Clinical and molecular insights into BAFopathies

Eline van der Sluijs

Clinical and molecular insights into BAFopathies

P.J. van der Sluijs

ISBN: 978-94-6522-516-6

Printing: Provided by thesis specialist Ridderprint, ridderprint.nl
Layout and design: Tara Schollema, persoonlijkproefschrift.nl
Cover design and artwork: Resi M. Baan-Schoonderwoerd

Clinical and molecular insights into BAFopathies

The printing of this thesis was supported by the Department of Clinical Genetics at LUMC, ChipSoft, and the Scholten-Cordes Fonds.

Clinical and molecular insights into BAFopathies

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op donderdag 30 oktober 2025

klokke 10:00 uur

door

Pleuntje Jantine van der Sluijs

geboren te Gorinchem

in 1995

Promotores

Prof. Dr. C.J. van Asperen

Prof. Dr. G.W.E. Santen

Co-promotores

Dr. L. Clemens-Daxinger

Leden promotiecommissie

Dr. S.A.M.J. Lesnik-Oberstein

Prof. Dr. M.M. van Haelst[1]

Prof. Dr. T. Kleefstra[2]

Dr. B.B.A. de Vries[3]

Affiliatie

1. Amsterdam UMC, Amsterdam
2. Erasmus MC, Rotterdam
3. Radboud UMC, Nijmegen

TABLE OF CONTENTS

Chapter 1 General introduction 7

Part I – Insights into clinical and epigenetic phenotypes

Chapter 2 The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin-Siris syndrome 25
Genet Med. 2019 Jun;21(6):1295-1307. doi: 10.1038/s41436-018-0330-z.

Chapter 3 Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort 49
Genet Med. 2022 Aug;24(8):1753-1760. doi: 10.1016/j.gim.2022.04.010.

Chapter 4 A Case Series of Familial ARID1B Variants Illustrating Variable Expression and Suggestions to Update the ACMG Criteria 65
Genes (Basel). 2021 Aug 20;12(8):1275. doi: 10.3390/genes12081275.

Chapter 5 ARID1B-related disorder in 87 adults: Natural history and self-sustainability. 87
Genet Med Open. 2024 Jul 23:2:101873. doi: 10.1016/j.gimo.2024.101873.

Chapter 6 Microduplications of ARID1A and ARID1B cause a novel clinical and epigenetic distinct BAFopathy 121
Genet Med. 2025 Jan;27(1):101283. doi: 10.1016/j.gim.2024.101283.

Part II – Methods for studying mechanisms of disease and possible treatment

Chapter 7 Clonazepam repurposing in ARID1B patients through conventional RCT and N-of-1 trials: an experimental strategy for orphan disease development 149
J Med Genet . 2024 Dec 31:jmg-2024-109951. doi: 10.1136/jmg-2024-109951.

Chapter 8 Developing an in vitro model system to study the effects of Coffin-Siris Syndrome pathogenic variants on neuronal differentiation 175
Manuscript in preparation.

Chapter 9 Discussion 201

Addendum Summary 218
Samenvatting 221
List of Publications 224
Curriculum Vitae 229
Dankwoord 230

Chapter 1

Introduction

INTRODUCTION

BAFopathies are disorders caused by pathogenic germline variants in genes encoding part of the BAF (BRG_1/BRM associated factor) chromatin remodeling complex (Figure 1). Altogether, BAFopathies are one of the main genetic causes of intellectual disability (ID)[1,2]. Based on the phenotype and molecular cause, BAFopathies can be divided into different entities, of which Coffin-Siris syndrome (CSS, OMIM#135900) is the most frequent.

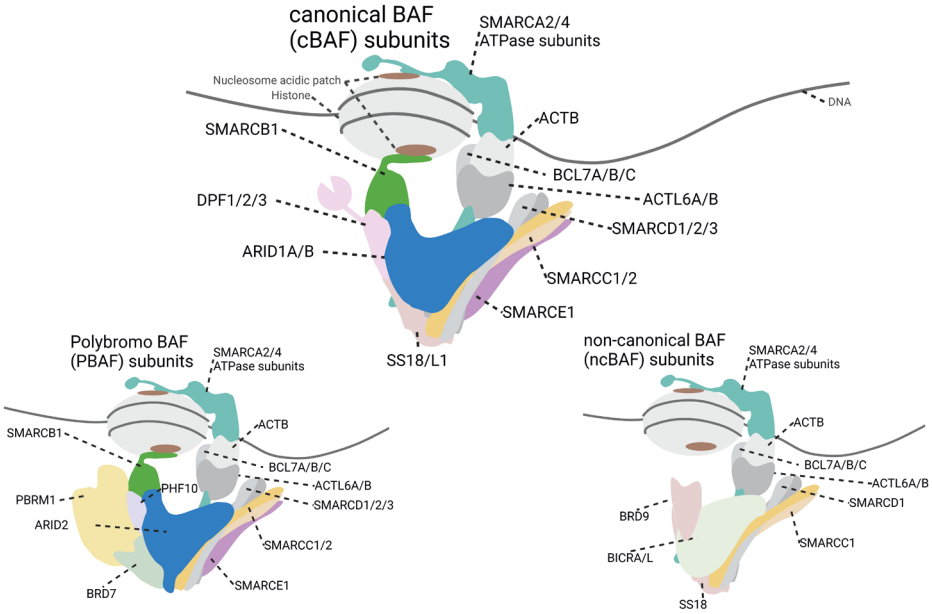


Figure 1: BAF-complex.
BAF-complex structure is derived from Mashtalir et al. 2021 [53], Wei et al 2023 [54], and Valencia et al 2023[1].

Coffin-Siris syndrome

CSS was first described in 1970 by dr. Coffin, a pediatrician, and dr. Siris, a radiologist[3]. Originally, CSS (also known as '5th digit syndrome') was diagnosed based on clinical criteria, including a short 5th finger, 5th finger or toenail hypoplasia, coarse facial features, hypertrichosis and intellectual disability or developmental delay frequently coinciding with a range of other features[4,5].

The causative genes of CSS remained unknown until next generation sequencing (NGS) in 2012 allowed the identification of the first genes causing CSS (i.e. *ARID1B*, *ARID1A*,

SMARCA4, *SMARCB1* and *SMARCE1*)[6,7] in patients with a clinical diagnosis of CSS (Figure 2). The publication describing *ARID1B* as a cause of CSS also identified a number of patients with a deletion of *ARID1B*, some without the typical features of CSS, indicating that perhaps *ARID1B* could be an important cause of ID[6]. This was confirmed around the same time by another publication identifying pathogenic variants in *ARID1B* as a frequent cause of ID[8]. Later, large-scale NGS studies in less selected patient groups (i.e. ID, neurodevelopmental delay (NDD), autism) confirmed that *ARID1B* is one of the most frequent causes of ID at about 1%[2,9-13]. Given the often-subtle features of CSS it is perhaps not surprising that the diagnosis of CSS has since then transitioned from a clinical diagnosis to one more and more based on genotype.

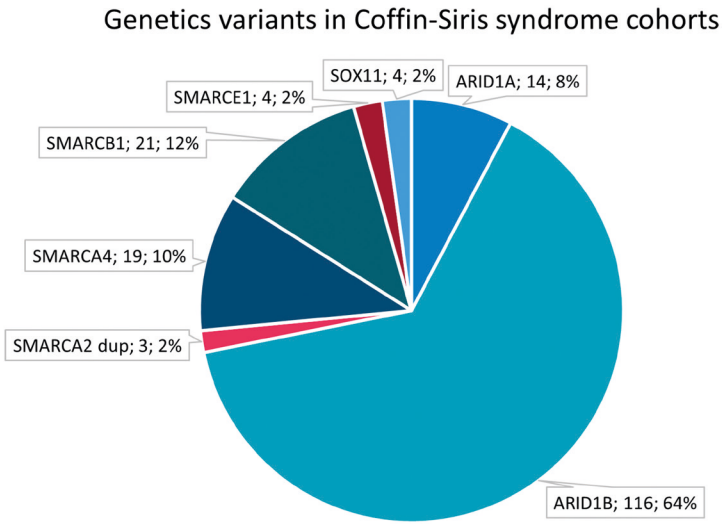


Figure 2: CSS genes pie chart based on reported CSS cohorts [7,55-58]*
**PHF6* and *SMARCA2* variants were also identified in a number of clinical CSS patients, in retrospect these patients were diagnosed as having Börjeson-Forssman-Lehmann syndrome (*PHF6*) and Nicolaides-Baraitser syndrome (*SMARCA2*).

Other BAFopathies

Aside from CSS, several other syndromic causes of ID have been linked to the BAF-complex. Nicolaides Baraitser syndrome (NCBRS, OMIM 601358) has been associated with pathogenic missense variants in *SMARCA2*[14], while whole gene duplications of *SMARCA2* have been linked to a CSS-like phenotype[15]. Other BAF-complex genes linked to a CSS-like phenotype are *ARID2*[16], *BICRA*[20], *DPF2*[17], *SMARCC2*[18], *SMARCD1*[27] and, *SOX11*[19]. Furthermore, pathogenic variants in *ACTL6A*[21], *ACTL6B*[22,23] and, *SMARCC1*[24-26] have each been associated with a gene-specific (syndromic) ID or NDD phenotype. As illustrated in Figure 1 all these genes encode parts of the BAF-complex except *SOX11*.

SOX11 is a downstream transcriptional factor of the BAF complex[19,28]. Interestingly, specific pathogenic variants in a number of the CSS and NCBRS genes lead to other non-CSS phenotypes. A phenotype of severe ID and choroid plexus hyperplasia has been observed in patients with a specific *SMARCB1* variant p.(Arg37His)[29], *SMARCA2* variants outside the helicase domains of *SMARCA2* have been observed in patients with blepharophimosis intellectual disability syndrome[30] (BIS, OMIM 619293) and, whole gene duplications of *ARID1A* cause a distinct syndrome ID[31] (Figure 3). All of this illustrates the broad phenotypic spectrum associated with these BAFopathies.

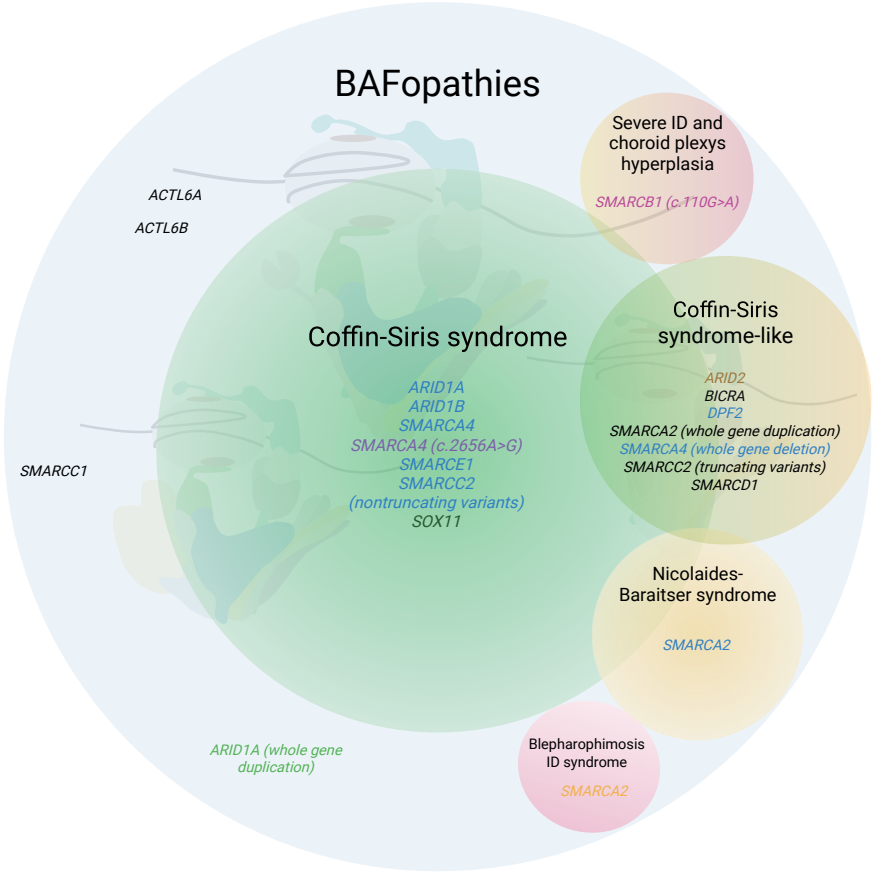


Figure 3: Genetic and epigenetic landscape of BAFopathies.
BAFopathies, caused by pathogenic variants in genes encoding components of the BAF complex, can be categorized into distinct entities based on genotype and phenotype. Disorders with specific names are written in regular font, while gene names are italicized. Additionally, colors indicate the episignature associated with pathogenic variants in these genes, where applicable. The size of the circles approach the number of patient in these groups, but are not an exact estimate.

BAF-complex

The genes mutated within BAFopathies all encode subunits of the BAF-complex or are a transcription factor downstream of BAF-complex (e.g. *SOX11*). As illustrated by Figure 1, there are three different configurations of the BAF-complex called canonical BAF-complex, noncanonical BAF-complex and polybromo BAF-complex. These distinct subunit assemblies of the BAF-complex occur at different time points during development and in different tissues, indicating specific roles of each subunit.

Current issues in BAFopathies

Over the past decade the application of NGS has brought us the identification of BAFopathies and enabled the discovery, diagnosis and description of several gene-specific phenotypes. However, this development has also revealed several challenges concerning 1) the interpretation of genetic variants, 2) shortage of genotype-phenotype studies and, 3) the lack of available treatments for the increasing number of diagnosed patients with BAFopathies.

1) Variant interpretation

The increasing application of NGS in genetic diagnostics has led to the identification of many variants in BAFopathy genes. Some of these variants are clearly pathogenic, but an increasing number is considered a variant of uncertain significance (VUS). Interpretation of these VUS is particularly challenging because of the broad, and often not very specific, phenotype associated with BAFopathies.

To aid and standardize the interpretation of genetic variants the American College of Medical Genetics and Genomics (ACMG) formulated a framework[32]. The ACMG framework can be used to interpret the variant into five classes: benign, likely benign, VUS, likely pathogenic, and pathogenic. Important aspects in the ACMG framework in the context of BAFopathies are whether the variant occurs in control populations, whether the variant segregates with the disease or is *de novo*; and whether the phenotype fits the known phenotype associated with the gene.

Whilst the ACMG criteria generally serve their purpose, they are formulated in a very general manner to apply to all genes and genetic variants. However, there are instances where these criteria may prove insufficient, necessitating (gene-specific) amendments[33-37]. For example, based on the ACMG framework a missense variant in a known ID-gene could be regarded likely pathogenic when it occurs *de novo* and absent in control populations. The UK Association for Clinical Genomic Science (UK-ACGS) published a specification for variant interpretation in which they specified that in the context of a high penetrant monogenic disorder a *de novo* ACMG-criterium (PS2)

should only be applied if the patient's phenotype is consistent with the gene-phenotype association.

Phenotype assessment

Classically, in the context of ID syndromes the phenotype has always been a key determinant in classifying a VUS. In this assessment it may be very helpful to consider whether the clinician suspected this gene or condition before genetic testing. This is to avoid a Texas sharpshooter fallacy, where differences are ignored and similarities overemphasized, like a sharpshooter drawing targets around random bullet holes. Just because certain clinical characteristics match those commonly associated with a gene, this does not necessarily make the characteristics specific for the disorder linked to that gene. Similar features could manifest in patients with other conditions. The assessment of whether a phenotype is fitting for a gene depends on human interpretation and is therefore subjective. A more objective approach to assessing the phenotype is through Human Phenotype Ontology (HPO) terms[38].

HPO-terms and PhenoScore

HPO is a standardized vocabulary of phenotypic abnormalities encountered in human disease. Each term represents a phenotypic abnormality. All terms are connected via a tree-like structure in a class-subclass relationship. For example, the HPO-term 'intellectual disability (HP:0001249)' is a subclass of 'Neurodevelopmental abnormality (HP:0012759)'. HPO-terms can be used in interpretation pipelines for genetic testing. Genes associated with the phenotype of the patient can be prioritized using HPO-terms.

HPO-terms can also be used in deep phenotyping. Deep phenotyping is the precise and comprehensive analysis of phenotypic abnormalities. An example of deep phenotyping via HPO-terms is PhenoScore[39]. PhenoScore is an artificial intelligence-based phenomics framework, which integrates facial recognition technology and HPO data analysis to quantify phenotypic similarity. PhenoScore can assess, based on HPO-terms and facial photographs, whether a patient group can be distinguished from matched controls. Consequently, PhenoScore can be used in VUS interpretation. PhenoScore can calculate a numerical score representing the similarity between a patient's phenotype and the expected phenotype of the disorder. In this manner PhenoScore provides an objective measure of phenotypic similarity, can identify group specific features, and aid in VUS interpretation. For example, the PhenoScore clustering result of a *de novo* VUS could be used to determine whether a *de novo* criterium (PS2) and specific phenotype criterium (PP4) may be applied.

If a VUS has been identified, assessing whether a patient's phenotype fits the expected phenotype of the VUS-gene is especially complicated in the context of BAFopathies. This is because there have already been several instances described where specific variants in a single gene lead to different syndromes. Examples are that (1) deletion and truncating variants in *ARID1A* lead to Coffin-Siris syndrome[7], but whole gene duplications lead to a different ID-phenotype[31], or, (2) missense variants in *SMARCB1* C-terminal domain lead to CSS[7], while one specific missense variant in the winged helix domain (near the N-terminal) leads to a specific non-CSS phenotype[29]. Deep phenotyping, extensive clinical descriptions and PhenoScore analyses could assist interpretation in BAFopathies.

Other tools of variant classification

There are several tools that could provide insight into the potential pathogenicity of a VUS. For example, functional readouts—such as introducing the variant into cellular or animal models—can help clarify the biological impact of the variant. These approaches are usually time-consuming, expensive and require the presence of an existing disease model.

DNA methylation

Another approach to aid variant interpretation, requiring only the leftover DNA from genetic testing is the analysis of DNA methylation patterns. DNA methylation is essential for gene regulation. DNA methylation patterns differ between tissue types[40], and can also differ between patients and controls. Specific DNA methylation patterns in blood have been linked to several disorders, such as cancer, neurodevelopmental disorders[41], and genetic syndromes[42,43], such as BAFopathies[44] (Figure 3). These methylation patterns can be used to define patient groups and can afterwards be used for VUS interpretation. For example, pathogenic variants in CSS and NCBRS genes lead to a BAFopathy episignature in blood. Several patients with a missense variant in *SMARCA2* and an ID phenotype, but not specific for NCBRS did not have this BAFopathy episignature in their blood, but were later found to have a different BIS-specific episignature.

2) Genotype-phenotype studies and their potential biases

A second issue for BAFopathies is the lack of genotype-phenotype studies. A genetic diagnosis, based on a (likely) pathogenic variant in a BAFopathy gene enables the clinician to counsel parents about the (expected) phenotype and give screening recommendations. This information is based on genotype-phenotype studies of the diagnosed BAFopathy. Genotype-phenotype studies asses clinical phenotypes of patients and divide this clinical description in gene and, if applicable, genotype

specific groups. Since genotype-phenotype studies serve as the information source for counseling patients, parents, and caregivers, it is essential to acknowledge potential pitfalls and biases inherent in such studies. For example, in the case of CSS, it is important that phenotypes are assessed per gene because there are important differences between patients with pathogenic variants in different CSS genes, and even among patients harboring pathogenic variants in the same gene. Aside from this, there are several potential biases inherent in genotype-phenotype studies.

Ascertainment bias is such an example. Patients diagnosed after a suspicion of a specific syndrome may have a different phenotypic spectrum than a patient diagnosed after whole exome or genome sequencing. Patients with a pre-test suspicion may be more likely to have the syndrome specific features compared to patients identified to genome-wide sequencing ID without any suggestive features. Research examining potential differences among these patient groups is therefore essential.

