaine des maladies rares^{23,24,25}. Furent ainsi labellisées 23 **filières de santé maladies rares** (FSMR)²⁶, régissant à l'heure actuelle 603 **centres de référence** (CRMR)²⁷ et 1723 **centres de compétence** (CCMR)²⁷ maladies rares sur l'ensemble du territoire français, ainsi que 19 **plateformes d'expertise maladies rares** (PEMR)²² afin de renforcer l'articulation inter-filière au sein des établissements universitaires sièges de plusieurs centres labellisés. Une **banque nationale de données maladies rares** (BNMDR)²⁴ fut également créée afin de centraliser les données de santé des patients atteints d'une maladie rare.

Cependant, malgré l'avènement du séquençage nouvelle génération (NGS) et des examens pangénomiques (*i.e.* séquençage de l'exome et du génome), le diagnostic génétique des maladies rares demeure un véritable challenge. En effet, les patients atteints d'une maladie rare, ainsi que leur famille, sont le plus souvent confrontés à une période d'**errance diagnostique** avec un **délai moyen d'un an et demi**

¹⁹ Ministère de la Santé et de la Prévention. Les Maladies Rares. Mise à jour le 28/07/2023. <https://sante.gouv.fr/soins-et-maladies/prises-en-charge-specialisees/maladies-rares/article/les-maladies-rares>

²⁰ Maladies Rares Occitanie. Qu'est-ce qu'une maladie rare ? 2022. <https://www.maladies-rares-occitanie.fr/toutes-les-questions-que-vous-vous-posez-sur-la-maladie-rare/>

²¹ European Commission. Rare diseases. Updated on 2023.09.13. https://health.ec.europa.eu/non-communicable-diseases/expert-group-public-health/rare-diseases_en#:~:text=Up%20to%2036%20mill on%20people,70%25%20already%20start%20in%20childhood

²² Légifrance. Loi n°2004-804 du 9 août 2004 relative à la pol t que de santé publique (1). <https://www.legifrance.gouv.fr/loda/ d/JORFTEXT00000787078/>

²³ Ministère de la Santé et de la Prévention. Plan National Maladies Rares 2005-2008. https://sante.gouv.fr/IMG/pdf/plan_national_maladies_rares_2005-2008.pdf

²⁴ Ministère de la Santé et de la Prévention. Plan National Maladies Rares 2011-2014. https://sante.gouv.fr/IMG/pdf/plan_national_maladies_rares_2011-2014.pdf

²⁵ Ministère de la Santé et de la Prévention. Plan National Maladies Rares 2018-2022. https://sante.gouv.fr/IMG/pdf/pnrm_3_v25-09pdf.pdf

²⁶ Filières de Santé Maladies Rares. « Et si c'était une maladie rare ? ». <https://www.filièresmaladiesrares.fr/>

²⁷ Anne-Sophie Lapointe et Hélène Berruë-Gaillard. En route vers un 4^{ème} Plan National Maladies Rares. 9^{ème} congrès de la Société Française de Médecine Prédictive et Personnalisée (SFMP). 22/09/2023.

entre l'apparition des premiers symptômes et le diagnostic de la maladie. Ce délai s'étend même au-delà de **quatre ans** pour un quart des patients²⁰, dont 45% recevront au moins **diagnostic erroné** durant leur parcours de soins²⁸. Les premiers signes cliniques apparaissent le plus souvent à l'âge **pédiatrique**, avec 50% de patients âgés de moins de 2 ans et 75% ayant moins de 12 ans, nécessitant ainsi l'existence de stratégies diagnostiques précoces afin d'accompagner l'enfant au plus tôt dans son développement et prévenir toute perte de chance.

Aussi, au de-là de la recherche de microréarrangements chromosomiques par ACPA et des techniques de séquençage pangénomique (*i.e.* exome ou génome), nombreuses sont les innovations technologiques ayant récemment vu le jour dans le but d'améliorer et de faciliter le diagnostic génétique des maladies rares à l'ère de la Médecine Personnalisée de demain²⁹ (Figure B).