Another type of bias is phenotype reporting bias. In larger cohort studies, the presence or absence of certain features may be unknown in a subset of patients. Excluding these patients from prevalence estimation can lead to overestimation, while counting them as lacking the feature may result in underestimation. To address this type of bias, it is advisable to report a prevalence range for such features or use data collection techniques that allow annotation of the unknown status, providing a more accurate representation.

Yet another type of bias is associated with the age of patients at inclusion. Due to the advancements in NGS children are predominantly diagnosed with rare disorders, leading to an underrepresentation of adult-aged patients in reported cohorts. This discrepancy can affect the evaluation of age-related skills or features, such as seizures or other potential adult-onset manifestations. It is therefore valuable to analyze developmental features, like the age of first words, in survival analyses and separately examine adult-aged patients to offer a more comprehensive overview of features associated with older age. This ensures a more nuanced understanding of age-related characteristics of the studied disorder over time.

3) Etiology and treatment of BAFopathies

To enhance counseling, interpretation of a genetic variant, obtaining a genetic diagnosis and genotype-phenotype studies are paramount, but ongoing research into the pathophysiology of BAFopathies is also necessary for the improvement of counseling and care. Therefore, the third challenge within BAFopathies is the lack of potential treatment options. To propose viable treatment options understanding of the etiology

is needed. Exploring the underlying pathophysiology of BAFopathies through disease models, such as animal (*in vivo*) models and patient-derived cell (*in vitro*) models or organoids is valuable because these models do not only deepen our understanding of the disorder but also pave the way for potential disease-specific therapeutic interventions.

BAFopathy-based animal models, such a Arid1b-haploinsufficient mice[45], can be used to study disease mechanisms. An advantage of an animal model is that disease progression can be studied in a living organism, an organism with a shorter lifespan and (brain) tissue samples can be taken. And, after a clear phenotype is established, therapeutic interventions can be tested. An example is described by Jung et al[46]. In their Arid1b-haploinsufficient mouse model they identified an increased apoptosis and decreased proliferation of inhibitory GABAergic interneurons. Consequently, they hypothesized and demonstrated that clonazepam through enhancing the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) may rescue the mice's phenotype. A disadvantage of animal models is the question whether findings from animal models can be extrapolated to humans. A possible next step could be validating the finding in human cell models[47-50] or organoids[51,52].

Patient-derived cell models and organoids have the advantage of investigating the effect of a variant in the genetic background of the individual. Although in extrapolating the findings to other patients this could be a disadvantage and iso-genic cell lines (e.g. CRISPR-Cas) may be preferred. Cell models can be used to examine (aberrant) DNA methylation, gene expression pattern via RNA-seq and investigate dysregulated pathways.

These models can be used not only to investigate the pathophysiology of genetic variants but also to explore potential treatment effects. Proposing and testing a potential disease-specific therapy within a disease model is an essential step in therapy development, before efficacy can be investigated in patients. In rare disorders like ID, limited patient numbers and heterogeneity within the disorder pose challenges to designing effective interventions. Publication of research into potential therapeutic approaches for rare ID disorders not only benefits specific patient groups but also serves as a valuable example for future therapeutic studies involving ID patients.

Aim and scope of this thesis

All of the above illustrates the intricacies of variant interpretation, the importance of genotype-phenotype studies, and the need for etiological and treatment studies in BAFopathies. To address these issues the research in this thesis aims 1) to provide

further insight in the clinical and epigenetic phenotypes associated with pathogenic variants in BAFopathy genes; and; 2) further examine the methods for studying mechanisms of disease and possible treatment by creating a model to further study possible underlying etiological aspects and by studying the potential effect of clonazepam treatment in *ARID1B* patients.

Outline of this thesis

1) Insights into clinical and epigenetic phenotypes

In **chapter 2** we describe 143 patients with pathogenic variants in *ARID1B* to answer the question whether there is a difference between patient diagnosed with pathogenic variant after a suspicion of CSS and patients diagnose after whole exome or genome sequencing. In this chapter we broaden the phenotypic spectrum associated with pathogenic *ARID1B* variants, we investigate the effects of ascertainment bias and phenotype reporting bias in our cohort and examine different data collection methods. In **chapter 3** we investigate the fetal CSS phenotype by describing prenatal anomalies detected among patients with pathogenic variants in CSS-associated genes, highlighting a new part of the phenotype associated with CSS, and providing an explanation why variants in *ARID1A* are so much rarer than variants in *ARID1B*.

A subset of CSS patients have a distinct DNA methylation pattern present in their blood, also referred to as the BAFopathy episignature[44]. This episignature provides a valuable tool in diagnostic care in VUS interpretation. In **chapter 4** we use this episignature, together with the predecessor of the PhenoScore algorithm, to interpret inherited *ARID1B* variants and evaluate the ACMG criteria in the context of *ARID1B*. Our analysis sheds light on the challenges associated with interpreting inherited *ARID1B* variants and provides recommendations concerning the ACMG-criteria to aid the interpretation of such variants. In **chapter 5** we address age-dependent bias by describing 87 adult-aged patients with pathogenic variants in *ARID1B* to give insight in the development of these patients and any features that develop with age. Based on this information we formulate screening recommendations.

In **chapter 6**, where we report the *ARID1B* microduplication syndrome and compare the DNA methylation pattern and phenotype of patients with a whole gene duplication with data of patients with an *ARID1A* duplication.

2) Methods for studying mechanisms of disease and possible treatment

In **chapter 7** we describe an *in vitro* model system to study CSS variant effects in neuronal differentiation of induced Pluripotent Stem Cells (iPSC) of patients with pathogenic variants in *ARID1B* and *SMARBC1*. These iPSCs are differentiated towards

GABAergic interneurons to investigate DNA methylation throughout neuronal development and whether the increased apoptosis observed in Arid1b haploinsufficient mice[46] also occurs in human cells.

In this same mouse model[46], these researchers hypothesized and demonstrated that due to the reduced number of inhibitory GABAergic interneurons, administering clonazepam, which enhances the activity of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA), partially rescues the phenotype of Arid1b haploinsufficient mice. In **chapter 8** we investigate via a randomized, double-blind, crossover study followed by an n-of-1 design whether clonazepam administration has a similar effect in *ARID1B* patients.

REFERENCES

1. Valencia AM, Sankar A, van der Sluijs PJ, et al. Landscape of mSWI/SNF chromatin remodeling complex perturbations in neurodevelopmental disorders. *Nat Genet*. Aug 2023;55(8):1400-1412. doi:10.1038/s41588-023-01451-6
2. Gillentine MA, Wang T, Eichler EE. Estimating the Prevalence of De Novo Monogenic Neurodevelopmental Disorders from Large Cohort Studies. *Biomedicines*. Nov 9 2022;10(11) doi:10.3390/biomedicines10112865
3. Coffin GS, Siris E. Mental retardation with absent fifth fingernail and terminal phalanx. *Am J Dis Child*. May 1970;119(5):433-9.
4. Fleck BJ, Pandya A, Vanner L, Kerkering K, Bodurtha J. Coffin-Siris syndrome: review and presentation of new cases from a questionnaire study. *Am J Med Genet*. Feb 15 2001;99(1):1-7. doi:10.1002/1096-8628(20010215)99:1<1::aid-ajmg1127>3.0.co;2-a
5. Schrier SA, Bodurtha JN, Burton B, et al. The Coffin-Siris syndrome: a proposed diagnostic approach and assessment of 15 overlapping cases. *Am J Med Genet A*. Aug 2012;158A(8):1865-76. doi:10.1002/ajmg.a.35415
6. Santen GW, Aten E, Sun Y, et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet*. Mar 18 2012;44(4):379-80. doi:10.1038/ng.2217
7. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet*. Mar 18 2012;44(4):376-8. doi:10.1038/ng.2219
8. Hoyer J, Ekici AB, Endele S, et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am J Hum Genet*. Mar 09 2012;90(3):565-72. doi:10.1016/j.ajhg.2012.02.007
9. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet*. Apr 04 2015;385(9975):1305-14. doi:10.1016/S0140-6736(14)61705-0
10. Deciphering Developmental Disorders S. Large-scale discovery of novel genetic causes of developmental disorders. *Nature*. Mar 12 2015;519(7542):223-8. doi:10.1038/nature14135
11. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. Feb 23 2017;542(7642):433-438. doi:10.1038/nature21062
12. Hamdan FF, Srour M, Capo-Chichi JM, et al. De novo mutations in moderate or severe intellectual disability. *PLoS Genet*. Oct 2014;10(10):e1004772. doi:10.1371/journal.pgen.1004772
13. Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*. Jan 23 2020;doi:10.1016/j.cell.2019.12.036
14. Van Houdt JK, Nowakowska BA, Sousa SB, et al. Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. *Nat Genet*. Feb 26 2012;44(4):445-9, S1. doi:10.1038/ng.1105
15. Miyake N, Abdel-Salam G, Yamagata T, et al. Clinical features of SMARCA2 duplication overlap with Coffin-Siris syndrome. *Am J Med Genet A*. Oct 2016;170(10):2662-70. doi:10.1002/ajmg.a.37778
16. Shang L, Cho MT, Retterer K, et al. Mutations in ARID2 are associated with intellectual disabilities. *Neurogenetics*. Oct 2015;16(4):307-14. doi:10.1007/s10048-015-0454-0
17. Vasileiou G, Vergarajauregui S, Endele S, et al. Mutations in the BAF-Complex Subunit DPF2 Are Associated with Coffin-Siris Syndrome. *Am J Hum Genet*. Mar 1 2018;102(3):468-479. doi:10.1016/j.ajhg.2018.01.014

18. Machol K, Rousseau J, Ehresmann S, et al. Expanding the Spectrum of BAF-Related Disorders: De Novo Variants in SMARCC2 Cause a Syndrome with Intellectual Disability and Developmental Delay. *Am J Hum Genet*. Jan 3 2019;104(1):164-178. doi:10.1016/j.ajhg.2018.11.007
19. Tsurusaki Y, Koshimizu E, Ohashi H, et al. De novo SOX11 mutations cause Coffin-Siris syndrome. *Nat Commun*. Jun 2 2014;5:4011. doi:10.1038/ncomms5011
20. Barish S, Barakat TS, Michel BC, et al. BICRA, a SWI/SNF Complex Member, Is Associated with BAF-Disorder Related Phenotypes in Humans and Model Organisms. *Am J Hum Genet*. Dec 3 2020;107(6):1096-1112. doi:10.1016/j.ajhg.2020.11.003
21. Marom R, Jain M, Burrage LC, et al. Heterozygous variants in ACTL6A, encoding a component of the BAF complex, are associated with intellectual disability. *Hum Mutat*. Oct 2017;38(10):1365-1371. doi:10.1002/humu.23282
22. Fichera M, Failla P, Saccuzzo L, et al. Mutations in ACTL6B, coding for a subunit of the neuron-specific chromatin remodeling complex nBAF, cause early onset severe developmental and epileptic encephalopathy with brain hypomyelination and cerebellar atrophy. *Hum Genet*. Feb 2019;138(2):187-198. doi:10.1007/s00439-019-01972-3
23. Bell S, Rousseau J, Peng H, et al. Mutations in ACTL6B Cause Neurodevelopmental Deficits and Epilepsy and Lead to Loss of Dendrites in Human Neurons. *Am J Hum Genet*. May 2 2019;104(5):815-834. doi:10.1016/j.ajhg.2019.03.022
24. Al Mutairi F, Alzahrani F, Ababneh F, Kashgari AA, Alkuraya FS. A mendelian form of neural tube defect caused by a de novo null variant in SMARCC1 in an identical twin. *Ann Neurol*. Feb 2018;83(2):433-436. doi:10.1002/ana.25152
25. Furey CG, Choi J, Jin SC, et al. De Novo Mutation in Genes Regulating Neural Stem Cell Fate in Human Congenital Hydrocephalus. *Neuron*. Jul 25 2018;99(2):302-314 e4. doi:10.1016/j.neuron.2018.06.019
26. Chen CA, Lattier J, Zhu W, et al. Retrospective analysis of a clinical exome sequencing cohort reveals the mutational spectrum and identifies candidate disease-associated loci for BAFopathies. *Genet Med*. Feb 2022;24(2):364-373. doi:10.1016/j.gim.2021.09.017
27. Nixon KCJ, Rousseau J, Stone MH, et al. A Syndromic Neurodevelopmental Disorder Caused by Mutations in SMARCD1, a Core SWI/SNF Subunit Needed for Context-Dependent Neuronal Gene Regulation in Flies. *Am J Hum Genet*. Mar 12 2019;doi:10.1016/j.ajhg.2019.02.001
28. Ninkovic J, Steiner-Mezzadri A, Jawerka M, et al. The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. *Cell Stem Cell*. Oct 3 2013;13(4):403-18. doi:10.1016/j.stem.2013.07.002
29. Diets IJ, Prescott T, Champaigne NL, et al. A recurrent de novo missense pathogenic variant in SMARCB1 causes severe intellectual disability and choroid plexus hyperplasia with resultant hydrocephalus. *Genet Med*. Jun 15 2018;doi:10.1038/s41436-018-0079-4
30. Cappuccio G, Sayou C, Tanno PL, et al. De novo SMARCA2 variants clustered outside the helicase domain cause a new recognizable syndrome with intellectual disability and blepharophimosis distinct from Nicolaides-Baraitser syndrome. *Genet Med*. Nov 2020;22(11):1838-1850. doi:10.1038/s41436-020-0898-y
31. Bidart M, El Atifi M, Miladi S, et al. Microduplication of the ARID1A gene causes intellectual disability with recognizable syndromic features. *Genet Med*. Jun 2017;19(6):701-710. doi:10.1038/gim.2016.180
32. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. May 2015;17(5):405-24. doi:10.1038/gim.2015.30

33. Ellard S BE, Berry I, Forrester N, Turnbull C, Owens M, Eccles DM, Abbs S, Scott R, Deans Z, Lester T, Campbell J, Newman W, McMullan D ACGS best practice guidelines for variant classification 2020: association for clinical genetics science (ACGS). May 25th, 2021. https://www.acgs.uk.com/quality/best-practice-guidelines/#VariantGuidelines
34. Houge G, Laner A, Cirak S, de Leeuw N, Scheffer H, den Dunnen JT. Stepwise ABC system for classification of any type of genetic variant. *Eur J Hum Genet*. Feb 2022;30(2):150-159. doi:10.1038/s41431-021-00903-z
35. Brandt T, Sack LM, Arjona D, et al. Adapting ACMG/AMP sequence variant classification guidelines for single-gene copy number variants. *Genet Med*. Feb 2020;22(2):336-344. doi:10.1038/s41436-019-0655-2
36. Davieson CD, Joyce KE, Sharma L, Shovlin CL. DNA variant classification-reconsidering "allele rarity" and "phenotype" criteria in ACMG/AMP guidelines. *Eur J Med Genet*. Oct 2021;64(10):104312. doi:10.1016/j.ejmg.2021.104312
37. Patel MJ, DiStefano MT, Oza AM, et al. Disease-specific ACMG/AMP guidelines improve sequence variant interpretation for hearing loss. *Genet Med*. Nov 2021;23(11):2208-2212. doi:10.1038/s41436-021-01254-2
38. Kohler S, Vasilevsky NA, Engelstad M, et al. The Human Phenotype Ontology in 2017. *Nucleic Acids Res*. Jan 04 2017;45(D1):D865-D876. doi:10.1093/nar/gkw1039
39. Dingemans AJM, Hinne M, Truijen KMG, et al. PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. *Nat Genet*. Aug 7 2023;doi:10.1038/s41588-023-01469-w
40. Roost MS, Slieker RC, Bialecka M, et al. DNA methylation and transcriptional trajectories during human development and reprogramming of isogenic pluripotent stem cells. *Nat Commun*. Oct 13 2017;8(1):908. doi:10.1038/s41467-017-01077-3
41. Levy MA, Relator R, McConkey H, et al. Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders. *Hum Mutat*. Nov 2022;43(11):1609-1628. doi:10.1002/humu.24446
42. Kerkhof J, Rastin C, Levy MA, et al. Diagnostic utility and reporting recommendations for clinical DNA methylation episignature testing in genetically undiagnosed rare diseases. *Genet Med*. Jan 18 2024:101075. doi:10.1016/j.gim.2024.101075
43. Levy MA, McConkey H, Kerkhof J, et al. Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. *HGG Adv*. Jan 13 2022;3(1):100075. doi:10.1016/j.xhgg.2021.100075
44. Aref-Eshghi E, Bend EG, Hood RL, et al. BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin-Siris and Nicolaides-Baraitser syndromes. *Nat Commun*. Nov 20 2018;9(1):4885. doi:10.1038/s41467-018-07193-y
45. Moffat JJ, Jung EM, Ka M, et al. The role of ARID1B, a BAF chromatin remodeling complex subunit, in neural development and behavior. *Prog Neuropsychopharmacol Biol Psychiatry*. Aug 24 2018;89:30-38. doi:10.1016/j.pnpbp.2018.08.021
46. Jung EM, Moffat JJ, Liu J, Dravid SM, Gurumurthy CB, Kim WY. Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. *Nat Neurosci*. Dec 2017;20(12):1694-1707. doi:10.1038/s41593-017-0013-0
47. Liu Y, Liu H, Sauvey C, Yao L, Zarnowska ED, Zhang SC. Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. *Nat Protoc*. Sep 2013;8(9):1670-9. doi:10.1038/nprot.2013.106
48. Pagliaroli L, Porazzi P, Curtis AT, et al. Inability to switch from ARID1A-BAF to ARID1B-BAF impairs exit from pluripotency and commitment towards neural crest formation in ARID1B-related neurodevelopmental disorders. *Nat Commun*. Nov 9 2021;12(1):6469. doi:10.1038/s41467-021-26810-x

49. Kang E, Kang M, Ju Y, et al. Association between ARID2 and RAS-MAPK pathway in intellectual disability and short stature. *J Med Genet*. Nov 2021;58(11):767-777. doi:10.1136/jmedgenet-2020-107111
50. Yuan X, Puvogel S, van Rhijn JR, et al. A human in vitro neuronal model for studying homeostatic plasticity at the network level. *Stem Cell Reports*. Nov 14 2023;18(11):2222-2239. doi:10.1016/j.stemcr.2023.09.011
51. Li C, Fleck JS, Martins-Costa C, et al. Single-cell brain organoid screening identifies developmental defects in autism. *Nature*. Sep 2023;621(7978):373-380. doi:10.1038/s41586-023-06473-y
52. Martins-Costa C, Wiegers A, Pham VA, et al. ARID1B controls transcriptional programs of axon projection in an organoid model of the human corpus callosum. *Cell Stem Cell*. May 6 2024;doi:10.1016/j.stem.2024.04.014
53. Mashtalir N, Dao HT, Sankar A, et al. Chromatin landscape signals differentially dictate the activities of mSWI/SNF family complexes. *Science*. Jul 16 2021;373(6552):306-315. doi:10.1126/science.abf8705
54. Wei J, Patil A, Collings CK, et al. Pharmacological disruption of mSWI/SNF complex activity restricts SARS-CoV-2 infection. *Nat Genet*. Mar 2023;55(3):471-483. doi:10.1038/s41588-023-01307-z
55. Santen GW, Aten E, Vulto-van Silfhout AT, et al. Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum Mutat*. Nov 2013;34(11):1519-28. doi:10.1002/humu.22394
56. Tsurusaki Y, Okamoto N, Ohashi H, et al. Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clin Genet*. Jun 2014;85(6):548-54. doi:10.1111/cge.12225
57. Sekiguchi F, Tsurusaki Y, Okamoto N, et al. Genetic abnormalities in a large cohort of Coffin-Siris syndrome patients. *J Hum Genet*. Sep 17 2019;doi:10.1038/s10038-019-0667-4
58. Wieczorek D, Bogershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet*. Dec 20 2013;22(25):5121-35. doi:10.1093/hmg/ddt366

Part I

Insights into clinical and epigenetic phenotypes


ID
ARID1B
CSS
ARID1B

Chapter 2

The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin–Siris syndrome

Pleuntje J van der Sluijs, Sandra Jansen, Samantha A Vergano, Miho Adachi-Fukuda, Yasemin Alanay, Adila AlKindy, Anwar Baban, Allan Bayat, Stefanie Beck-Wödl, Katherine Berry, Emilia K Bijlsma, Levinus A Bok, Alwin F J Brouwer, Ineke van der Burgt, Philippe M Campeau, Natalie Canham, Krystyna Chrzanowska, Yoyo W Y Chu, Brain H Y Chung, Karin Dahan, Marjan De Rademaeker, Anne Destree, Tracy Dudding-Byth, Rachel Earl, Nursel Elcioglu, Ellen R Elias, Christina Fagerberg, Alice Gardham, Blanca Gener, Erica H Gerkes, Ute Grasshoff, Arie van Haeringen, Karin R Heitink, Johanna C Herkert, Nicolette S den Hollander, Denise Horn, David Hunt, Sarina G Kant, Mitsuhiro Kato, Hülya Kayserili, Rogier Kersseboom, Esra Kilic, Malgorzata Krajewska-Walasek, Kylin Lammers, Lone W Laulund, Damien Lederer, Melissa Lees, Vanesa López-González, Saskia Maas, Grazia M S Mancini, Carlo Marcelis, Francisco Martinez, Isabelle Maystadt, Marianne McGuire, Shane McKee, Sarju Mehta, Kay Metcalfe, Jeff Milunsky, Seiji Mizuno, John B Moeschler, Christian Netzer, Charlotte W Ockeloen, Barbara Oehl-Jaschkowitz, Nobuhiko Okamoto, Sharon N M Olminkhof, Carmen Orellana, Laurent Pasquier, Caroline Pottinger, Vera Riehmer, Stephen P Robertson, Maian Roifman, Caroline Rooryck, Fabienne G Ropers, Monica Rosello, Claudia A L Ruivenkamp, Mahmut S Sagiroglu, Suzanne C E H Sallevelt, Amparo Sanchis Calvo, Pelin O Simsek-Kiper, Gabriela Soares, Lucia Solaeche, Fatma Mujgan Sonmez, Miranda Splitt, Duco Steenbeek, Alexander P A Stegmann, Constance T R M Stumpel, Saori Tanabe, Eyyup Uctepe, G Eda Utine, Hermine E Veenstra-Knol, Sunita Venkateswaran, Catheline Vilain,, Catherine Vincent-Delorme, Anneke T Vulto-van Silfhout, Patricia Wheeler, Golder N Wilson, Louise C Wilson, Bernd Wollnik, Tomoki Kosho, Dagmar Wieczorek, Evan Eichler, Rolph Pfundt, Bert B A de Vries, Jill Clayton-Smith, Gijs W E Santen

Published: Genet Med. 2018 Nov 8;21(6):1295–1307. doi: 10.1038/s41436-018-0330-z

ABSTRACT

Purpose: Pathogenic variants in *ARID1B* are one of the most frequent causes of intellectual disability (ID) as determined by large-scale exome sequencing studies. Most studies published thus far describe clinically diagnosed Coffin–Siris patients (ARID1B-CSS) and it is unclear whether these data are representative for patients identified through sequencing of unbiased ID cohorts (ARID1B-ID). We therefore sought to determine genotypic and phenotypic differences between ARID1B-ID and ARID1B-CSS. In parallel, we investigated the effect of different methods of phenotype reporting.

Methods: Clinicians entered clinical data in an extensive web-based survey.

Results: 79 ARID1B-CSS and 64 ARID1B-ID patients were included. CSS-associated dysmorphic features, such as thick eyebrows, long eyelashes, thick alae nasi, long and/or broad philtrum, small nails and small or absent fifth distal phalanx and hypertrichosis, were observed significantly more often ($p < 0.001$) in ARID1B-CSS patients. No other significant differences were identified.

Conclusion: There are only minor differences between ARID1B-ID and ARID1B-CSS patients. ARID1B-related disorders seem to consist of a spectrum, and patients should be managed similarly. We demonstrated that data collection methods without an explicit option to report the absence of a feature (such as most Human Phenotype Ontology-based methods) tended to underestimate gene-related features.

Keywords: ARID1B, Coffin–Siris syndrome, intellectual disability, bias

INTRODUCTION

The overall prevalence of intellectual disability (ID) has been estimated at around 1%. Given the increasing number of genes involved in ID, exome sequencing is becoming the first method of choice to identify the underlying genetic cause in patients with ID.[1] This unbiased approach detects clearly pathogenic variants in patients without the typically associated phenotype, indicating that variability in expression is higher than previously thought, confirming the existence of ascertainment bias. This bias may be mitigated by establishing the frequency of cardinal features in patients diagnosed through sequencing of unselected populations, although selection criteria for patients who undergo sequencing still cause a degree of bias. However, studies in unbiased populations suffer from another, less appreciated bias: they may *under*estimate the frequency of cardinal features because of the way that data are collected in large research studies. We would like to coin this phenomenon "phenotype underreporting bias": typically, a busy clinician is requested to supply several Human Phenotype Ontology (HPO) terms, and there is no guarantee that all features have been assessed or that all the clinical information is reported by the clinician. Consequently, it is not possible to make a distinction between the absence of a feature and unknown status, especially when specific diagnostic procedures are required to assess a feature (e.g., a magnetic resonance image [MRI] scan for agenesis of the corpus callosum [ACC]).

The *ARID1B* phenotype represents a good case study to investigate these biases, because the associated phenotypes range from clearly recognizable Coffin–Siris syndrome (ARID1B-CSS) to less specific ID[2,3] (ARID1B-ID). *ARID1B* is by far the most frequently mutated gene (51–75%) (refs. [4–6]) in Coffin Siris syndrome (CSS) (OMIM 135900) and large-scale exome sequencing studies invariably find that pathogenic variants in *ARID1B* are among the most frequently identified causes in unspecified ID cohorts (usually around 1%) (refs. [1,2]).

CSS is characterized by "developmental or cognitive delay, hypotonia, sparse scalp hair, distinctive facial features, aplasia or hypoplasia of the distal phalanx or nail of the fifth and additional digits, and hypertrichosis."[7] Approximately 70 ARID1B-CSS patients[4–6,8–12] have been described. Roughly 30 ARID1B-ID patients have been described in some detail[2,13–24] and an additional 23 patients with pathogenic variants in *ARID1B* were identified in genome-wide research studies[1,24] where detailed clinical information is generally unavailable. Data on the frequency of typical CSS features in the ARID1B-ID population are lacking because a substantial number of these patients were published before the link between *ARID1B* and CSS was known. For example, in Hoyer et al.[2] ACC was not specifically reported, although this is now known as a frequent feature in

ARID1B-CSS.[11] Therefore, the precise prevalence of these features cannot be estimated. The availability of unbiased information is crucial now that exome sequencing is being performed increasingly in neonatal and prenatal settings, and a reliable prognosis can only be given based on unbiased data.

The first aim of this study was therefore to overcome ascertainment and phenotype underreporting biases, by acquiring detailed clinical data of a large cohort of ARID1B-ID patients. The second aim was to determine whether the frequencies of features differ between ARID1B-CSS and ARID1B-ID patients, as would be expected at least for typical CSS features.

MATERIALS AND METHODS

Patient ascertainment

We developed a web-based survey (www.arid1bgene.com) based on previously reported features of *ARID1B* patients. This website is part of the Human Disease Genes website series (HDG), a collection of websites aimed at informing professionals about genes and copy-number variations and their clinical consequences (http://humandiseasegenes.com/). The survey was open to all clinicians of patients with pathogenic variants in *ARID1B*. Data were contributed by pediatric neurologists, pediatricians, and in most instances by clinical geneticists. Some clinicians contacted us, others were approached based on publications, conference presentations, submissions to databases like DECIPHER, or through large laboratories. We also included patients from our previous studies[4–6,9,11] and those referred to our national CSS expertise center in Leiden, The Netherlands.

The institutional review board of the Leiden University Medical Center, Leiden, The Netherlands provided an approval waiver for this study.

Data assessment

When only partial variant data was given, the remaining information was recovered using Alamut version 2.6.0. When the standard deviation score (SDS) was not reported, but raw data on weight, height, or occipital–frontal circumference (OFC) was available, the SDS was determined using published growth charts.[25]

After initial analyses we recontacted contributing clinicians by email to inquire about features frequently reported in open-ended questions.

We used a nominal *p* value of 0.05 as a cut-off for significance. However, given that we assessed 90 features, multiplicity correction by Bonferroni suggests that *p* values above 0.0006 (0.05/90) should be treated with care. All analyses were executed using SPSS version 23. R version 3.4.1 was used to create graphs, including the *survival* package (version 2.41-3).

RESULTS

Data from 143 individuals with pathogenic variants in *ARID1B* were included in the database. We received additional data regarding features recurrently indicated in our open-ended questions (Table 1 features marked with "++", Supplementary Table S4) of 95 patients. Supplementary Figure S1 displays facial photographs and hands or feet from two ARID1B-ID and two ARID1B-CSS patients. Parents provided consent for publication of these pictures.

Table 1. Clinical characteristics of *ARID1B* patients

	Total		**ARID1B-CSS**		**ARID1B-ID**			
Clinical features[d]	***n*= 143**	**%**	***n* =79**	**%**	***n* =64**	**%**	***p* value**	**Test**
Sex (female)	143	48.3	79	57.0	64	37.5	0.028	[a]
Growth parameters & development								
Gestational age, weeks (mean; SD)	133	39.0; 2.1	75	39.1; 2.0	58	38.9; 2.4	0.879	[b]
Birthweight (<–2 SDS)	129	5.4	74	6.8	55	3.6	0.506	[c]
Height at birth (<–2 SDS)	43	9.3	27	11.1	16	6.3	0.660	[c]
OFC at birth (<–2 SDS)	51	3.9	35	2.9	16	6.3	0.232	[c]
Age last measurements, years (median; min–max)	143	10; 0–51	79	10; 0–36	64	9; 0.5–51	0.682	[b]
Weight (<–2 SDS)	92	6.5	46	8.7	46	4.3	0.571	[c]
Height (<–2 SDS)	122	30.3	70	37.1	52	21.2	0.177	[c]
OFC (<–2 SDS)	105	2.9	63	3.2	42	2.4	0.670	[c]
Motor skills gross, delayed	103	99.0	46	97.8	57	100.0	0.447	[c]
Motor skills fine, delayed	100	95.0	44	97.7	56	92.9	0.381	[c]
Speech, delayed	131	65.6	75	68.0	56	62.5	0.106	[c]
Obstructive sleep apnea[e]	71	8.5	34	0.0	37	16.2	0.026	[c]
Laryngomalacia[e]	90	19.8	47	17.0	44	22.7	0.466	[a]
Feeding difficulties	121	69.4	62	62.9	59	76.3	0.111	[a]
Start of feeding difficulties	71		34		37		0.345	[c]
Birth		76.1		76.5		75.7		

Table 1 continues on next page

Table 1. Continued

	Total		ARID1B-CSS		ARID1B-ID			
Clinical features[d]	***n*= 143**	**%**	***n* =79**	**%**	***n* =64**	**%**	***p* value**	**Test**
Before 6 months		16.9		20.6		13.5		
After 6 months		7.0		2.9		10.8		
Duration of feeding problems	58		23		35		0.639	[c]
Brief		46.6		39.1		51.4		
Several years		6.9		8.7		5.7		
Ongoing		46.6		52.2		42.9		
Tube feeding	65	16.9	22	13.6	43	18.6	0.409	[c]
0–6 months		10.8		4.5		14.0		
6–12 months		3.1		4.5		2.3		
1–3 years		1.5		0.0		2.3		
Recurrent infections	75	57.3	30	63.3	45	53.3	0.391	[a]
Upper airway tract		17.3		10.0		22.2		
Lower airway tract		2.7		0.0		4.4		
ENT infections		12.0		10.0		13.3		
Otitis media		14.7		13.3		15.6		
Urinary tract		2.7		3.3		2.2		
Neurological features								
Intellectual disability	127	99.2	70	98.6	57	100.0	0.015	[c]
Normal–mild		3.1		1.4		5.3		
Mild		28.3		38.6		15.8		
Mild–moderate		15.7		8.6		24.6		
Moderate		22.0		22.9		21.1		
Moderate–severe		16.5		17.1		15.8		
Severe		13.4		10.0		17.5		
Hypotonia	116	81.0	71	80.3	45	82.2	0.795	[a]
Seizures	142	27.5	78	28.2	64	26.6	0.880	[c]
No seizures, but abnormal EEG		5.6		6.4		4.7		
Seizure frequency	18		9		9		0.671	[c]
Once		27.8		11.1		44.4		
Less than once a year		11.1		22.2		0.0		
Once a year		33.3		44.4		22.2		
Once a month		11.1		11.1		11.1		
1/2 a month		5.6		0.0		11.1		
≥2 per month		5.6		0.0		11.1		
Agenesis of the corpus callosum	101	28.7	62	29.0	39	28.2	0.344	[c]

Table 1. Continued

	Total		ARID1B-CSS		ARID1B-ID			
Clinical features[d]	***n*= 143**	**%**	***n* =79**	**%**	***n* =64**	**%**	***p* value**	**Test**
Partial/hypoplasia		13.9		17.7		7.7		
Neuroradiology	47	87.2	17	94.1	30	83.3	0.305	[a]
Delayed myelination		17.0		11.8		20.0		
Mega cisterna magna		14.9		23.5		10.0		
Colpocephaly		10.6		11.8		10.0		
Hypoplasia		4.3		0.0		6.7		
Enlarged Virchow–Robin spaces		4.3		5.9		3.3		
Vision and hearing impairments								
Vision impaired	109	48.6	62	45.2	47	53.2	0.406	[a]
Vision problems	68	70.6	33	78.8	35	62.9	0.320	[c]
Astigmatism		16.2		24.2		8.6		
Strabismus		30.9		36.4		25.7		
Optic nerve hypoplasia		2.9		6.1		0.0		
Nystagmus		8.8		6.1		11.4		
Refraction error		10.3		9.1		11.4		
Myopia	102	27.5	59	18.6	43	39.5	0.020	[a]
Hypermetropia	50	18.0	21	28.6	29	10.3	0.140	[c]
Abnormal eye exam	40	17.5	15	6.7	25	24.0	0.224	[c]
Hearing loss	122	22.1	71	18.3	51	27.5	0.157	[a]
Hearing loss, conductive		6.6		1.4		13.7		
Hearing loss, bilateral		11.5		8.5		15.7		
Hearing loss, unilateral		4.9		5.6		3.9		
Eartubes		4.9		4.2		5.9		
Start hearing problems, congenital	11	63.6	3	66.7	8	62.5	0.109	[c]
Hearing aid	5	80.0	2	100.0	3	66.7	0.665	[c]
Dysmorphic features								
Coarse face	121	81.8	62	90.3	59	72.9	0.013	[a]
Hairline (low anterior and/or posterior)	91	69.2	45	75.6	46	63.0	0.196	[a]
Scalp hair, abnormal	129	79.1	78	83.3	51	72.5	0.141	[a]
Sparse		58.1		62.8		51.0		
Forehead (broad or narrow)	95	42.1	49	28.6	46	56.5	0.000	[c]
Broad		22.1		6.1		39.1		
Narrow		20.0		22.4		17.4		
Eyelashes, long	131	63.4	79	75.9	52	44.2	0.000	[a]
Eyebrows, thick	134	81.3	78	91.0	56	67.9	0.001	[a]

Table 1 continues on next page

Table 1. Continued

	Total		ARID1B-CSS		ARID1B-ID			
Clinical features[d]	***n*= 143**	**%**	***n* =79**	**%**	***n* =64**	**%**	***p* value**	**Test**
Ptosis	133	20.3	77	20.8	56	19.6	0.872	[a]
Tear duct nonfunctioning or absent	93	15.1	54	16.7	39	12.8	0.609	[a]
Nasal bridge, abnormal	100	61.0	59	62.7	41	58.5	0.050	[a]
Wide		34.0		40.7		24.4		
Flat		21.0		20.3		22.0		
Broad		12.0		5.1		22.0		
Nasal tip, abnormal	129		76		53		0.002	[a]
Broad		58.1		61.8		52.8		
Upturned (anteverted nares)		29.5		39.5		15.1		
Nose, abnormal	83	47.0	43	53.5	40	40.0	0.022	[a]
Short		26.5		39.5		12.5		
Long		20.5		14.0		27.5		
Alae nasi, thick	107	55.1	69	66.7	38	34.2	0.001	[a]
Nasal base, broad	88	48.9	48	43.8	40	55.0	0.392	[a]
Philtrum, abnormal	109	78.9	72	86.1	37	64.9	0.001	[a]
Short		24.8		29.2		16.2		
Long		44.0		48.6		35.1		
Broad		34.9		44.4		16.2		
Mouth, large	131	68.7	75	76.0	56	58.9	0.037	[a]
Upper vermillion, abnormal	127	56.7	75	60.0	52	51.9	0.366	[a]
Thin		35.4		45.3		21.2		
Thick		21.3		14.7		30.8		
Lower vermillion, thick	125	69.6	76	78.9	49	55.1	0.005	[a]
Lower lip, drooping	71	56.3	30	76.7	41	41.5	0.004	[a]
Cleft palate/submucous cleft	90	6.7	35	14.3	55	1.8	0.031	[a]
Cleft palate		2.2		5.7		0.0		
Bifid uvula		2.2		5.7		0.0		
Submucous cleft		3.3		5.7		1.8		
High arched palate	85	16.5	31	22.6	54	13.0	0.250	[a]
Ears, abnormal	122	52.5	66	57.6	56	46.4	0.433	[a]
Low-set		9.8		13.6		5.4		
Posterior rotated		7.4		9.1		5.4		
Hypertrichosis	128	86.7	76	94.7	52	75.0	0.001	[a]
Musculoskeletal anomalies								
Scoliosis	123	26.0	70	27.1	53	24.5	0.743	[a]

Table 1. Continued

	Total		ARID1B-CSS		ARID1B-ID			
Clinical features[d]	***n*= 143**	**%**	***n* =79**	**%**	***n* =64**	**%**	***p* value**	**Test**
Pectus, excavatum	104	13.5	57	14.0	47	12.8	0.850	[a]
Primary dentition, delayed	65	44.6	40	50.0	25	36.0	0.313	[a]
Permanent dentition, delayed	33	48.5	18	33.3	15	66.7	0.056	[a]
Widely spaced teeth	72	41.7	40	40.0	32	43.8	0.748	[a]
Bone age, delayed	40	47.5	30	46.7	10	50.0	1.000	[c]
Joint laxity	88	60.2	52	61.5	36	58.3	0.763	[a]
Early arthritis	75	5.3	36	5.6	39	5.1	1.000	[c]
Clinodactyly	77	36.4	42	45.2	35	25.7	0.076	[a]
Short phalanges	49	34.7	34	41.2	15	20.0	0.151	[a]
Complete absent or small 5th distal phalanx	110	40.0	66	60.6	44	9.1	0.000	[c]
Prominent distal phalanges;	102	24.5	64	31.3	38	13.2	0.040	[a]
Prominent interphalangeal joints	103	21.4	64	28.1	39	10.3	0.032	[a]
Brachydactyly general	60	16.7	19	15.8	41	17.1	1.000	[c]
Brachydactyly fifth finger	68	30.9	22	50.0	46	21.7	0.018	[a]
Small nails	122	54.9	73	68.5	49	34.7	0.000	[a]
Which nails, 5th finger and/or toe	106	55.7	67	65.7	39	38.5	0.007	[a]
Which nails, all	53	11.3	29	20.7	24	0.0	0.027	[c]
Fetal finger pads	100	29.0	50	26.0	50	32.0	0.509	[a]
Intestinal								
Inguinal hernia	90	7.8	46	2.2	44	13.6	0.056	[c]
Intestinal problems	105	48.6	60	36.7	45	64.4	0.000	[a]
Constipation		30.5		21.7		42.2		
Gastroesophageal reflux		17.1		13.3		22.2		
Diarrhea		4.8		3.3		6.7		
Pyloric Stenosis		2.9		5.0		0.0		
Umbilical hernia		4.8		1.7		8.9		
Cardiac & genitourinary anomalies								
Cardiac anomalies	113	19.5	69	21.7	44	15.9	0.492	[a]
ASD		10.6		13.0		6.8		
VSD		5.3		5.8		4.5		
Renal anomalies	95	12.6	53	11.3	42	14.3	0.666	[c]
Renal sonography, abnormal	43	25.6	19	21.1	24	29.2	0.105	[c]
Cryptorchidism	65	55.4	28	39.3	37	67.6	0.023	[a]