Le **séquençage long reads** (SLR) représente sans doute l'une des avancées technologiques les plus prometteuses pour le diagnostic des maladies génétiques. Cette nouvelle technologie repose sur le séquençage et l'alignement sur le génome de référence de fragments d'ADN de grande taille, le plus souvent situés entre 10 et 60 kilobases (kb), là où le NGS ne peut séquençer que des fragments de 100 à 300 paires de bases (pb)³⁰. Le SLR présente ainsi de nombreux avantages face au NGS tels qu'une meilleure détection des variations de structure équilibrées (SVs) et des microduplications segmentaires avec la caractérisation de leurs points de cassure respectifs, une meilleure analyse des régions riches en GC étant le plus souvent difficiles d'étude en NGS, l'étude par haplotype permettant le phasage des variants identifiés (*i.e.* variants en *cis* ou en *trans*), et l'identification des marques de méthylation de l'ADN, au premier rang desquels la 5-méthylcytosine^{31,32}. Le SLR est d'ores et déjà à l'étude dans plusieurs pays anglo-saxons avec comme objectif, à terme, le remplacement progressif du NGS en faveur de cette nouvelle technologie de séquençage de 3^{ème} génération³³.

²⁸ Molster, C., Urwin, D., Di Pietro, L. *et al.* Survey of healthcare experiences of Australian adults living with rare diseases. *Orphanet J Rare Dis* 11, 30 (2016).

²⁹ Marwaha S, Knowles JW, Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med.* 2022 Feb 28;14(1):23.

³⁰ Logsdon, G.A., Vollger, M.R. & Eichler, E.E. Long-read human genome sequencing and its applications. *Nat Rev Genet* 21, 597–614 (2020).

³¹ Sedlazeck FJ, Lee H, Darby CA, Schatz MC. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat Rev Genet.* 2018 Jun;19(6):329-346.

³² Mantere T, Kersten S, Hoischen A. Long-Read Sequencing Emerging in Medical Genetics. *Front Genet.* 2019 May 7;10:426. doi: 10.3389/fgene.2019.00426.

³³ Gorzynski JE, Goenka SD, Shafin K, Jensen TD, Fisk DG, Grove ME, Spiteri E, Pesout T, Monlong J, Ba d G, Bernstein JA, Ceresnak S, Chang PC, Christle JW, Chubb H, Dalton KP, Dunn K, Garalde DR, Guillory J, Knowles JW, Kolesnikov A, Ma M, Moscarello T, Nattestad M, Perez M, Ruzhnikov MRZ, Samadi M, Setia A, Wright C, Wusthoff CJ, X ong K, Zhu T, Jain M, Sedlazeck FJ, Carroll A, Paten B, Ashley EA. Ultrarapid Nanopore Genome Sequencing in a Critical Care Setting. *N Engl J Med.* 2022 Feb 17;386(7):700-702.