Table 1 continues on next page

Table 1. Continued

	Total		ARID1B-CSS		ARID1B-ID			
Clinical features[d]	***n*= 143**	**%**	***n* =79**	**%**	***n* =64**	**%**	***p* value**	**Test**
Endocrinological abnormalities [e]								
Diabetes mellitus	71	7.0	43	7.0	28	7.1	1.000	[c]
Type 2 diabetes mellitus	4	75.0	2	100.0	2	50.0	1.000	[c]
Age (years) diagnosis (min–max)	2	18–46	1	18.0	1	46.0	0.317	[b]
Hypothyroidism	63	19.0	38	15.8	25	24.0	0.417	[a]
Age (years) diagnosis (median; min–max)	10	8; 1–40	4	4; 1.3–36.0	6	6; 1.0–40.0	0.394	[b]
Growth hormone deficiency	51	13.7	33	18.2	18	5.6	0.398	[c]
Growth hormone supplementation	50	12.0	31	16.1	19	5.3	0.387	[c]
Behavioral characteristics								
Behavioral abnormalities	71	83.1	28	85.7	43	81.4	0.945	[a]
Hyperkinetic		15.5		14.3		16.3		
Short attention		25.4		25.0		25.6		
Impulsiveness		14.1		14.3		14.0		
Obsessive		15.5		14.3		16.3		
Rigid		8.5		3.6		11.6		
Anger outbursts		16.9		10.7		20.9		
Aggressive		16.9		14.3		18.6		
Anxious		23.9		17.9		27.9		
Poor sociability		19.7		17.9		20.9		
Hyperactivity	63	42.9	25	60.0	38	31.6	0.026	[a]
High pain threshold[e]	47	40.4	28	28.6	19	57.9	0.044	[a]
Psychiatric disorders								
ADHD	48	33.3	16	50.0	32	25.0	0.083	[a]
Autistic traits	77	57.1	27	66.7	50	52.0	0.215	[a]
Malignancies	97	1.0	53	0.0	44	2.3	0.454	[c]

Only characteristics present in ≥5% of all patients or in either patient group, characteristics differing between groups, and distinctive features are shown.
ADHD attention deficit hyperactivity disorder, ARID1B-CSS patient group with a suspicion of Coffin–Siris syndrome before genetic testing, ARID1B-ID patient group with no suspicion of Coffin–Siris syndrome before genetic testing, *ASD* atrial septal defect, *CSS* Coffin–Siris syndrome, *EEG* electroencephalography, *ENT* ear nose throat, *OFC* occipitofrontal circumference, *SDS* standard deviation score, *VSD* ventricular septal defect.
[a]Chi-square.
[b]Mann–Whitney U.
[c]Fisher's.
[d]The total number of a feature can differ from the sum of subcategories, because in some cases it was possible to answer with more than one option or to report the existence of a feature without specifying.
[e]Data regarding these features were collected through email after first analyses.

Genotype

Figure 1 and Supplementary Table S1 provide an overview of the submitted pathogenic variants. Sixty-two patients were previously reported in literature.[4–6,9,19,26] Pathogenic variants were apparently de novo in all cases where parents could be tested (107/107). In two sisters the same pathogenic variant was found, while paternal DNA could not be obtained. Most pathogenic variants were frameshift or nonsense (*n* = 118), 18 were deletions involving multiple or all exons, and 7 involved canonical splice sites. One patient with a missense variant was initially submitted but later retracted when the variant turned out to be inherited from the unaffected father.

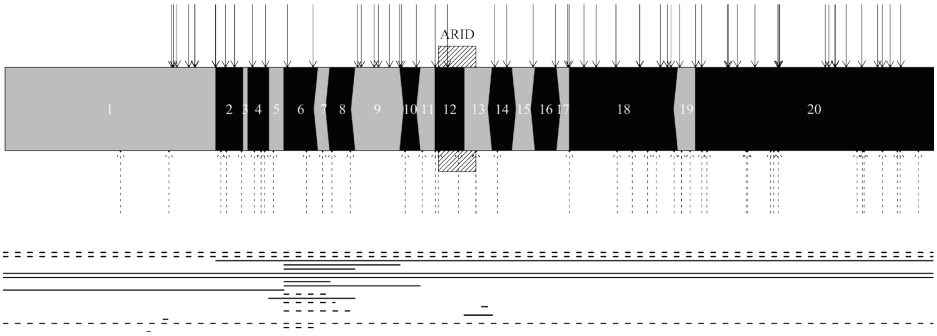


Figure 1. Overview of the location of pathogenic variants in *ARID1B*.
Numbers represent exon numbers and a graphical representation of in-frame and out-frame exons. When an exon ends with a complete codon, a vertical line is displayed. If it has one additional base an arrow to the left is displayed, and two additional bases are indicated by an arrow to the right. In-frame exons thus have the same boundary on both sides of the exon. Small variants (defined as events ≤20 bases) are identified by the arrows above the exon structure; larger variants are shown as lines under the intron–exon structure. All large events were deletions. Only unique variants are shown. Uninterrupted lines represent variants in ARID1B-CSS patients; interrupted lines represent variants in ARID1B-ID patients.

Phenotype

Patients' characteristics are summarized in Table 1 and Supplementary Table S2. Of the 143 individuals, 69 (48.3%) were female and the age at follow-up varied between 0 and 51 years with a median of 10 years. ARID1B-ID patients were less likely to be female than ARID1B-CSS patients (38 vs. 57%, p = 0.028). Two individuals have died, one of brain swelling potentially due to low-grade brain stem encephalitis that led to cardiorespiratory arrest at the age of 9 years, and one suddenly after the age of 24 years of unknown cause.

Diagnosis before genetic testing

Seventy-nine ARID1B-CSS patients were included. Most (91.5%) of the 64 ARID1B-ID patients were retrospectively classified to fit the CSS spectrum by the referring clinician. An overview of the clinical diagnoses prior to identification of the pathogenic variant can be found in Supplementary Table S3. No statistically significant differences in phenotypic features were found between patients who were retrospectively classified to fit the CSS spectrum by their referring clinicians (n = 54) and those who were not (n = 5).

Comparison of the ARID1B-ID and ARID1B-CSS groups

As expected, ARID1B-CSS patients more frequently displayed features associated with CSS than ARID1B-ID patients, including thick eyebrows, long eyelashes, thick alae nasi, long and/or broad philtrum, small nails and small or absent fifth distal phalanx, and hypertrichosis ($p < 0.0001–0.001$, Table 1).

ARID1B-ID patients appeared to have a higher prevalence of myopia (p = 0.020), cryptorchidism (p = 0.023), constipation (p = 0.002), sleep apnea (p = 0.026), hyperactivity (p = 0.026), and high pain threshold (p = 0.044), although these differences are not statistically significant under the Bonferroni adjusted significance level.

Regarding all other features, no significant differences were found between the ARID1B-CSS and ARID1B-ID groups (Table 1). Therefore, in the remainder of this section no distinction is made between both groups.

Overall phenotype

Growth and development

Histograms of the standard deviations of height, weight, and OFC are shown in Figure 2a–c. A height below –2 SDS was observed in 30.3%. Confirming our previous data,[11] the head circumference is normally distributed around 0 SDS and only 2.9% have an OFC below –2 SDS. Developmental milestones are shown in Figure 2d, e. Speech, and gross and fine motor skills were delayed in almost all patients. Figure 2e shows the Kaplan–Meier plot of the age at first words, and this plot suggests that about 25% of patients do not develop speech.

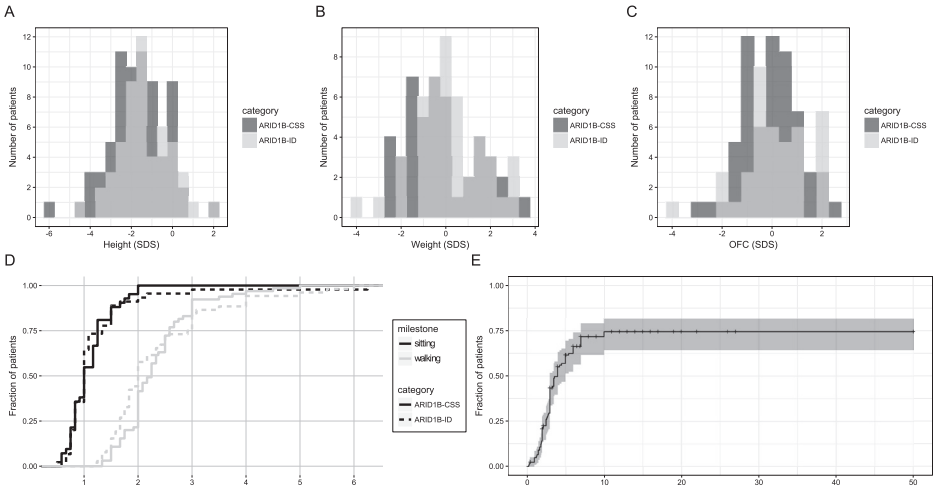


Figure 2. Biometry and developmental milestones.
Histograms of the standard deviation score (SDS) of (**a**) height, *n* = 122; (**b**) weight, *n* = 92; and (**c**) occipitofrontal circumference (OFC), *n* = 105. (**d**) Cumulative distribution of the developmental milestones walking (*n* = 117) and sitting (*n* = 85) for ARID1B-CSS and ARID1B-ID. (**e**) Kaplan–Meier plot for the whole cohort of the age at which patients spoke their first words (*n* = 126). Confidence intervals of Kaplan–Meier plots are generated by R's survfit function.

Feeding difficulties were frequent (69.4%) and led to tube feeding in 16.9% of patients. Bone age was delayed in 47.5% and scoliosis occurred in 26.0% of patients. Frequent infections (57.3%) and gastrointestinal problems (48.6%, mostly constipation) were also reported.

Congenital anomalies

Cardiac anomalies (19.5%), cryptorchidism (55.4%), laryngomalacia (19.8%), and a nonfunctioning or absent tear duct (15.1%) were frequently reported. The cardiac anomalies consisted mostly of abnormalities/defects of the cardiac septa (*n* = 19) and/or mitral or aortic valves (*n* = 6). Renal anomalies were present in 12.6% of patients. The most frequent renal anomalies were hydronephrotic kidney (*n* = 3) and nephrolithiasis (*n* = 3).

Neurological features

Almost all patients exhibited a variable degree of ID. Figure 3a, b show the distribution of reported ID severity and IQ scores (*n* = 35). Remarkably, several patients did have an IQ in the normal range.

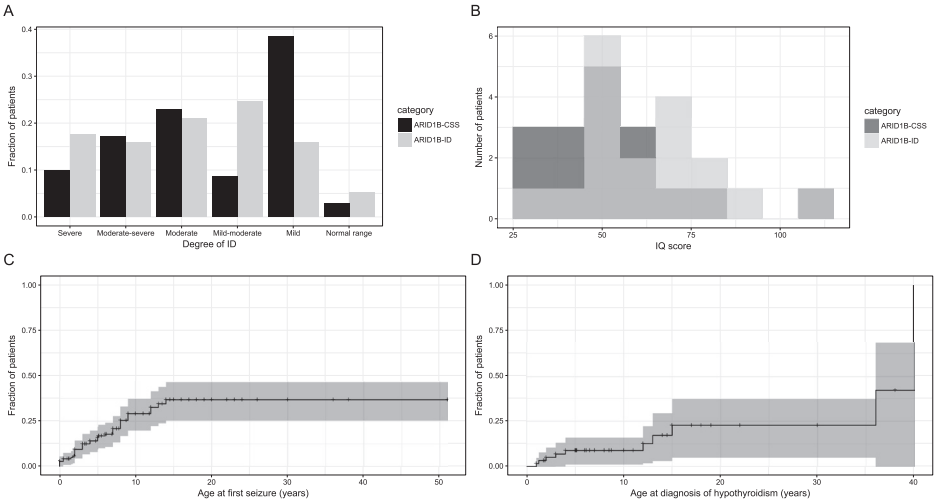


Figure 3. Degree of intellectual disability and survival analyses of seizures and hypothyroidism. (**a**) Intellectual disability (ID) category as assessed by the treating physician, *n* = 127; and (**b**) IQ scores (determined at different ages using different scales), *n* = 35 for ARID1B-CSS and ARID1B-ID patients. (**c**) Kaplan–Meier plot for age of onset of seizures, *n* = 37 and (**d**) Kaplan–Meier plot for the age at which hypothyroidism was diagnosed, *n* = 10. Confidence intervals of Kaplan–Meier plots are generated by R's survfit function.

Hypotonia occurred in 81.0% of patients and 27.5% suffered from seizures, while an additional 5.6% had an abnormal electroencephalogram (EEG). The age of onset of the seizures varied from 0 to 14 years, with a median of 4 years (Figure 3c). The Kaplan–Meier plot suggests that about 40% of patients are expected to experience one or more epileptic seizure during their life. All patients responded well to seizure medication (*n* = 18), while 4 did not receive medication. Complete or partial ACC (42.6%) and delayed myelination (17.0%) were the most prevalent cerebral anomalies.

Visual and hearing impairment

Vision problems (48.6%) were twice as frequent as hearing problems (22.1%). Myopia was the most reported vision problem (27.5%); severity ranged from –24 to –1.0 (median –6.35). Hearing loss was congenital in most patients (63.6%) and when present usually bilateral (51.9%); four patients needed hearing aids.

Endocrinological features

Hypothyroidism was documented in 19.0% and one patient had hypothyroidism for 2 years, which spontaneously resolved. The Kaplan–Meier plot of age at which hypothyroidism developed (Figure 3d) suggests that the prevalence may be 25%, but because the numbers are small there is a large confidence interval. Of 71 individuals,

5 (7.0%) were diagnosed with diabetes, 1 with hyperinsulinism, and 7/51 (13.7%) had a growth hormone (GH) deficiency, of which 6 received GH supplementation. We had no information on whether GH deficiency was partial or complete.

Effect of location of pathogenic variant

The location of the pathogenic variant did not appear to correlate with the severity of the phenotype (Supplementary Figure S2). Because it has been reported that patients with pathogenic variants in in-frame exons sometimes have less severe phenotypes because of naturally occurring exon skipping,[27] we compared the phenotypes of patients with pathogenic variants in in-frame exons versus those with pathogenic variants in out-frame exons. We classified patients with pathogenic variants in the first and last exon and whole-gene deletions as out-frame, because such pathogenic variants cannot be rescued by exon skipping. Most pathogenic variants were classified as out-frame (84.6%), which was expected because 86.1% of the coding region consists of out-frame exons. We did not find any differences in the phenotype between these two groups.

DISCUSSION

Genotype

Information on 143 patients with pathogenic variants in *ARID1B* was collected. In accordance with previous studies, all pathogenic variants were truncating (nonsense, frameshift, splice-site, and deletions of various numbers of exons including whole-gene deletions). In addition, four translocations[14,15,28,29] and three duplications[2,16,21] affecting *ARID1B* have been reported in literature. While previous reviews[30,31] have suggested that a pathogenic missense variant was reported in Tsurusaki et al.,[9] it is worth noting that a nonsense variant in *ARID1B* was also found in the same patient. Although parental DNA was not available for this patient, it seems most likely that the missense variant was harmless, while the nonsense variant was causal. Yu et al.[16] described two pathogenic de novo missense variants in patients with short stature, but without ID or speech delay. Mignot et al.[32] described a patient with mild ID and ACC who inherited a pathogenic missense variant from a mildly affected mother. Because the pathogenic variant (c.6092 T>C; p.Ile2031Thr) arose de novo in the mother, it is highly likely to be pathogenic. Thus, missense variants appear to be a much less common cause of ID than truncating variants. Missense variants identified in the absence of parental DNA should be interpreted with caution and are much more likely to be harmless. Care should also be taken with the interpretation of de novo missense variants in patients with ID.

As suggested by Johnston et al.,[33] it is important to take note if pathogenic variants in the same exon have been previously described when interpreting variants, because some annotated exons might in fact not be relevant for disease-associated transcripts. It is therefore noteworthy that we have not identified any pathogenic variants that only affect exon 3 (Figure 1). Exon 3 is a small in-frame exon and we previously indicated[11] that truncating variants in this exon are most likely benign. Indeed, Johnston et al.[33] report an inherited truncating variant in exon 3 in a patient without an *ARID1B*-related phenotype. Taking all evidence together, it is most likely that pathogenic variants in exon 3 are nonpathogenic, and that transcript NM_017519.2, which lacks exon 3 of NM_020732.3, is in fact the more relevant transcript. Another remarkable aspect that is apparent from Figure 1 is that no pathogenic variants have been reported in the first 849 bases of *ARID1B*. Although some frameshift variants are reported in gnomAD, these are dubious calls because the variants are in a low-complexity region. The lack of pathogenic variants may be by chance or due to sequencing difficulties in this GC-rich area, but it is also possible that there is an alternative start site that renders truncating variants in the first part of exon 1 neutral. Therefore, we advise caution when interpreting such variants.

Although a previous analysis seemed to suggest that pathogenic variants in the last exon might result in a more severe phenotype,[11] repeating this analysis with the current data did not confirm this suspicion (Supplementary Figure S2).

Penetrance

Penetrance of rare variants cannot be directly estimated, but the inheritance status and prevalence of apparently pathogenic variants in population databases can be used as proxy.[34] Pathogenic variants were de novo in all cases where inheritance could be established (n = 107). We did observe two sisters with the same pathogenic variant, and although this could be due to paternal inheritance, the most likely explanation is gonadal mosaicism, as previously described for *ARID1B*.[17]

The gnomAD browser may be viewed as a large population data set, consisting of exomes and genomes of unrelated, unaffected individuals sequenced as part of disease-specific and population genetic studies. In transcript NM_020732.2 (ENST00000346085), fewer missense variants in *ARID1B* are reported than would be expected if these were to occur randomly (ExAC accessed 10 April 2018; 744.4 expected, 555 observed, z-score of 3.39). Ten loss-of-function variants are reported; half of these might not be pathogenic because they are either in the start of exon 1, or in exon 3 (Supplementary Table S5). The other five variants, three splice-site variants and two frameshift variants, are potentially pathogenic. This might suggest that the penetrance of pathogenic variants

in *ARID1B* is not complete, and that these individuals had an average IQ as some of the individuals in our study. Therefore, while the penetrance of pathogenic variants in *ARID1B* appears to be very high, we still recommend parental testing, especially when future pregnancies are considered.

Phenotype

The main differences between ARID1B-CSS and ARID1B-ID seem to be related to dysmorphic features. Therefore, we conclude from this data set that the ARID1B-related disorders represent a spectrum. Not every feature is present in all patients, and depending on the combination of features present, and the experience of the clinician concerned, a patient might receive a clinical diagnosis of CSS. Although the patients in our previous *ARID1B-CSS* cohort[11] had an equal sex ratio, our current cohort finds significantly more females than males in the *ARID1B-CSS* group compared with the ARID1B-ID group. This may be because some features, most notably hypertrichosis, are easier to recognize as abnormal in females than in males.

Endocrinological features

Hypothyroidism was reported in 12/63 patients, and most were diagnosed before the age of 15 years (Figure 3d). Diabetes was diagnosed in 5/71 patients; 4 of these patients were reported to have type 2 diabetes and had a relatively high weight, while one was diagnosed with type 1 diabetes, and 1 patient was diagnosed with hyperinsulinism. This patient with hyperinsulinism has been described previously,[19] and one additional patient has been reported to have hyperinsulinism.[10] Of our patients, 7/51 were diagnosed with GH deficiency, and 6 of those received GH supplementation. Likewise, several patients with a growth delay due to GH deficiency were described[16] and a similar phenomenon was replicated in an Arid1b heterozygous mouse model.[35] Considering 30% of patients had short stature (<–2 SDS),[36] GH deficiency could be an underrecognized feature of ARID1B patients.

Cancer

Somatic variants in *ARID1B* have been associated with several types of cancer.[37] However, only one case is known of a patient with a pathogenic germline variant in *ARID1B* and (thyroid) cancer.[38] Similarly, in our cohort only one patient had malignancies. This boy had a Sertoli–Leydig cell tumor at the age of 3 and a temporal glioneuronal tumor at 12 years of age. No genetic testing was performed in this patient to detect the presence of specific tumor syndromes. Based on our patient cohort it seems unlikely that pathogenic germline variants in *ARID1B* confer an increased cancer risk, but longer follow-up of our patients is needed to make a definitive statement.

Phenotype delineation methods and biases

The increased awareness of ascertainment bias and the increasing number of newly discovered disease genes have resulted in new methods for data collection. Paper questionnaires and Excel spreadsheets are rapidly being replaced by Internet forms and HPO-based methods along with more formal registry software interfaces such as PhenoTips.[39] For the current study we have chosen to use an online questionnaire based on our previous data, with mostly fixed answers. More general genetic databases (such as DECIPHER[40]) accommodate inclusion of phenotypic data by using HPO terms. Whereas HPO-based methods allow for more straightforward identification of new, unexpected findings, our approach with specific questions allows calculation of true frequencies of features, because we can discriminate between absence of a feature and missing data. To investigate whether this influences results, we compared our data regarding the phenotype of *ARID1B* patients with the features reported for *ARID1B* patients in the DECIPHER database. It should be noted that there are some patients who are both in our cohort and the DECIPHER database. On 2 October 2017 DECIPHER contained 54 open-access *ARID1B* patients with 247 phenotypic features. Except for laryngomalacia and excessive salivation and/or drooling, all characteristics enriched in ARID1B patients in DECIPHER had been included in our questionnaire. The DECIPHER frequency of all reported features was lower than ours, a difference that was statistically significant for most (Supplementary Table S6).

In addition to other phenotypic features, growth and development are graphically represented on the DECIPHER website. We replicated some of these graphs and found only minor differences, which may be explained by the increased volume of data available to us (Supplementary Table S7 and Supplementary Figure S3).

Based on these results, we conclude that the DECIPHER database is a valuable tool to assess the potential features that ought to be included in gene-specific phenotype questionnaires, but the reported percentages are potential underestimations, likely due to phenotype underreporting bias owing to their data collection method. This bias could be mitigated by developing an adaptive questionnaire, so that submitters are requested to include or actively exclude features that have been mentioned several times in patients with the same underlying genetic cause.

Another bias that is present in most cohorts is an age-related bias. Most of our population is young, and this precludes us from detecting features that typically present at later age. This bias can partially be resolved by performing survival analysis, as we have done for speech, seizures, and hypothyroidism. In all cases there was a clear difference in our global frequency estimate, and the estimate provided by the Kaplan–

Meier plot. We therefore recommend that authors of clinical cohorts include the age of occurrence of age-dependent features and perform survival analysis.

Study limitations

Our study has several limitations. Data entry was performed by many different clinicians, which could have led to different interpretation of subjective questions, such as whether the patient fits the CSS spectrum. During this study, it became evident from literature and the DECIPHER database that previously unknown features, such as laryngomalacia, were part of the *ARID1B* spectrum and should have been added to our questionnaire. The effect of these incompletions was mitigated in part by the presence of an open field and our request to the contributing clinicians to send us an update on their patient's characteristics for frequently reported additional features.

Although we have attempted to account for different sources of bias, we cannot exclude that we currently overestimate the presence of some features, because our calculation is based on those patients having the feature, divided by those who are reported not to have the feature. It is possible that in some cases the absence of a feature is not consciously recognized or recorded in the clinical charts, which may result in scoring "unknown" rather than "absent" for a given feature. This bias can only be mitigated by thoroughly phenotyping every patient by a limited number of physicians, and is something we are planning to do with our national CSS cohort in the near future.

Conclusion

We conclude that the *ARID1B*-related disorders encompass a spectrum of features. The typical *ARID1B* patient has ID, feeding difficulties, laryngomalacia, speech delay, motor delay, hypertrichosis, and cryptorchidism. Our data suggest that endocrinological abnormalities, in particular hypothyroidism, may also be part of the *ARID1B* spectrum, but further research is needed to confirm this finding. There are few differences between ARID1B-CSS and ARID1B-ID, and we recommend that patients should be managed similarly. Based on the clinical data presented here we update our previous recommendations[11] for the management of *ARID1B* patients:

At the time of diagnosis

- Rule out congenital anomalies by performing renal and cardiac ultrasounds. Thoroughly check for cryptorchidism and laryngomalacia as indicated.
- Refer to clinical geneticist for counseling.

At the time of diagnosis and at follow-up

- Evaluate growth; consider referral to endocrinologist when significantly delayed.

- Evaluate hearing and vision regularly.
- Consider an EEG if there is a suspicion of seizures.
- Evaluate feeding problems. Offer dietary advice and feeding management. Consider evaluation by a gastroenterologist and/or a swallowing study.
- Treat constipation adequately.
- If indicated, early intervention using speech and/or physical therapy.
- Yearly evaluation of development by a specialized pediatrician and implementation of a specialized education plan for school or daycare as indicated.
- Periodic evaluations for scoliosis.
- On indication, determine thyroid status and glucose concentrations.

We also recommend periodic evaluation by a team of professionals specializing in *ARID1B*-related disorders (e.g., pediatrician, pediatric physiatrist, physical therapist, speech therapist, behavioral specialist, pediatric neurologist, clinical geneticist), which may take place at a distance.

ELECTRONIC SUPPLEMENTARY MATERIAL

Supplementary Information (1.6MB, pdf)

ACKNOWLEDGEMENTS

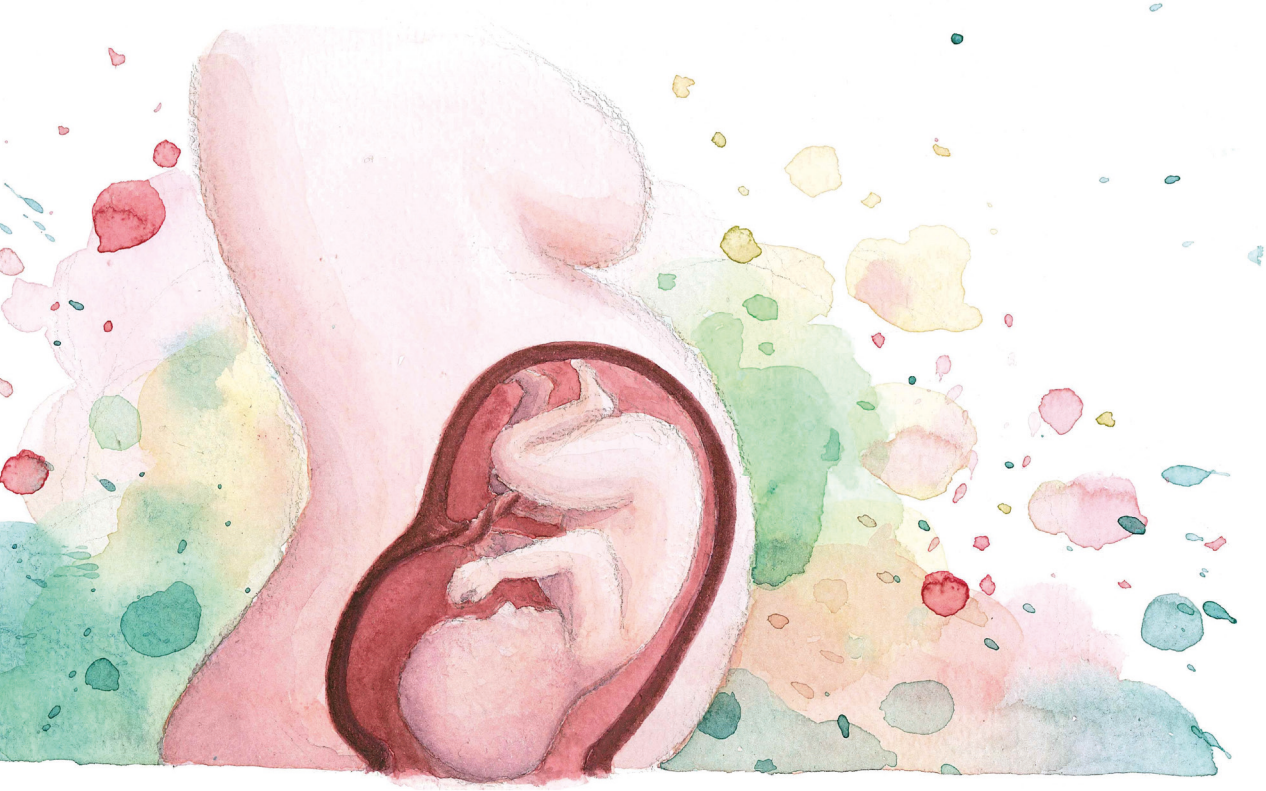
We are grateful for the assistance of Pepijn Cox in setting up the website www.arid1bgene.com. This study has made use of data generated by the Human Disease Genes website series, www.humandiseasegenes.com. This work was financially supported by grants from the Netherlands Organisation for Health Research and Development (917-86-319 to B.B.A.d.V., 912-12-109 to B.B.A.d.V.)

REFERENCES

1. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet. 2015;385:1305–14. doi: 10.1016/S0140-6736(14)61705-0.
2. Hoyer J, Ekici AB, Endele S, et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am J Hum Genet. 2012;90:565–72. doi: 10.1016/j.ajhg.2012.02.007.
3. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. Nature. 2015;519:223–8. doi: 10.1038/nature14135.
4. Santen GW, Aten E, Vulto-van Silfhout AT, et al. Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum Mutat. 2013;34:1519–28. doi: 10.1002/humu.22394.
5. Wieczorek D, Bogershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum Mol Genet. 2013;22:5121–35. doi: 10.1093/hmg/ddt366.
6. Tsurusaki Y, Okamoto N, Ohashi H, et al. Coffin-Siris syndrome is a SWI/SNF complex disorder. Clin Genet. 2014;85:548–54. doi: 10.1111/cge.12225.
7. Schrier Vergano S, Santen G, Wieczorek D, Wollnik B, Matsumoto N, Deardorff MA. Coffin-Siris syndrome. In: Pagon RA, Adam MP, Ardinger HH, et al., eds. GeneReviews. Seattle (WA) University of Washington 1993.
8. Santen GW, Aten E, Sun Y, et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. Nat Genet. 2012;44:379–80. doi: 10.1038/ng.2217.
9. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet. 2012;44:376–8. doi: 10.1038/ng.2219.
10. Vals MA, Oiglane-Shlik E, Noukas M, et al. Coffin-Siris syndrome with obesity, macrocephaly, hepatomegaly and hyperinsulinism caused by a mutation in the ARID1B gene. Eur J Hum Genet. 2014;22:1327–9. doi: 10.1038/ejhg.2014.25.
11. Santen GW, Clayton-Smith J. ARID1B-CSS consortium. The ARID1B phenotype: what we have learned so far. Am J Med Genet C Semin Med Genet. 2014;166C:276–89. doi: 10.1002/ajmg.c.31414.
12. Mari F, Marozza A, Mencarelli MA, et al. Coffin-Siris and Nicolaides-Baraitser syndromes are a common well recognizable cause of intellectual disability. Brain Dev. 2015;37:527–36. doi: 10.1016/j.braindev.2014.08.009.
13. Nord AS, Roeb W, Dickel DE, et al. Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. Eur J Hum Genet. 2011;19:727–31. doi: 10.1038/ejhg.2011.24.
14. Halgren C, Kjaergaard S, Bak M, et al. Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. Clin Genet. 2012;82:248–55. doi: 10.1111/j.1399-0004.2011.01755.x.
15. Sim JC, White SM, Fitzpatrick E, et al. Expanding the phenotypic spectrum of ARID1B-mediated disorders and identification of altered cell-cycle dynamics due to ARID1B haploinsufficiency. Orphanet J Rare Dis. 2014;9:43. doi: 10.1186/1750-1172-9-43.
16. Yu Y, Yao R, Wang L, et al. De novo mutations in ARID1B associated with both syndromic and nonsyndromic short stature. BMC Genom. 2015;16:701. doi: 10.1186/s12864-015-1898-1.
17. Ben-Salem S, Sobreira N, Akawi NA, et al. Gonadal mosaicism in ARID1B gene causes intellectual disability and dysmorphic features in three siblings. Am J Med Genet A. 2016;170A:156–61. doi: 10.1002/ajmg.a.37405.

18. Smith JA, Holden KR, Friez MJ, Jones JR, Lyons MJ. A novel familial autosomal dominant mutation in ARID1B causing neurodevelopmental delays, short stature, and dysmorphic features. Am J Med Genet A. 2016;170:3313–8. doi: 10.1002/ajmg.a.37945.
19. Sonmez FM, Uctepe E, Gunduz M, et al. Coffin-Siris syndrome with cafe-au-lait spots, obesity and hyperinsulinism caused by a mutation in the ARID1B gene. Intractable Rare Dis Res. 2016;5:222–6. doi: 10.5582/irdr.2014.01040.
20. Takenouchi T, Yoshihashi H, Sakaguchi Y, et al. Hirschsprung disease as a yet undescribed phenotype in a patient with ARID1B mutation. Am J Med Genet A. 2016;170:3249–52. doi: 10.1002/ajmg.a.37861.
21. Seabra CM, Szoko N, Erdin S, et al. A novel microduplication of ARID1B: clinical, genetic, and proteomic findings. Am J Med Genet A. 2017;173:2478–84. doi: 10.1002/ajmg.a.38327.
22. Zahir FR, Mwenifumbo JC, Chun HE, et al. Comprehensive whole genome sequence analyses yields novel genetic and structural insights for intellectual disability. BMC Genom. 2017;18:403. doi: 10.1186/s12864-017-3671-0.
23. Zweier M, Peippo MM, Poyhonen M, et al. The HHID syndrome of hypertrichosis, hyperkeratosis, abnormal corpus callosum, intellectual disability, and minor anomalies is caused by mutations in ARID1B. Am J Med Genet A. 2017;173:1440–3. doi: 10.1002/ajmg.a.38143.
24. Trujillano D, Bertoli-Avella AM, Kumar Kandaswamy K, et al. Clinical exome sequencing: results from 2819 samples reflecting 1000 families. Eur J Hum Genet. 2017;25:176–82. doi: 10.1038/ejhg.2016.146.
25. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. BMC Pediatr. 2013;13:59. doi: 10.1186/1471-2431-13-59.
26. Baban A, Moresco L, Divizia MT, et al. Pituitary hypoplasia and growth hormone deficiency in Coffin-Siris syndrome. Am J Med Genet A. 2008;146A:384–8. doi: 10.1002/ajmg.a.32111.
27. Kapahnke M, Banning A, Tikkanen R. Random splicing of several exons caused by a single base change in the target exon of CRISPR/Cas9 mediated gene knockout. Cells. 2016;5:E45.
28. Backx L, Seuntjens E, Devriendt K, Vermeesch J, Van Esch H. A balanced translocation t(6;14) (q25.3; q13.2) leading to reciprocal fusion transcripts in a patient with intellectual disability and agenesis of corpus callosum. Cytogenet Genome Res. 2011;132:135–43. doi: 10.1159/000321577.
29. Malli T, Duba HC, Erdel M, et al. Disruption of the ARID1B and ADAMTS6 loci due to a t(5;6) (q12.3; q25.3) in a patient with developmental delay. Am J Med Genet A. 2014;164A:3126–31. doi: 10.1002/ajmg.a.36738.
30. Sim JC, White SM, Lockhart PJ. ARID1B-mediated disorders: mutations and possible mechanisms. Intractable Rare Dis Res. 2015;4:17–23. doi: 10.5582/irdr.2014.01021.
31. Son EY, Crabtree GR. The role of BAF (mSWI/SNF) complexes in mammalian neural development. Am J Med Genet C Semin Med Genet. 2014;166C:333–49. doi: 10.1002/ajmg.c.31416.
32. Mignot C, Moutard ML, Rastetter A, et al. ARID1B mutations are the major genetic cause of corpus callosum anomalies in patients with intellectual disability. Brain. 2016:e64.
33. Johnston JJ, Lewis KL, Ng D, et al. Individualized iterative phenotyping for genome-wide analysis of loss-of-function mutations. Am J Hum Genet. 2015;96:913–25. doi: 10.1016/j.ajhg.2015.04.013.
34. Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. Genet Med. 2013;15:478–81. doi: 10.1038/gim.2012.164.
35. Celen C, Chuang JC, Luo X, et al. Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. Elife. 2017;6:e25730.

36. Pedicelli S, Peschiaroli E, Violi E, Cianfarani S. Controversies in the definition and treatment of idiopathic short stature (ISS) J Clin Res Pediatr Endocrinol. 2009;1:105–15. doi: 10.4008/jcrpe.v1i3.53.
37. Biegel JA, Busse TM, Weissman BE. SWI/SNF chromatin remodeling complexes and cancer. Am J Med Genet C Semin Med Genet. 2014;166C:350–66. doi: 10.1002/ajmg.c.31410.
38. Vengoechea J, Carpenter L, Zarate YA. Papillary thyroid cancer in a patient with interstitial 6q25 deletion including ARID1B. Am J Med Genet A. 2014;164A:1857–9. doi: 10.1002/ajmg.a.36515.
39. Girdea M, Dumitriu S, Fiume M, et al. PhenoTips: patient phenotyping software for clinical and research use. Hum Mutat. 2013;34:1057–65. doi: 10.1002/humu.22347.
40. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 2009;84:524–33. doi: 10.1016/j.ajhg.2009.03.010.



Chapter 3

Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort

Pleuntje J van der Sluijs, Marieke Joosten, Caroline Alby, Tania Attié-Bitach, Kelly Gilmore, Christele Dubourg, Mélanie Fradin, Tianyun Wang, Evangeline C Kurtz-Nelson, Kaitlyn P Ahlers, Peer Arts, Christopher P Barnett, Myla Ashfaq, Anwar Baban, Myrthe van den Born, Sarah Borrie, Tiffany Busa, Alicia Byrne, Miriam Carriero, Claudia Cesario, Karen Chong, Anna Maria Cueto-González, Jennifer C Dempsey, Karin E M Diderich, Dan Doherty, Stense Farholt, Erica H Gerkes, Svetlana Gorokhova, Lutgarde CP Govaerts, Pernille A Gregersen, Scott E Hickey, Mathilde Lefebvre, Francesca Mari, Jelena Martinovic, Hope Northrup, Melanie O'Leary, Kareesma Parbhoo, Sophie Patrier, Bernt Popp, Fernando Santos-Simarro, Corinna Stoltenburg, Christel Thauvin-Robinet, Elisabeth Thompson, Anneke T Vulto-van Silfhout, Farah R Zahir, Hamish S Scott, Rachel K Earl, Evan E Eichler, Neeta L Vora, Yael Wilnai, Jessica L Giordano, Ronald J Wapner, Jill A Rosenfeld, Monique C Haak, Gijs W E Santen

Published: Genet Med. 2022 May 18;24(8):1753–1760. doi: 10.1016/j.gim.2022.04.010

ABSTRACT

Purpose: Genome-wide sequencing is increasingly being performed during pregnancy to identify the genetic cause of congenital anomalies. The interpretation of prenatally identified variants can be challenging and is hampered by our often limited knowledge of prenatal phenotypes. To better delineate the prenatal phenotype of Coffin-Siris syndrome (CSS), we collected clinical data from patients with a prenatal phenotype and a pathogenic variant in one of the CSS-associated genes.

Methods: Clinical data was collected through an extensive web-based survey.