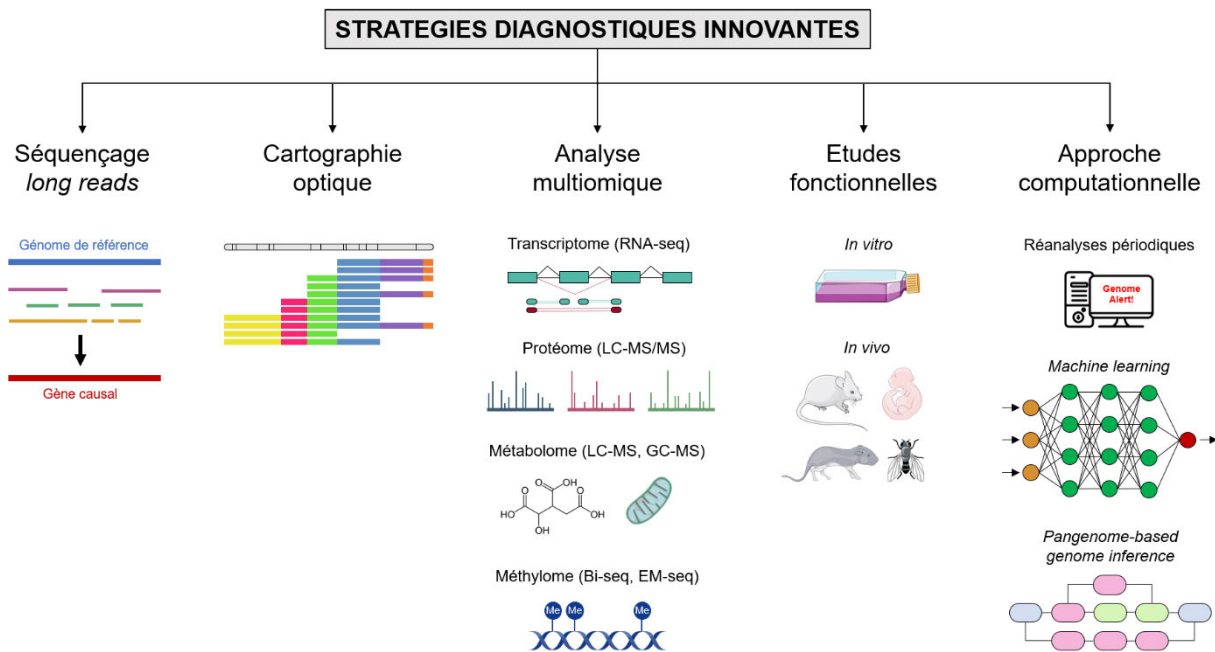


Figure B. Technologies innovantes pour le diagnostic génétique des maladies rares. Figure adaptée de [Marwaha *et al.*, *Genome Medicine*, 2022]. Abréviations : LC-MS/MS (chromatographie en phase liquide-spectrométrie de masse), GC-MS (chromatographie en phase gazeuse-spectrométrie de masse), Bi-seq (*bisulfite sequencing*), EM-seq (*enzymatic methyl sequencing*).

La **cartographie optique** (CO) représente quant à elle une avancée technologique majeure pour le monde de la génétique chromosomique. Cet examen, basé sur le traitement par enzymes de restriction de molécules d'ADN étirées et fixées sur un appareil microfluidique, permet l'obtention d'une carte de restriction pangénomique (*whole genome map*) de l'individu analysé. La CO permet l'analyse des variations de structures chromosomiques jusqu'à une résolution théorique de 500 pb, avec d'un côté une étude quantitative et la détection des variations du nombre de copies (CNVs) et de l'autre une analyse qualitative et l'identification des SVs³⁴. Cette nouvelle technologie « 2 en 1 », déjà utilisée en routine clinique pour l'étude des remaniements chromosomiques complexes et des phénomènes de *chromoanagenesis* (*i.e. chromothripsis, chromoanasyntesis* et *chromoplexy*)³⁵ s'apprête ainsi à occuper une place croissante dans les laboratoires de cytogénomique, aussi bien constitutionnels que somatiques.

³⁴ Bocklandt S, Hastie A, Cao H. Bionano Genome Mapping: High-Throughput, Ultra-Long Molecule Genome Analysis System for Precision Genome Assembly and Haploid-Resolved Structural Variation Discovery. *Adv Exp Med Biol*. 2019;1129:97-118.

³⁵ Pellestor F, Gaillard JB, Schneider A, Puechberty J, Gatinois V. Chromoanagenesis, the mechanisms of a genomic chaos. *Semin Cell Dev Biol*. 2022 Mar;123:90-99.