Results: We included 44 patients with a variant in a CSS-associated gene and a prenatal phenotype; 9 of these patients have been reported before. Prenatal anomalies that were frequently observed in our cohort include hydrocephalus, agenesis of the corpus callosum, hypoplastic left heart syndrome, persistent left vena cava, diaphragmatic hernia, renal agenesis, and intrauterine growth restriction. Anal anomalies were frequently identified after birth in patients with *ARID1A* variants (6/14, 43%). Interestingly, pathogenic *ARID1A* variants were much more frequently identified in the current prenatal cohort (16/44, 36%) than in postnatal CSS cohorts (5%–9%).

Conclusion: Our data shed new light on the prenatal phenotype of patients with pathogenic variants in CSS genes.

Keywords**:** *ARID1A*, *ARID1B* BAFopathy, BAF-complex, Fetal, *SMARCA4*, *SMARCB1*

INTRODUCTION

Coffin-Siris syndrome (CSS) (OMIM 135900 and others) is characterized by neurodevelopmental delay and may be accompanied by hypoplasia of the fifth digit and/or nail, distinct facial features, hypertrichosis, and a varying degree of congenital anomalies.[1–6] In the current literature, most patients present with no prenatal ultrasound (US) anomalies. Congenital diaphragmatic hernia,[7] hypoplastic left heart syndrome (HLHS),[8] and intrauterine growth restriction (IUGR)[9] have been occasionally described, whereas agenesis of the corpus callosum (ACC)[10] is observed more frequently.

Application of prenatal genome-wide sequencing is rapidly becoming standard care.[11–15] The interpretation of prenatally identified variants can be challenging because of the limited knowledge of prenatal phenotypes, because most genome-wide sequencing in diagnostics and research has focused on postnatal phenotypes. CSS is no exception to this rule. For example, we recently detected a novel de novo missense variant in *SMARCA4* (OMIM 614609) in a fetus with HLHS. HLHS had been described in 1 patient with *ARID1A*-CSS[8] before. With this little information, it was difficult to be certain about the pathogenicity of the variant, showing a typical difficulty encountered in current genetic practice.

This study aimed to increase knowledge of the prenatal phenotype of CSS by presenting a detailed description of the prenatal features in a cohort of patients with 1 or more prenatal US anomalies and a pathogenic variant in one of the CSS-related genes.

MATERIALS AND METHODS

Patient collection

Patients were collected through our outpatient CSS expertise center in the Leiden University Medical Center, Leiden, the Netherlands. Additional patients were included through international contacts, from physicians contacting us for second opinions, through contacts with the Baylor Genetics Laboratory, and through laboratories who reported pathogenic variants in LOVD[16] and ClinVar.[17] We also included patients who were presented at conferences and via the fetal sequencing consortium, which consists of an international group of medical professionals who discuss results of prenatal diagnostics in a biweekly case meeting.

Patient selection

Inclusion criteria were (1) the presence of any anomaly that was mentioned in US reports, excluding so called soft-markers (eg, echogenic focus in the left ventricle)

and (2) a pathogenic or likely pathogenic genetic variant in a CSS-associated gene (*ARID1A, ARID1B, ARID2, BICRA, DPF2, SMARCA2* duplication, *SMARCA4, SMARCB1, SMARCC2, SMARCD1, SMARCE1, SOX11, SOX4*). The exclusion criterium was the ACC as the sole prenatal US anomaly in combination with a pathogenic *ARID1B* variant, because the association between the ACC and pathogenic *ARID1B* variants is well established.[10,18]

Timing of the identification of the genetic variant could be prenatal, postnatal, or postmortem.

Data collection

Data were collected through an online questionnaire. If required, genetic data were completed using Alamut version 2.11.

A patient was considered to have IUGR if it was reported as such by the contributing clinician.

RESULTS

Genetic variants

In total, 44 patients with variants in *ARID1A* (*n* = 16, 36%), *ARID1B* (*n* = 19, 43%), *SMARCA4* (*n* = 3, 7%), and *SMARCB1* (*n* = 6, 14%) and a prenatal phenotype detected by US were identified (Supplemental Table 1). Variants were classified using the American College of Medical Genetics and Genomics/Association for Molecular Pathology criteria.

Phenotype

The clinical features are summarized in Tables 1 and 2. The median time in weeks of gestation at which the first US anomalies were detected tends to be lowest among patients with *ARID1A* variants (14+5 weeks), followed by *ARID1B* (20+0 weeks), *SMARCA4* (20+0 weeks), and *SMARCB1* (24+0 weeks) (Supplemental Figure 1). Unless stated otherwise, frequencies in the text refer to anomalies that were detected prenatally. The total prevalence of features (detected prenatally or by postnatal assessment) in this cohort is shown in Table 1.

Genotype-phenotype correlation

ARID1A

Of the 16 *ARID1A* pregnancies, 11 (69%) were terminated on request of the parents because the anomalies diagnosed by US carried a poor prognosis. One pregnancy led

to an intrauterine fetal death, and one resulted in a child who died shortly after birth. Brain anomalies were most frequently present (10/15, 73%), especially hydrocephalus (*n* = 7) and ACC (*n* = 5). Cardiac anomalies were reported in 60% (9/15). HLHS was detected in 2 fetuses and hypoplastic right heart in 1 fetus. Persistent left vena cava (*n* = 3) and double outlet right ventricle (*n* = 3) were cardiovascular anomalies reported in a subset of fetuses. In one case, the ventricular septal defect and patent ductus arteriosus spontaneously resolved at 4 months of age. Renal anomalies were observed in 27% (4/15) of cases, including 1 case with unilateral renal agenesis who received a peritoneal dialysis catheter at the age of 1 year and a renal transplant 6 years later. Diaphragmatic hernia was detected in 13% (2/15) of fetuses, and in 3 cases, a diaphragmatic hernia was identified during postnatal investigation (34%, 5/15). Postnatal investigation also revealed supernumerary ribs in 2 patients and anal anomalies in 43% (6/14); 4 patients had anal atresia and 2 patients had an anteriorly placed anus.

ARID1B

Despite excluding isolated ACC cases for *ARID1B*, brain anomalies were detected in 76% (13/17) of *ARID1B* cases, but only in a subset during pregnancy (29.4%, 5/17). Most often, ACC was not detected prenatally. Cardiac anomalies were present in 19% (3/16), including 1 fetus with HLHS. Fetus 18 with a complex cardiac defect also had a 22q11 deletion. This patient died during her first day of life because of respiratory insufficiency. Another fetus had unilateral renal agenesis.

SMARCA4

Of the 3 fetuses with a *SMARCA4* variant, 2 had a severe cardiac anomaly (ie, HLHS). One of these patients died postnatally as a result of her congenital anomalies. The fetus without a severe cardiac anomaly had hydrops, diaphragmatic hernia, double renal collecting system, and abnormal feet position. Parents decided to continue this pregnancy after they received the genetic results at 19+4 weeks of gestation.

SMARCB1

IUGR was reported for 50% (3/6) of fetuses with a *SMARCB1* variant. Brain anomalies were reported for 2 fetuses. One fetus had an abnormal cardiac position (ie, ventricular axis and apex to the right of the midline) with biventricular hypertrophy and a minor size discrepancy of the great arteries, and 1 fetus had a cardiac septal defect that was detected postnatally. Two fetuses had hydronephrosis on prenatal US, of which 1 case resolved postnatally. During postnatal assessment, it appeared that both males with *SMARCB1* variants had cryptorchidism.

Table 1. Clinical characteristics of included patients

	ARID1A					*ARID1B*				
Clinical Features	***n* = 16**[a]	**Detected by US**[b]	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**	***n* = 19**	**Detected by US**[b]	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**
Sex, female	15	—	—	9	60%	19	—	—	5	26%
First US anomaly, time of AD in days, *n* (min-max, median)	12	—	—	(84–224, 103)		14	—	—	(60–259, 140)	
Anomalies										
Placenta anomaly	15	1	—	1	7%	19	—	—	—	—
Single umbilical artery	15	7	—	7	47%	19	3	1	4	21%
Oligohydramnios	15	3	—	3	20%	19	3	—	3	16%
Polyhydramnios	15	—	—	—	—	19	3	—	3	16%
IUGR	15	2	—	2	13%	19	4	—	4	21%
Short femur	15	2	—	2	13%	19	—	—	—	—
Brain anomaly	15	1-	1	11	73%	17	5	8	13	76%
Hydrocephalus	15	7	—	7	47%	17	1	—	1	6%
ACC	14	5	1	6	43%	17	—	7	7	41%
Ventriculomegaly	14	4	—	4	29%	17	2	—	2	12%
Dandy Walker malformation	14	1	2	3	21%	17	—	—	—	—
intracranial cysts	14	2	2	4	29%	17	1	—	1	6%
Enlarged cisterna magna	14	—	1	1	7%	17	2	—	2	12%
Cerebellar vermis atrophy or hypoplasia	14	2	3	5	36%	17	1	1	2	12%
Increased NT	15	7	—	7	47%	19	3	—	3	16%
Cardiovascular anomaly	15	9	3	12	80%	16	3	3	6	38%
Vascular anomaly	15	—	—	8	53%	15	2	2	4	27%
Aortic arch anomaly	15	2	1	3	20%	14	2	—	2	14%
Involving the vena cava	15	3	1	4	27%	14	—	—	—	—
Patent ductus arteriosus[c]	15	n.a.	1	1	7%	14	1	—	1	7%
Cardiac anomaly	15	9	1	10	67%	15	3	—	3	20%
Hypoplastic left heart	15	2	—	2	13%	16	1	—	1	6%
Hypoplastic right heart	15	1	—	1	7%	16	—	—	—	—
Atrioventricular canal	15	1	1	2	13%	14	—	—	—	—

	SMARCA4					*SMARCB1*				
	***n* = 3**	**Detected by US[b]**	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**	***n* = 6**	**Detected by US[b]**	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**
	3	—	—	2	67%	6	—	—	4	67%
	3	—	—	(82–145, 140)		6	—	—	(140–231, 168)	
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	1	—	1	17%
	3	—	—	—	—	6	3	—	3	50%
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	3	1	4	67%
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	5	2	—	2	40%
	3	—	—	—	—	5	1	1	2	40%
	3	—	—	—	—	5	—	—	—	—
	3	—	—	—	—	5	1	—	—	—
	3	—	—	—	—	5	1	1	2	40%
	3	—	—	—	—	5	—	1	1	20%
	3	1	—	1	33%	6	—	—	—	—
	3	2	—	2	67%	6	1	1	2	33%
	3	1	—	1	33%	6	1	—	1	17%
	3	1	—	1	33%	6	1	—	1	17%
	3	—	—	—	—	6	—	—	—	—
	3	n.a.	—	—	—	6	n.a.	—	—	—
	3	2	—	2	67%	6	1	1	2	33%
	3	2	—	2	67%	6	—	—	—	—
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	—	—	—	—

Table 1 continues on next page

Table 1. Continued

	ARID1A					*ARID1B*				
Clinical Features	***n* = 16**[a]	**Detected by US**[b]	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**	***n* = 19**	**Detected by US**[b]	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**
DORV	15	3	—	3	20%	14	—	—	—	—
Septal defect (isolated or in combination with other structural cardiac anomalies)	15	4	5	9	60%	14	—	2	2	14%
Valve defect	15	—	—	3	20%	14	—	—	—	—
Cardiac position/Dextrocardia	15	1	—	1	7%	14	—	—	—	—
Cardiovascular intervention	4	—	—	—	—	14	—	—	1	7%
Diaphragmatic hernia	15	2	3	5	33%	16	1	—	1	6%
Anal anomaly	14	n.a.	6	6	43%	15	n.a.	—	—	—
Renal anomaly	15	4	3	7	47%	16	1	1	2	13%
Renal dysplasia	14	—	3	3	21%	15	—	1	1	7%
Renal agenesis	14	3	—	2–3	13–20%	15	—	—	—	—
Hydronephrosis[c]	4	—	—	—	—	15	1	—	1	7%
Genitourinary anomaly	12	—	5	5	42%	15	—	8	8	53%
Urinary anomaly	14	—	2	2	14%	15	—	3	3	20%
Genital anomaly	14	—	2	2	14%	15	—	3	3	20%
Cryptorchidism[c]	6	n.a.	—	—	—	14	n.a.	10	10	71%
Hypospadia	6	n.a.	1	1	17%	14	—	—	—	—
Digit anomaly	14	n.a.	10	10	71%	11	—	4	4	36%
Nail anomaly	11	n.a.	6	6	55%	11	n.a.	5	5	45%
TOP	11[i]					—			—	—
Mean time AD, days	10	—	—	(101–161, 132.1)	67%	1	—	—	(144)	5%
IUFD										
Mean time AD, days	1	1	—	(126–126, 126)	7%	19	—	—	—	—
Died	17	—	—	13	87%	19		—	2	11%

ACC, agenesis of the corpus callosum; *AD*, amenorrhea duration; *DORV*, double outlet right ventricle; *IUFD*, intrauterine fetal death; *IUGR*, intrauterine growth restriction; *n.a.*, not applicable, *NT*, nuchal translucency; *TOP*, termination of pregnancy; *US*, ultrasound.

[a] One fetal case/TOP no additional information available.

[b] Separately reported as US anomaly.

[c] Only data of non-TOP and IUFD cases.

	SMARCA4					*SMARCB1*				
	***n* = 3**	**Detected by US[b]**	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**	***n* = 6**	**Detected by US[b]**	**Detected by Postnatal Investigation**	**Total Affected**	**% Total**
	3	—	—	—	—	6	—	—	—	—
	3	1	—	1	33%	6	—	1	1	17%
	3	2	—	2	67%	6	—	—	—	—
	3	—	—	—	—	6	1	—	1	17%
	3	—	—	1	33%	6	—	—	—	—
	3	2	—	2	67%	6	—	—	—	—
	3	n.a.	—	—	—	6	n.a.	—	—	—
	3	1	—	1	33%	6	3	—	1	17%
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	—	—	—	—
	3	—	—	—	—	6	2	—	1	17%
	3	—	—	—	—	6	—	3	3	50%
	3	—	—	—	—	6	—	1	1	17%
	3	—	—	—	—	3	—	2	2	67%
	1	n.a.	n.a.	n.a.	—	2	n.a.	2	2	100%
	1	n.a.	n.a.	n.a.	—	2	n.a.	1	1	50%
	2	—	2	2	100%	4	—	3	3	75%
	2	n.a.	2	2	100%	2	n.a.	1	1	50%
	—			—	—	—			—	—
	—	—	—	—	—	1	—	—	(159)	17%
	3	—	—	—	—	6	—	—	—	—
	3	—	—	1	33%	6	—	—	1	17%

Table 2. Distribution of genetic variants per gene within our cohort and within published CSS cohorts

Gene	Total	%	CSS cohort Santen et al[a,b]		CSS cohort Wieczorek et al[a,b]		CSS cohort Tsurusaki/ Sekiguchi et al[a,b]		Cumulative Frequency of CSS Cohorts Combined[a,b]		P-value Fisher's exact
ARID1A	16/44	36.4%	4/45	8.9%	1/21	4.8%	6/78	7.7%	11/144	7.6%	<.01
ARID1B	19/44	43.2%	28/45	62.2%	14/21	76.2%	48/78	61.5%	90/144	62.5%	.04
SMARCA4	3/44	6.8%	4/45	8.9%	0/21	0.0%	7/78	9.0%	11/144	7.6%	1.00
SMARCB1	6/44	13.6%	4/45	8.9%	1/21	4.8%	8/78	10.3%	13/144	9.0%	.40

CSS, Coffin-Siris syndrome.
[a] Only patients with pathogenic variants in a CSS-associated gene were taken into account.
[b] Numbers may not add up to 100%, because pathogenic variants in other CSS-associated genes were detected.

DISCUSSION

Genotype

In total, we identified 44 CSS cases with prenatal US anomalies, making this the largest case series to date. *ARID1A* (*n* = 16) or *ARID1B* (*n* = 19) variants were most frequent in this prenatal cohort. The frequency of *ARID1A* variants in particular was remarkable. *ARID1A* variants were overrepresented among the cases that did not survive into the neonatal period (76.5%, *n* = 13 for *ARID1A* vs 23.5%, *n* = 4 for the non-*ARID1A* group, *P* < .001). Furthermore, the median gestational age at which the anomalies were detected was earlier (14+5 weeks for *ARID1A* vs 20+0 weeks for the non-*ARID1A* group) (Supplemental Figure 1), although the difference was not significant (*P* = .17). This may fit with our previous finding that *ARID1A* variants in postnatal cases are frequently present in a mosaic state,[1] which suggests that nonmosaic pathogenic *ARID1A* variants may lead to an embryonically severe or lethal phenotype. The likely explanation for the relative lack of postnatal *ARID1A* variants (Table 2) is that most fetuses with full pathogenic variants display severe congenital anomalies, leading to a large proportion of pregnancy terminations as well as fetal and neonatal demise. Thus far, such cases were less likely to undergo extensive diagnostic sequencing. Contrary to our previous findings where all 4 patients with *ARID1A* variants appeared mosaic in lymphocytes,[1] all patients in this cohort appear to have the pathogenic variant in 50% of sequence reads.

Phenotype

Frequent findings in fetal cases with pathogenic *ARID1A* variants, such as hydrocephalus, hypoplastic left or right heart syndrome, renal agenesis, diaphragmatic hernia, and postnatally detected anal anomalies, have not been previously recognized as part of the CSS phenotypic spectrum. Sporadically, patients have been reported with one of

these anomalies and a pathogenic variant in a CSS-associated gene.[6,7,18–20] Unilateral renal agenesis was reported among patients with clinically diagnosed CSS in the era before identification of CSS genes.[21,22]

Pregnancies were terminated in 68.8% (11/16) of the *ARID1A* cases, which makes it impossible to assess severity of developmental delay or treatment response postnatally. Compared with reported congenital and/or structural anomalies among ARID1A patients,[1–4,23] the severity of the congenital anomalies in our *ARID1A* cases appears to be more severe (Table 1).

The postnatal *ARID1B* phenotype has been extensively reported on by our group and others.[10,18,24,25] Although we excluded patients with ACC as the sole prenatal presenting feature, 29.4% (5/17) had 1 or more brain anomalies. IUGR was detected more frequently in our cohort compared with previously reported frequencies in patients[18] (Fisher's exact, $P = .04$), which might indicate that it is more common in the subset of patients with congenital anomalies. The frequency of cardiac anomalies in the fetal series does not differ significantly from previously reported frequencies in patients[18] (Fisher's exact, $P = 1.00$).

It is remarkable that all *SMARCA4* fetuses had severe congenital anomalies, which have not previously been reported among patients with *SMARCA4* variants. Li et al[9] do report other prenatal findings in 6 of 13 patients with *SMARCA4* variants (ie, IUGR [$n = 4$], cardiac anomaly [$n = 2$], and ACC [$n = 1$]).

The most frequent US anomaly in *SMARCB1* variants was IUGR, which is a frequent, nonspecific prenatal US finding. Where the postnatal *SMARCB1* patient group usually presents with a phenotype considered at the severe end of the CSS spectrum,[1,6,25] there does not to appear to be a more severe prenatal presentation for this gene compared with patients with *ARID1A*, *ARID1B*, or *SMARCA4* variants. However, we show here that growth delay,[1,4] which is a prevalent feature among reported patients with CSS and is often more extreme in *SMARCB1* patients, is detected prenatally as IUGR in several patients.

We did not identify patients with a prenatal presentation and a pathogenic variant in the other CSS-related genes (*ARID2, BICRA, DPF2, SMARCA2* duplication, *SMARCC2*, *SMARCD1, SMARCE1, SOX11, SOX4*) during this study. It is possible that prenatal US anomalies are less prevalent among these patients. The number of patients known with pathogenic variants in these genes is, however, still relatively small, so a prenatal phenotype may be identified in the future.

Not all cases in this cohort underwent a similar US examination, which is reflected by the fact that not all anomalies were detected prenatally. Table 1, however, gives an overview of the potential for prenatal detection if every fetus would have been examined on the level of a tertiary referral center. Because not every case underwent extensive US examination, this may be an underestimation.

CONCLUSION

This study highlights the prenatal phenotypic spectrum of CSS in a prenatal cohort with a molecularly confirmed pathogenic variant in a CSS-associated gene. Pathogenic variants in *ARID1A* were identified much more frequently in this group than in cohorts of postnatal patients with CSS. Frequently observed prenatal anomalies include hydrocephalus, ACC, HLHS, double outlet right ventricle, persistent left vena cava, congenital diaphragmatic hernia, renal agenesis, and IUGR, indicating that the prenatal phenotype of CSS significantly differs from the postnatal phenotype. This difference may be explained by an increased rate of pregnancy termination for fetuses at the severe end of the spectrum and in increased chance of fetal or postnatal demise. Genome-wide sequencing has been less frequently applied in such cases until now, leaving this part of the CSS spectrum undiscovered.

SUPPLEMENTARY MATERIAL

Supplemental Table 1

NIHMS1809356-supplement-Supplemental_Table_1.pdf (321.6KB, pdf)

ACKNOWLEDGEMENTS

We would like to thank Ileana Minguel for her assistance in the collection of patient data and Aude Tessier and Anne-Claire Brehin for referring cases and conducting genetic analysis.

This work was supported, in part, by grants from the National Institutes of Health (Grant No. R01 MH101221 [to E.E.E.]). E.E.E. is an investigator of the Howard Hughes Medical Institute.

Sequencing and analysis for individual 30 was provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung, and Blood Institute (Grant Nos. UM1 HG008900 and R01 HG009141).

Sequencing and analysis of cases 5 and 18 was funded by the National Institute of Child Human Development (Grant Nos. K23 HD088742 and R01 HD105868 [to N.L.V.]).

CONFLICT OF INTEREST

All authors declare no conflicts of interest.

ETHICS DECLARATION

The Institutional Review Board of Leiden University Medical Center, Leiden, The Netherlands provided approval waivers for using de-identified data and publishing aggregated data (no: G18.098 and G21.129) without obtaining specific informed consent. Clinical data was de-identified by allocating patient numbers. Where possible, informed consent was obtained by the referring clinician.

DATA AVAILABILITY

De-identified patient data will be made available on request to the corresponding author.

REFERENCES

1. Santen GWE, Aten E, Vulto-van Silfhout AT, et al. Coffin–Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum Mutat. 2013;34(11):1519–1528. 10.1002/humu.22394.
2. Wieczorek D, Bögershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin–Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum Mol Genet. 2013;22(25): 5121–5135. 10.1093/hmg/ddt366.
3. Tsurusaki Y, Okamoto N, Ohashi H, et al. Coffin–Siris syndrome is a SWI/SNF complex disorder. Clin Genet. 2014;85(6):548–554. 10.1111/cge.12225.
4. Kosho T, Okamoto N, Coffin-Siris Syndrome International Collaborators. Genotype-phenotype correlation of Coffin–Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. Am J Med Genet C Semin Med Genet. 2014;166C(3): 262–275. 10.1002/ajmg.c.31407.
5. Kosho T, Okamoto N, Ohashi H, et al. Clinical correlations of mutations affecting six components of the SWI/SNF complex: detailed description of 21 patients and a review of the literature. Am J Med Genet A. 2013;161A(6):1221–1237. 10.1002/ajmg.a.35933.
6. Schrier Vergano S, Santen GWE, Wieczorek D, Wollnik B, Matsumoto N, Deardorff MA. Coffin–Siris syndrome. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. GeneReviews® [Internet]. Seattle, WA: University of Washington; 2013.
7. Sweeney NM, Nahas SA, Chowdhury S, et al. The case for early use of rapid whole-genome sequencing in management of critically ill infants: late diagnosis of Coffin–Siris syndrome in an infant with left congenital diaphragmatic hernia, congenital heart disease, and recurrent infections. Cold Spring Harb Mol Case Stud. 2018;4(3):a002469. 10.1101/mcs.a002469.
8. Petrikin JE, Cakici JA, Clark MM, et al. The NSIGHTI-randomized controlled trial: rapid whole-genome sequencing for accelerated etiologic diagnosis in critically ill infants. NPJ Genom Med. 2018:3:6. 10.1038/s41525-018-0045-8.
9. Li D, Ahrens-Nicklas RC, Baker J, et al. The variability of SMARCA4-related Coffin–Siris syndrome: do nonsense candidate variants add to milder phenotypes? Am J Med Genet A. 2020;182(9):2058–2067. 10.1002/ajmg.a.61732.
10. Mignot C, Moutard ML, Rastetter A, et al. ARID1B mutations are the major genetic cause of corpus callosum anomalies in patients with intellectual disability. Brain. 2016;139(ll):e64. 10.1093/brain/awwl81.
11. Lord J, McMullan DJ, Eberhardt RY, et al. Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. Lancet. 2019;393(101731):747–757. 10.1016/S0140-6736(18)31940-8.
12. Stanley KE, Giordano J, Thorsten V, et al. Causal genetic variants in stillbirth. N Engl J Med. 2020:383(12): 1107–1116. 10.1056/NEJMoal908753.
13. Ferretti L, Mellis R, Chitty LS. Update on the use of exome sequencing in the diagnosis of fetal abnormalities. Eur J Med Genet. 2019:62(8): 103663. 10.1016/j.ejmg.2019.05.002.
14. Petrovski S, Aggarwal V, Giordano JL, et al. Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. Lancet. 2019:393(101731):758–767. 10.1016/S0140-6736(18)32042-7.
15. Audain E, Wilsdon A, Breckpot J, et al. Integrative analysis of genomic variants reveals new associations of candidate haploin sufficient genes with congenital heart disease. PLoS Genet. 2021;17(7):el009679. Published correction appears in. PLoS Genet. 2021;17(9):el009809. 10.1371/journal.pgen.1009679.

16. Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat. 2011:32(5):557–563. 10.1002/humu.21438.
17. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014;42(Database issue):D980–D985. 10.1093/nar/gktl113.
18. van der Sluijs PJ, Jansen S, Vergano SA, et al. The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin–Siris syndrome. Genet Med. 2019:21(6): 1295–1307. Published correction appears in. Genet Med. 2019:21(9):2160–2161. 10.1038/s41436-018-0330-z.
19. Lefebvre M, Bruel AL, Tisserant E, et al. Genotype-first in a cohort of 95 fetuses with multiple congenital abnormalities: when exome sequencing reveals unexpected fetal phenotype-genotype correlations. J Med Genet. 2021;58(6):400–413. 10.1136/jmedgenet2020-106867.
20. Shang L, Cho MT, Retterer K, et al. Mutations in ARID2 are associated with intellectual disabilities. Neurogenetics. 2015; 16(4):307–314. 10.1007/s10048-015-0454-0.
21. Al Mosawi AJ. Genetic drift. Letter from Baghdad: Coffin–Siris syndrome in a girl with absent kidney. Am J Med Genet A. 2006; 140(16): 1789–1790. 10.1002/ajmg.a.31337.
22. Pollono D, Drut R, Cecotti N, Pollono A. Neuroblastoma in a patient with Coffin–Siris syndrome. Fetel Pediatr Pathol. 2009:28(4): 185–191. 10.1080/15513810902984129.
23. Miraldi Utz V, Brightman DS, Sandoval MA, Hufnagel RB, Saal HM. Systemic and ocular manifestations of a patient with mosaic ARID1A-associated Coffin–Siris syndrome and review of select mosaic conditions with ophthalmic manifestations. Am J Med Genet C Semin Med Genet. 2020;184(3):644–655. 10.1002/ajmg.c.31839.
24. Santen GWE, Clayton-Smith J, ARID1B-CSS consortium. The ARID1B phenotype: what we have learned so far. Am J Med Genet C Semin Med Genet. 2014;166C(3):276–289. 10.1002/ajmg.c.31414.
25. Vasko A, Drivas TG, Schrier Vergano SA. Genotype-phenotype correlations in 208 individuals with Coffin–Siris syndrome. Genes (Basel). 2021;12(6):937. 10.3390/genes12060937.

Chapter 4

A Case Series of Familial *ARID1B* Variants Illustrating Variable Expression and Suggestions to Update the ACMG Criteria

Pleuntje J van der Sluijs, Mariëlle Alders, Alexander J M Dingemans, Kareesma Parbhoo, Bregje W van Bon, Jennifer C Dempsey, Dan Doherty, Johan T den Dunnen, Erica H Gerkes, Ilana M Millller, Stephanie Moortgat, Debra S Regier, Claudia A L Ruivenkamp, Betsy Schmalz, Thomas Smol, Kyra E Stuurman, Catherine Vincent-Delorme, Bert B A de Vries, Bekim Sadikovic, Scott E Hickey, Jill A Rosenfeld, Isabelle Maystadt*, Gijs W E Santen*

* Co-last authorship.

Published: Genes (Basel). 2021 Aug 20;12(8):1275. doi: 10.3390/genes12081275

ABSTRACT

ARID1B is one of the most frequently mutated genes in intellectual disability (~1%). Most variants are readily classified, since they are de novo and are predicted to lead to loss of function, and therefore classified as pathogenic according to the American College of Medical Genetics and Genomics (ACMG) guidelines for the interpretation of sequence variants. However, familial loss-of-function variants can also occur and can be challenging to interpret. Such variants may be pathogenic with variable expression, causing only a mild phenotype in a parent. Alternatively, since some regions of the *ARID1B* gene seem to be lacking pathogenic variants, loss-of-function variants in those regions may not lead to *ARID1B* haploinsufficiency and may therefore be benign. We describe 12 families with potential loss-of-function variants, which were either familial or with unknown inheritance and were in regions where pathogenic variants have not been described or are otherwise challenging to interpret. We performed detailed clinical and DNA methylation studies, which allowed us to confidently classify most variants. In five families we observed transmission of pathogenic variants, confirming their highly variable expression. Our findings provide further evidence for an alternative translational start site and we suggest updates for the ACMG guidelines for the interpretation of sequence variants to incorporate DNA methylation studies and facial analyses.

Keywords: *ARID1B*, Coffin–Siris syndrome, ACMG guidelines, inherited, familial, variable expression, non-pathogenic, intellectual disability

INTRODUCTION

Pathogenic variants in *ARID1B* are one of the top hits (~1%) in large-scale sequencing studies in intellectual disability (ID) populations [1,2] and *ARID1B* is also the most frequently mutated gene in Coffin–Siris syndrome (CSS) (OMIM 135900) (51–76%) [3,4,5]. With the increasing application of genome-wide screening approaches, such as exome and genome sequencing, more and more variants in *ARID1B* are identified. Usually variants in *ARID1B* can be readily classified, since most identified variants are de novo and predicted to lead to loss of function. According to the ACMG guidelines [6], this results in one very strong (PVS1) and one strong criterium (PS2), leading to a clinical classification as a pathogenic variant.

ARID1B variants are regularly identified in patients without specific features [7], complicating the interpretation of variants of uncertain significance (VUS). Indeed, we have previously reported on the variable expression of the *ARID1B* phenotype by showing IQ values within the normal range [7,8], suggesting that a variant classified as pathogenic could be inherited from a mildly affected parent. To make matters more complex, we also noticed that no pathogenic variants have been reported in the 5' end of the 1542 bp exon 1 nor in the small 39 bp in-frame exon 3 of transcript NM_020732.3 [7], suggesting that loss-of-function variants in these regions may not be disease causing. Thus, interpretation of some loss-of-function variants, in particular those that are not proven de novo, may be challenging because of the difficulty to distinguish between variants not leading to disease and familial variants with variable expression.

To illustrate the difficulties in interpreting inherited loss-of-function variants in *ARID1B*, we describe here 12 index cases with potential loss-of-function variants that were either inherited or have an unknown inheritance. Furthermore, we describe our diagnostic evaluation to determine their true effect on clinical outcomes. This evaluation includes algorithms that combine regular facial recognition techniques with the modelling of human facial dysmorphisms [9], and DNA methylation patterns [10]. Together, this allowed for determining variants with variable expression versus non-pathogenic variants. Based on this cohort, we recommend updates for the ACMG guidelines for interpretation of sequence variants.

MATERIALS AND METHODS

Patient Ascertainment

Patients with potential loss-of-function variants that were either inherited or of unknown inheritance were included. We included patients from colleagues who had approached us about such variants, screened variants using ClinVar, and identified additional patients through contacts with the Baylor Genetics Laboratories.

The institutional review board (IRB) of the Leiden University Medical Center, Leiden, The Netherlands, provided an approval waiver for this study.

Clinical Data Collection

Clinical information was collected through an online questionnaire or by asking the clinician to give a brief summary of their patient's phenotype.

Transcript

Clinicians reported the genetic variant on transcript NM_020732.3, which is also used throughout the manuscript. The Matched Annotation from NCBI and EMBL-EBI (MANE) project (https://www.ncbi.nlm.nih.gov/refseq/MANE/ (accessed on 20 July 2021)), which aims to select a well-supported transcript for all genes to facilitate communication about variants, recently selected the NM_001374828.1 transcript for *ARID1B*. This transcript differs from the NM_020732.3 transcript, which is most widely used in clinical literature thus far, in three aspects: (1) the encoded protein has an N-terminal extension since an upstream start codon in exon 1 is used; (2) exon 3 of NM_020732.3 is not included; and (3) it has an additional in-frame 159 bp exon 11. The MANE project selected this transcript based on the conservation of the coding region and support for the transcription start site from FANTOM5 CAGE data as well as direct evidence of translation from uniquely mapping peptides (Human Peptide Atlas build 502). The MANE project also assigns the longest conserved coding region supported by evidence for the MANE Select. Where relevant, we mention the consequences of this transcript throughout the manuscript.

DNA Methylation

Analysis of the DNA methylation array data was performed by the clinical bioinformatics laboratory using Illumina Infinium EPIC arrays as previously described [11,12]. Methylation data for each sample were compared to the established DNA methylation signatures for *ARID1B* among 43 other disorders that were part of the EpiSign V2 clinical test. EpiSign analysis utilized the EpiSign Knowledge Database EKD, a clinical database with >5000 peripheral blood DNA methylation profiles, including disorder-specific reference cohorts and normal (general population samples with various age and racial backgrounds)

controls housed at the London Health Sciences Centre Molecular Diagnostics Laboratory (https://www.lhsc.on.ca/palm/molecular.html (accessed on 15 January 2020)). Individual DNA methylation data for each subject were compared to the EKD using the Support Vector Machine (SVM)-based classification algorithm for EpiSign disorders. A Methylation Variant Pathogenicity (MVP) score was generated, ranging between 0 and 1, representing the confidence of the prediction for the specific class the SVM was trained to detect. Classification for a specific EpiSign disorder includes an MVP score assessment with a general threshold of >0.5 for positive, < 0.1 for negative and 0.1–0.5 for inconclusive or low-confidence samples; hierarchical clustering; and the multidimensional scaling (MDS) of a subject's methylation data relative to the disorder-specific EpiSign probe sets and controls. A detailed description of this analytics protocol was described previously [11,13].

Analyses of Facial Features

Computer Vision Algorithms

Using previously reported computer vision algorithms [9,14], we first assessed whether the photographs of the *ARID1B* patients with a confirmed pathogenic variant and who were referred to our national CSS expertise center in Leiden clustered compared to age-, ethnicity-, sex-matched controls with an ID. We performed this analysis with different age-groups (0–10 years, 10–18 years and 18+ years) as well, since we have observed in our CSS population that with advancing age, facial features may become less typical. We next investigated how photographs from patients with inherited *ARID1B* variants clustered. If more facial photographs of one patient were available, the photograph with the best quality for the analysis was chosen for the initial assessment.

Face2Gene

We also analysed the photographs of our *ARID1B* familial cases by uploading them in Face2Gene (FDNA Inc., Boston, MA, USA). We assessed the rank of CSS in the top 30 suggested syndromes and their similarity to CSS according to Face2Gene on a scale of 0–1. Face2Gene analyses were performed in April–May 2021.

RESULTS

We collected 12 previously unreported index cases associated with a potential loss of function, as an inherited variant in *ARID1B* or a variant with unknown inheritance (Table 1 and Figure 1). Eight families had an inherited variant, one family consisted of two siblings of whom the mother was tested negative for the variant and there were

three index cases with variants of unknown inheritance. The carrier parent of Case 11 is inconclusively mosaic (see also case description) and the parent of Case 12 is mosaic; the carrier parents of all other patients with inherited variants appeared to have non-mosaic or germline variants. In the DNA from peripherical blood, there was no indication for mosaicism in the other carrier parents.

DNA Methylation

ARID1B DNA methylation EpiSign analysis was performed for 5/12 cases where DNA and parental consent were available. A typical BAFopathy methylation pattern was detected in three familial cases (Table 1).

Table 1. An overview of the included patients' variants in *ARID1B* and a summary of the variant and phenotype assessments. (Transcript NM_020732.3).

Case	Exon	cDNA	Protein Change	Inheritance	Variant	GnomAD (1 May 2021)	DNA Methylation Pattern
1	1	c.3G > A	p.(Met1?)	Not maternal	Start codon	No	no BAFopathy
2	1	c.361C > T	p.(Gln121*)	Paternal	Nonsense	No	n.a.
3	1	c.363_364insG	p.(Gln122fs*110)	Unknown	Frameshift	7x	n.a.
4	1	c.521dup	p.(Pro177fs*55)	Maternal	Frameshift	No	BAFopathy
5	1	c.1029_1056del	p.(Ala349Metfs*11)	Maternal	Frameshift	No	BAFopathy
6	1	c.1044_1071del	p.(Ala349Metfs*11)	Maternal	Frameshift	No	n.a.
7	1	c.1044_1062del	p.(Gly351Alafs*12)	Paternal	Frameshift	No	BAFopathy
8	3–4	exon 3–4 deletion	p.?	Paternal	In-frame deletion	-	no BAFopathy
9	7	c.2371+2T > C	r.spl?	Unknown	Splice site	No	n.a.
10	8	c.2372-2A > C	r.spl?	Maternal	Splice site	1x	n.a.
11	18	c.4870C > T	p.(Arg1624*)	Father is inconclusively mosaic. Mother is negative. Siblings.	Nonsense	No	n.a.
12	20	c.6322C > T	p.(Gln2108*)	Paternal, mosaic father	Nonsense	No	n.a.

Abbreviations: n.a.: not available; P: pathogenic; LP: likely pathogenic; VUS: variant of uncertain significance; LB: likely benign.

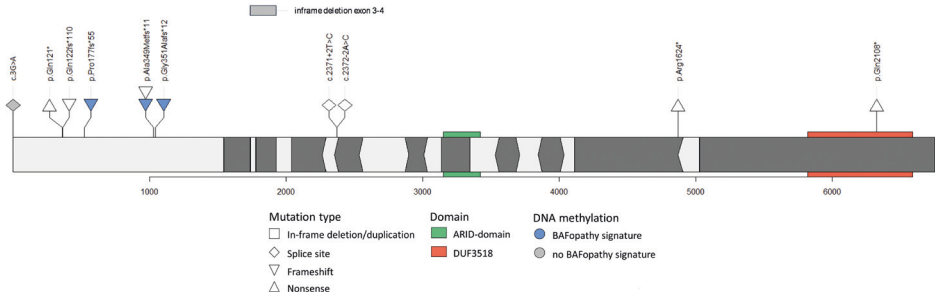


Figure 1. Overview of the location of variants in *ARID1B* and DNA methylation results of the included cases (transcript NM_020732.3).