Le génome humain est composé d'environ **20.000 gènes** dont uniquement **25%** ayant une fonction connue à l'heure actuelle³⁶. L'amélioration du diagnostic génétique des maladies rares passera ainsi par une meilleure compréhension de ces 15.000 gènes dont le rôle demeure encore inconnu à ce jour, tout en optimisant l'étude de l'impact fonctionnel des variations de séquence identifiées. Pour cela, **l'analyse multiomique**, déjà utilisée dans le monde de la recherche, apparaît comme une technologie d'avenir pour le patient. Cette technique d'analyse intégrée, associant à l'analyse du génome celle du **transcriptome**, du **protéome** et du **métabolome** permet l'étude des conséquences fonctionnelles d'une variation d'intérêt tout en apportant des précieuses informations quant aux mécanismes moléculaires sous-jacents de la maladie génétique étudiée³⁷. L'analyse multiomique, en s'associant à des **études fonctionnelles complémentaires *in vitro* et *in vivo***, jouera ainsi un rôle prépondérant dans la Médecine de Précision de demain notamment pour la reclassification des variations de signification incertaine (VSI) ainsi que l'interprétation et la priorisation des variations non codantes situées au niveau des sites d'épissage et des régions régulatrices. A l'instar de l'article original de cette thèse, l'analyse du **méthylome** à la recherche d'une **signature épigénétique** spécifique d'une maladie génétique d'intérêt représente également une stratégie diagnostique innovante pour les patients porteurs d'une variation pathogène dans un gène impliqué dans la dynamique chromatinienne³⁸. A l'heure actuelle, une cinquantaine de maladies génétiques sont d'ores et déjà associées à une épisignature pouvant être utilisée en pratique courante pour la reclassification des VSI ou le diagnostic de la pathologie chez un patient dont les explorations génétiques se sont avérées non contributives³⁹. Aussi, l'implémentation du méthylome comme partie intégrante de l'analyse multiomique constituerait une source d'informations complémentaires pour l'étude des mécanismes impliqués dans les maladies rares encore non élucidés.

³⁶ Nurk S, Koren S, Rhie A, Rautiainen M, Bizikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, Aganezov S, Hoyt SJ, Diekhans M, Logsdon GA, Alonge M, Antonarakis SE, Borchers M, Bouffard GG, Brooks SY, Caldas GV, Chen NC, Cheng H, Chin CS, Chow W, de Lima LG, Dishuck PC, Durbin R, Dvorkina T, Fiddes IT, Formenti G, Fulton RS [...] Sims Y, Smith AFA, Soto DC, Sović I, Storer JM, Streets A, Sullivan BA, Thibaud-Nissen F, Torrance J, Wagner J, Walenz BP, Wenger A, Wood JMD, Xiao C, Yan SM, Young AC, Zarate S, Surti U, McCoy RC, Dennis MY, Alexandrov IA, Gerton JL, O'Neill RJ, Timp W, Zook JM, Schatz MC, Eichler EE, Miga KH, Phillippy AM. The complete sequence of a human genome. *Science*. 2022 Apr;376(6588):44-53.

³⁷ Dar MA, Arafah A, Bhat KA, Khan A, Khan MS, Ali A, Ahmad SM, Rashid SM, Rehman MU. Multiomics technologies: role in disease biomarker discoveries and therapeutic cs. *Brief Funct Genomics*. 2023 Apr 13;22(2):76-96.

³⁸ Aref-Eshghi E, Kerkhof J, Pedro VP, France GD, Barat-Houari M, Ruiz-Pallares N, Andrau JC, Lacombe D, Van-Gils J, Fergelot P, Dubourg C, Cormier-Daire V, Rondeau S, Lecoquierre F, Saugier-Verber P, Nicolas G, Lesca G, Chatron N, Sanlaville D [...] Balci TB, Siu VM, Stuart A, Kadour M, Masters J, Takano K, Kleefstra T, de Leeuw N, Field M, Shaw M, Gecz J, Ainsworth PJ, Lin H, Rodenhiser DI, Friez MJ, Tedder M, Lee JA, DuPont BR, Stevenson RE, Skinner SA, Schwartz CE, Genevieve D, Sadikov c B. Evaluation of DNA Methylation Episignatures for Diagnosis and Phenotype Correlations in 42 Mendelian Neurodevelopmental Disorders. *Am J Hum Genet*. 2021 Jun 3;108(6):1161-1163.

³⁹ Sadikov c B, Levy MA, Kerkhof J, Aref-Eshghi E, Schenkel L, Stuart A, McConkey H, Henneman P [...] Louie RJ, Lee JA, Tedder ML, Alders M. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet Med*. 2021 Jun;23(6):1065-1074.

L'**approche computationnelle *in silico*** constitue la dernière clef du diagnostic des maladies génétiques de demain. En effet, à l'heure de l'accélération des connaissances dans le domaine des maladies rares, des **relectures bioinformatiques périodiques** des données de séquençage issues d'exams pangénomiques initialement non contributifs apparaissent comme essentielles. Ainsi, différents algorithmes bioinformatiques standardisés furent développés pour la réinterprétation des données de séquençage pangénomiques à la lumière des nouvelles connaissances de la littérature. A titre d'exemple, le pipeline de réinterprétation *Genome Alert!* permis un sauvetage diagnostique dans 356 exomes initialement non contributifs sur les 4929 réanalysés, soit plus de 7% de nouveaux diagnostics⁴⁰. A l'ère du *Big Data* et de la multiplication des bases données, le recours à l'**intelligence artificielle** (IA) et au **machine learning** s'avèrera également incontournable pour l'extraction, la synthèse et l'interprétation des données utiles au diagnostic génétique des maladies rares⁴¹. Enfin, l'optimisation des pipelines bioinformatiques dédiés à l'alignement des séquences sur le génome de référence constituera un véritable défi. Différentes technologies novatrices sont actuellement développées dans ce domaine au premier rang desquelles le **pangenome-based genome inference**⁴². Contrairement aux pipelines conventionnels, cet algorithme bioinformatique repose sur l'identification des variants de séquence (*variant calling*) en amont de leur alignement sur le génome de référence⁴². Cette technologie, basée sur le principe d'inférence, permet ainsi un meilleur alignement des séquences nucléotides dans les régions polymorphiques et riches en GC, réduisant alors drastiquement le risque de biais d'alignement⁴².

En conclusion, le diagnostic génétique des maladies rares représente un véritable challenge pour la communauté médicale et scientifique. A l'avenir, l'optimisation de celui-ci passera par l'association de ces différentes techniques complémentaires décrites comme innovantes aujourd'hui, mais dont l'utilisation généralisée dans le monde de demain les feront alors tomber en désuétude au profit de nouvelles technologies novatrices dépassant les limites de notre imagination actuelle.

⁴⁰ Yauy K, Lecoquierre F, Baert-Desurmont S, Trost D, Boughalem A, Luscan A, Costa JM, Geromel V, Raymond L, Richard P, Coutant S, Broutin M, Lanos R, Fort Q, Cackowski S, Testard Q, Diallo A, Soirat N, Holder JM, Duforet-Frebourg N, Bouge AL, Beaumeunier S, Bertrand D, Audoux J, Genevieve D, Mesnard L, Nicolas G, Thevenon J, Philippe N. Genome Alert!: A standardized procedure for genomic variant reinterpretation and automated gene-phenotype reassessment in clinical routine. *Genet Med.* 2022 Jun;24(6):1316-1327.

⁴¹ Banerjee J, Taroni JN, Allaway RJ, Prasad DV, Guinney J, Greene C. Machine learning in rare disease. *Nat Methods.* 2023 Jun;20(6):803-814.

⁴² Ebler J, Ebert P, Clarke WE, Rausch T, Audano PA, Houwaart T, Mao Y, Korbel JO, Eichler EE, Zody MC, Dilthey AT, Marschall T. Pangenome-based genome inference allows efficient and accurate genotyping across a wide spectrum of variant classes. *Nat Genet.* 2022 Apr;54(4):518-525.

ANNEXES



EURODYSMORPHO

33rd EDITION OF THE EUROPEAN WORKSHOP ON
DYSMORPHOLOGY
PORTUGAL, LISBON, 13-16 SEPTEMBER 2023

Saturday 16 September

Session 11: Syndrome delineation - Chair: **Sally Ann Lynch**

- | | | |
|-------|----------------------------------|--|
| 9:00 | Elis Tiivoja* | Expanding the clinical phenotype of individuals with novel variants in the X-linked WDR13 gene |
| 9:15 | Giulia Vitetta* | Prenatal and postnatal phenotype of ZTTK syndrome: report of two cases |
| 9:30 | Maria Chiara Baroni* | First patient with compound heterozygous variants in ZNF528: the combination of a frameshift and a missense variant could lead to a milder presentation of the disease |
| 9:45 | Anita Rauch | Further evidence for a silver-russell-like phenotype caused by truncating variants in PLAG1 |
| 10:00 | Quentin Sabbagh | Genotype-phenotype expansion of BCL11B-related disorder and identification of an epismutation: a series of 25 unreported individuals |
| 10:15 | Alessandro Mauro Spinelli | Independent confirmation of a clinically recognizable multisystem syndrome due to a recurrent de novo NSD2 gain-of-function variant |
| 10:30 | Coffee break | |

Genetics
in
Medicine
The Official Journal of the ACMG

Oct 01, 2023

RE: GENETMED-D-23-00500R2, entitled "Clinico-biological refinement of BCL11B-related disorder and identification of an epismutation: a series of 20 unreported individuals"

Dear Quentin SABBAGH,

I am pleased to inform you that your work has now been accepted for publication in Genetics in Medicine and will be processed to production for publication pending the review and receipt of the final editorial office items listed below.

Receipt of the editorial office items below is due on Oct 08, 2023.

We have also included reviewer comments on this version of the manuscript purely for your information, no response to these comments is required nor expected.

Reviewer comments:

Reviewer #3: All of the questions I had have been addressed. No further comments. Thank you!

Thank you for submitting your interesting and important work to the journal.

With Kind Regards,

Dr. Robert Steiner
Editor-in-Chief
Genetics in Medicine

SERMENT

En présence des Maîtres de cette école, de mes chers condisciples et devant l'effigie d'Hippocrate, je promets et je jure, au nom de l'Être suprême, d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la médecine.

Je donnerai mes soins gratuits à l'indigent et n'exigerai jamais un salaire au-dessus de mon travail.

Admis dans l'intérieur des maisons, mes yeux ne verront pas ce qui s'y passe, ma langue taira les secrets qui me seront confiés, et mon état ne servira pas à corrompre les mœurs, ni à favoriser le crime.

Respectueux et reconnaissant envers mes Maîtres, je rendrai à leurs enfants l'instruction que j'ai reçue de leurs pères.

*Que les hommes m'accordent leur estime si je suis fidèle à mes promesses.
Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.*

DESCRIPTION CLINICO-BIOLOGIQUE DE LA MALADIE LIÉE AU GÈNE *BCL11B* ET IDENTIFICATION D'UNE SIGNATURE ÉPIGÉNÉTIQUE

RÉSUMÉ

Rationnel : La maladie liée au gène *BCL11B* (MIM 618092) est caractérisée par un trouble du neurodéveloppement syndromique associant un décalage des acquisitions, des variations morphologiques mineures et un déficit immunitaire. A ce jour, 29 patients ont été rapportés dans la littérature. Cette étude repose sur une description clinico-biologique de 20 nouveaux individus porteurs d'une variation pathogène du gène *BCL11B*, ainsi que sur l'analyse pangénomique de leur profil de méthylation.

Méthodes : Les données cliniques, immunitaires et moléculaires des 20 patients porteurs d'une variation pathogène du gène *BCL11B* ont été récoltées à partir d'une collaboration internationale, *via* un questionnaire standardisé, et comparées aux données de la littérature. L'analyse du méthylome, menée en collaboration avec le *London Health Sciences Centre* (Ontario, Canada), a été réalisée à partir d'ADN extrait sur prélèvement sanguin en utilisant la technologie du séquençage au bisulfite (*Bi-Seq*).

Résultats : Nous décrivons des manifestations cliniques peu documentées dans la littérature telles qu'une morphologie crânio-faciale évocatrice du syndrome de Rubinstein-Taybi, ainsi qu'un surrisque de craniosténose, de maladies auto-immunes et de comorbidités psychiatriques (*i.e.* trouble anxieux et comportement auto/hétéro-agressif). De surcroît, nous identifions une signature épigénétique spécifique de la maladie liée au gène *BCL11B*.

Conclusion : Cette étude enrichit nos connaissances du spectre clinico-biologique de la maladie liée au gène *BCL11B*, optimisant ainsi le diagnostic et la prise en charge des patients atteints de cette maladie rare. De surcroît, l'identification d'une signature épigénétique sensible et spécifique représente une véritable avancée pour le diagnostic de la maladie liée au gène *BCL11B*. Cette nouvelle analyse pourra désormais être déployée en pratique quotidienne afin, notamment, d'aider à la reclassification des variations de signification incertaine du gène *BCL11B*.

MOTS CLEFS

BCL11B - Trouble du neurodéveloppement - Méthylation - Épisignature