Phenotype Suggestive for an *ARID1B* Related Disorder	Photograph Clusters with *ARID1B* Patients	CSS in Face2Gene (Rank/ Similarity)	ACMG Criteria	Interpretation ACMG	Updated ACMG Criteria	Interpretation Updated ACMG	Expert Opinion
+	+	2/0.32	PVS1, PM2	LP	PM2, PM7, PP4	VUS	VUS
+/−	n.a.	n.a.	PVS1, PM2	LP	PM2, PM7	VUS	VUS
−	n.a.	n.a.	PVS1, BS2	VUS	PM4, BS2	VUS	LB
+	+(parent−)	1/0.13 (−/0.13)	PVS1, PM2	LP	PS5, PM2, PM7	LP	P
+	+	19/0.08	PVS1, PM2	LP	PVS1, PS5, PM2, PM7	P	P
+	+	1/0.38	PVS1, PM2	LP	PVS1, PM2, PM7	P	P
+	+(parent−)	7/0.31 16/0.07	PVS1, PM2	LP	PVS1, PS5, PM2, PM7	P	P
-	n.a.	n.a.	PM2, PM4	VUS	PM2, PM4	VUS	LB
+	n.a.	n.a.	PM2, PM4, PP4	VUS	PM2, PM4, PP4	VUS	VUS
+	n.a.	n.a.	PP4	VUS	PP4	VUS	VUS
+	n.a.	n.a.	PVS1, PM2, PP4	P	PVS1, PM2, PP4	P	P
+	+(parent+)	1/0.78 1/0.34	PVS1, PM2, PP4	P	PVS1, PM2, PM7, PP1, PP4	P	P

Facial Analyses

Computer Vision Algorithms

Compared to matched controls, the photographs of patients with a confirmed pathogenic *ARID1B* variant (n = 34) did not cluster (p = 0.34). Selecting only *ARID1B* patients aged 10 and below (n = 21, Figure S1) showed that the photographs of this group did cluster (p = 0.038), while photographs of patients aged 10–18 and 18+ years did not cluster (respectively, n = 9, p = 0.73, and n = 4, p = 0.88).

The photographs of Index 1, 4, 5, 6, 7 and 12 and the parent of Index 12 clustered with photographs of the young *ARID1B* patients from our outpatient clinic compared to the controls, while the photographs of the parent of Index 2, 4 and 7 did not cluster (Table 1, Table S1 and Figure S1). The photographs of Case 2 at an age of two years and 11 years clustered with the young *ARID1B* patients, but her most recent photograph taken at an age of 11.5 years did not cluster.

Face2Gene

Face2Gene suggested CSS as a possible syndrome for all pictures of children and of the parents of whom childhood pictures were available for analyses. CSS was the first suggested syndrome for the Cases 4, 6 and 7, and for childhood pictures of Parents 7 and 12 (Table 1 and Figure S2a,b). CSS was not suggested for the adult pictures of parents of Cases 2, 4 and 7. For Cases 2, 7 and 12, multiple facial photographs taken at different ages were available. The concordance of these photographs to the CSS Face2Gene model appears to decline with age (Figure S2c,d).

Variant Interpretation

Variants at the 5' end of Exon 1 of ARID1B

Case 1

This 20-year-old boy has a start codon variant, which was not found in his mother. It should be noted that on the MANE transcript NM_001374828.1, this variant is a missense variants (p.(Met84Ile)). He has moderate ID, normal speech development, seizures since the age of three years, cannot read or write and has small nails. According to the ACMG guidelines, this variant is classified as likely pathogenic (Table 1). The location of this variant and the absence of other pathogenic variants in the region make interpretation complex. His photograph clusters with other *ARID1B* patients, but the methylation pattern in his blood was not compatible with a BAFopathy. Given the lack of variants in the 5' end of exon 1, our expert opinion is that the variant should be classified as a VUS.

Case 2

This six-year-old girl carries a paternally inherited nonsense variant in the 5' end of exon 1 (c.361C>T, p.(Gln121*)). Her primary symptoms were dysmorphic features, hyperactivity and disruptive/defiant behaviour. None of these features were seen in the father and both father and daughter have normal intelligence. Although according to ACMG guidelines (Table 1) this variant would classify as likely pathogenic, taking our experience with the inheritance, location and phenotype into account, we would consider this variant a VUS. The photograph of the child at age two years clustered with other *ARID1B* patients; one of her photographs taken at 11 years also clustered, while another photo at this age did not cluster. The photograph of the parent at age 47 years did not cluster. No DNA samples were available for further testing. After adding the photograph data to our assessment, we still consider this variant a VUS.

Case 3

This is a singleton case aged >50 years from the AURORA study (conducted in the early 2000s in patients with end-stage kidney failure). A frameshift variant in the 5' end of exon 1 was identified (c.363_364insG, p.(Gln122fs*110)). This variant has been reported several times in gnomAD. No additional information about this patient's phenotype is available and inheritance was not tested. According to the ACMG guidelines (Table 1), this variant classifies as a VUS. However, based on the frequency in gnomAD and the reason why sequencing was performed, we consider this variant as likely benign. No photographs or DNA were available for further testing.

Case 4

This is a 17-month-old child who inherited a frameshift variant located at the 5' end of exon 1 (c.521dup, p.(Pro177fs*55)) from a parent. Although based on ACMG guidelines this variant would be classified as likely pathogenic (Table 1), we initially considered this variant a VUS, based on the inheritance and the location at the 5'end of exon 1. The photograph of the child clustered with other *ARID1B* patients, while the photograph of the parent at age 36 years did not cluster. A picture of the parent during childhood was not available. DNA methylation in both the child and the parent was compatible with a BAFopathy. The phenotype of the child and the parent fit within the *ARID1B* spectrum and this variant is considered pathogenic.

Case 5

This is a 10-year-old boy who inherited a frameshift variant in exon 1 (c.1029_1056del, p.(Ala349Metfs*11)) from his mother. He has mild ID, speech delay, behavioural issues and attends special education school. His mother also attended special education school (as a child and adult) and has no behaviour issues. The maternal grandmother

was also reported to have ID. The boy lives with his father. According to the ACMG guidelines this variant could be classified as likely pathogenic (Table 1). The photograph of this patient clustered with *ARID1B* patients, his DNA methylation pattern was compatible with a BAFopathy and this variant was considered pathogenic.

Case 6

This is a four-year-old girl with a maternally inherited frameshift variant in exon 1 (c.1044_1071del, p.(Ala349Metfs*11)). Further segregation analyses showed that the variant was present in another affected sibling and absent in an unaffected sibling. She has global developmental delay, severe speech delay, autistic features and feeding difficulties. Her mother has ID and her affected brother has developmental delay, mild ID and short stature. ACMG guidelines support a variant classification of likely pathogenic (Table 1). The photograph of the child clustered with *ARID1B* patients and the variant is considered pathogenic.

Case 7

This is a nine-year-old girl who inherited a frameshift variant in exon 1 (c.1044_1062del, p.(Gly351Alafs*12)) from her father. She had neonatal hypotonia with transient feeding difficulties, developmental delay (walked independently at 30 months, speech delay), mild ID (total IQ 73 at 7.5 years, verbal comprehension 84, perceptual reasoning 75, working memory 82, processing speed 76), behavioural problems (i.e. tantrums, ADHD), epilepsy and progressive obesity. The father had mild learning difficulties but no intellectual deficiency (total IQ 98 but heterogeneous profile, verbal comprehension 84, perceptual reasoning 106, working memory 112, processing speed 94). At physical examination, he had coarse facial features. He had epilepsy between 2 and 12 years of age. Based on the ACMG guidelines, this variant could be classified as likely pathogenic (Table 1). The photograph of the child clustered with *ARID1B* patients, while the photograph of her father at age 14 years did not cluster. The DNA methylation pattern in the parent and child was compatible with a BAFopathy. This variant is therefore classified as pathogenic.

Splice Sites of in-Frame Exons or in-Frame Deletions in ARID1B

Case 8

This is a seven-year-old boy who inherited an in-frame deletion of exon 3 and 4 from his father. He has a clinical diagnosis of familial Mediterranean fever. A microarray was performed in search of an intragenic MEFV deletion because only one pathogenic variant in the FMF gene was identified. He has normal development and no dysmorphic features. ACMG guidelines support a classification as a VUS (Table 1). No photograph for

facial analysis was available and DNA methylation did not show a BAFopathy pattern. This variant is considered likely benign.

Case 9

This is an 18-year-old girl with a splice site variant near the in-frame exon 7 (c.2371+2T>C) with unknown inheritance. This variant is located close to a variant (c.2371+5G>A, de novo) reported by Lord et al. [15]. She has severe ID, currently no speech, lost 3–4 words between age 1–2 years, seizures, myopia, hypoplasia of the corpus callosum, intracerebral lipoma of the quadrigeminal plate, self-injurious behaviour, mild scoliosis and camptodactyly of the 3rd and 5th finger. She is currently a senior in a high school program for individuals with developmental disabilities. Suggestive features for an *ARID1B*-related disorder in this case are hypoplasia of the corpus callosum, myopia, seizures and scoliosis. In accordance with the ACMG guidelines, this variant was considered a VUS (Table 1). No photographs or DNA were available for further testing.

Case 10

This is an 11-year-old girl who inherited a splice site variant near the in-frame exon 8 (c.2372-2A>C) from her mother. This variant has been reported in two siblings with autism spectrum disorder without ID who inherited this variant from their father. This variant was absent in their sibling without autism [16]. This variant was also reported in a stillborn case with features consistent with prenatal CSS with unknown inheritance [17].

She has severe global developmental delay, complex partial epilepsy, microcephaly, generalized hypotonia, cortical visual impairment, laryngomalacia, oromotor dysphagia, constipation, feeding difficulties and hip dislocation. No information about her mother is available. Although the combination of laryngomalacia with developmental delay, visual impairment and seizures is suggestive of *ARID1B*-related ID. The fact that the variant is inherited both in the literature case and in this family argues against pathogenicity. Unfortunately, no information on the mother, and no photographs or DNA could be obtained, and this variant remains a VUS, conforming to the ACMG guidelines (Table 1).

Mosaicism in Parents

Case 11

The proband is an eight-year-old boy who has a sister with a nonsense variant in exon 18 (c.4870C>T, p.(Arg1624*)) that has been reported as pathogenic several times in the literature [7,18,19,20]. The mother tested negative for the variant and the father was potentially mosaic, with a very low signal on Sanger sequencing (data not shown). The proband has delayed motor milestones, delayed speech, moderate ID, autism spectrum disorder, hypotonia, dysmorphia, exotropia, short stature, hypoplastic left

heart syndrome with coarctation of the aorta, undescended testes and inguinal hernias, feeding difficulty (G-tube dependent) and a history of strabismus. Brain MRI showed mild cerebellar vermis hypoplasia. His 10-year-old sister (diagnosed at 4.5 years) has developmental delay, moderate ID, no speech, behavioural anomalies, normal growth, no cardiac defect, feeding difficulties in infancy and, besides optic nerve hypoplasia, no brain anomalies. In accordance with the ACMG guidelines and based on the phenotype and previously reported loss-of-function variants located in this exon, this variant was considered pathogenic in the initial assessment (Table 1). No photographs could be obtained.

Case 12

This is a three-year-old boy (now eight years) who inherited a nonsense variant in exon 20 (c.6322C>T, p.(Gln2108*)) from his mosaic father (data not shown). Several patients with loss-of-function variants near this variant have been reported [3,4,5,7,19,20]. This boy's phenotype matches the *ARID1B* phenotype well. His father is considered to have normal intellectual functioning, went to regular primary school and was able to keep up with peers, does not have developmental delay, does have a range of behavioural issues, mild scoliosis and currently lives independently with his partner and children. In accordance with the ACMG guidelines (Table 1), this variant was considered pathogenic. The photographs of the child at the age of eight years and the parent at the age of 18 years clustered with *ARID1B* patients, confirming the pathogenicity of this variant.

DISCUSSION

Variable Expression

Variable expression of the *ARID1B* phenotype is evident in the inherited cases we have described (Table 1). For example, the parents of Cases 4–7 appear to be considerately less severely affected than their offspring. Although sporadic cases of inherited variants from similarly affected parents [21,22,23] or variants inherited through (gonadal) mosaicism [7,24,25,26,27] have been reported (Table S2), here we describe for the first time that a pathogenic, non-mosaic variant is inherited from a very mildly affected parent with normal IQ values (i.e., the father in Case 7). Unfortunately, the grandparents could not be tested in our cases with inherited variants. Parents in our cases were generally less severely affected than their child. This is likely explained by selection bias since mildly affected patients are more likely to have offspring than severely affected patients. Another clear illustration of variable expression is Case 11, where two siblings with the same variant share a similar cognitive phenotype (moderate ID), whilst only one of them has congenital heart disease. Due to this variable phenotype, it can be

challenging to distinguish truly benign variants from familial variants with high variable expression.

We show that DNA methylation studies and algorithms that model human facial dysmorphism and facial recognition can be helpful tools to determine the pathogenicity of such complex variants in several illustrative cases. Based on our experience, we propose an expansion of the evidence framework of the ACMG guidelines [6] (Box 1) to improve variant interpretation, which we discuss in more detail below.

Box 1. Suggested updates of the ACMG guidelines for the interpretation of sequence variants.

PVS1.

Adjusting the 2nd caveat of PVS1 into:

- Use caution interpreting LOF variants at the extreme 5' and 3' ends of a gene.

Adding **PS5**

A DNA methylation signature matching the gene in which the variant was identified is present.

Caveat:

- If variants in related genes lead to the same signature, care must be taken to exclude the relevant variants in those genes.

Adding **PM7**

Facial photograph clusters with photographs of patients with a confirmed pathogenic variant and separate from the matched controls.

Variant Location and the PVS1 Category

The PVS1 criterion in the ACMG framework is triggered when a predicted loss-of-function variant is detected in a gene known to have haploinsufficiency as a mechanism. One of the caveats to the PVS1 criterion is that 3' variants may escape nonsense-mediated decay (NMD) [6,28]. Another problem is the possibility that 5' loss-of-function variants may be rescued by an alternative translation start site. For *ARID1B*, we have previously noted that there are no convincing pathogenic variants reported at the 5'

end of exon 1 [7]. This could be a chance finding, could reflect that such variants are not pathogenic or perhaps that such variants are lethal and therefore not identified in live born patients. Since no BAFopathy DNA methylation signature was shown in Case 1, and the phenotypes of Cases 2 and 3 are not consistent with *ARID1B*-related ID, such a start site may be located between cDNA location 363 and 521. It may be that such a start site would rescue the complete phenotype, but an intriguing alternative is that it rescues part of the phenotype only. This could explain why we find a relative abundance of inherited variants in exon 1. There is one methionine in between (p.Met154), but other start codons have also been described to initiate translation [29]. We therefore propose to adjust the 2nd caveat of PVS1 to also include possibilities of alternative translation start sites: "Use caution interpreting LOF variants at the extreme 5' and 3' ends of a gene".

The absence of pathogenic variants in exon 3 of *ARID1B* reinforces the importance of another caveat for PVS1: to use caution when the genes is known to generate multiple transcripts [6,30]. The in-frame exon 3 in transcript NM_020732.3 is not included in another *ARID1B* transcript NM_017519.2. This suggests that exon 3 is alternatively spliced and may not be required in all tissues, which is confirmed by several reported cases with loss-of-function variants in exon 3 without a matching phenotype [31,32].

Our Case 8, where in-frame exons 3 and 4 were deleted in an unaffected patient and parent, illustrates the importance of this caveat, but also of the caveat relating to skipping of in-frame exons. Interestingly, Pascolini et al. [33] describe monozygotic twins with an in-frame deletion of exon 2–4 of *ARID1B*, where the father could not be tested. The location, in-frame nature and unknown inheritance cast doubt upon the pathogenicity of this particular variant [34], although a very similar variant was identified in a CSS patient by Fujita et al. [35]. They identified an in-frame deletion of exons 2–5, of unknown inheritance, but the phenotype and the Face2Gene analysis (rank 1, similarity 0.62) make the pathogenicity of this variant more likely. Unfortunately, no DNA of this reported case could be obtained to determine the DNA methylation signature.

The described cases illustrate how important it is to keep in mind the caveats mentioned in the ACMG guidelines. For example, if these were to be disregarded, the variants of Cases 1, 2, 8 and 9 would have been classified as likely pathogenic or pathogenic instead of as a VUS. This also stresses the importance of having gene-specific knowledge concerning genotype and phenotype when interpreting variants.

Transcript Choice

We have considered switching to the MANE transcript for this paper. However, to prevent unnecessary confusion in the literature, we decided to only switch to a new transcript when it is fully backed up by clinical evidence (such as variants being identified in 'new' exonic regions (i.e. the additional 5'part of exon 1, and the in-frame exon 11). Thus far, this is not the case, although care should be taken that these 'new' regions are properly sequenced and called. Although they appear to be captured in most recent exome capture kits, it is possible that bio-informatic pipelines do not yet annotate these new regions causing clinical variants to be missed. We recommend that clinical labs check their pipelines for compatibility with the MANE transcript. We have also queried GnomAD for truncating variants in the new exonic regions, which would be an argument against the new transcript. We found five truncating variants in GnomAD (accession date 6 August 2021), but each variant was found once, and four of them are in low complexity regions, so this does not provide definitive evidence either way. Thus, even though NM_020732.3 is not the perfect transcript for *ARID1B* as it incorporates the in-frame exon 3, which is most likely not clinically relevant, we have chosen to keep using NM_020732.3 until there is more evidence for the new transcript. We have added the description of all variants detected in our cases on the MANE transcript in the Supplementary Data (Table S1).

One important consequence of the new transcript is that the start codon variant of patient one changes into a missense variant (c.252G > A, p.(Met84Ile)). This might explain the absence of a DNA methylation signature in this patient, since for some genes specific signatures for missense variants are identified. The consequence of this variant is on the protein level is not yet known.

Incorporating New Developments into ACMG Guidelines for the Interpretation of Sequence Variants

Since the introduction of the ACMG guidelines there have been several relevant developments. Facial analysis has become more widely used and DNA methylation analysis has been shown to be an effective functional assay in clinical classification of an expanding number of genetic disorders [11,12], including CSS [10].

Facial Analyses

By means of computer vision algorithms we have shown that facial photographs of *ARID1B* patients aged <10 years cluster separate from age-, sex-, ethnicity-matched controls with ID and that with advancing age this clustering becomes less pronounced (Figure S1). This apparent absence of a clear facial phenotype among older *ARID1B* patients could be explained by the notion that with advancing age the facial phenotype

becomes less specific, but the low number of available photographs of older *ARID1B* patients may have played a role as well. We illustrated that facial analysis using childhood photographs can be helpful in the interpretation of difficult variants and we suggest that a positive facial gestalt match could be incorporated as a moderate ACMG criteria (PM7): "Facial photograph clusters with photographs of patients with confirmed pathogenic variants and separate from matched controls". We think this better reflects the objective nature of facial analyses then allocating the PP4 criterion of a phenotype highly specific for gene. Of course, this criterion should be limited to the syndromes that display facial dysmorphism.

DNA Methylation

Multiple studies have shown the potential of DNA methylation signatures for variant interpretation [11,12]. However, caution is warranted because the sensitivity of DNA methylation signatures may be not be complete for all pathogenic variants in a given gene. For example, multiple episignatures have been described in single genes in some disorders [36,37]. Similarly, the DNA methylation result of Case 1 is not compatible with a BAFopathy, whilst his photograph clusters with *ARID1B* patients and his phenotype may fit within the *ARID1B* spectrum. Hence, while existence of an episignature is generally considered strong functional evidence to confirm a pathogenicity of a variant, lack of an episignature, albeit indicative, should not be considered conclusive evidence of the lack of pathogenicity [12]. The DNA methylation results all mirror the assessment of pathogenicity before testing in this study, except for Case 4. In this case, a loss-of-function variant was found at the 5' end of exon 1, at cDNA position 521. This variant was reclassified after the DNA methylation showed a BAFopathy methylation pattern in both the child and parent. These data, along with the growing body of published literature on other disorders, support incorporating DNA methylation as a strong criterion for classifying pathogenic variants (PS5): "A DNA methylation signature matching the gene in which the variant was identified is present". In our view this should be a separate criterion from PS3 (well-established functional studies show a deleterious effect). This is also in line with ACGS Best Practice Guidelines for Variant Classification 2019 [38], since the evidence of a methylation signature supports phenotype specificity and is not directly related to the identified variant. Perhaps at a later stage when more is certain about the sensitivity of this DNA methylation test, a criterion could also be added to classify variants as benign in the absence of a DNA methylation signature.

Incorporating Suggested ACMG Updates

Taking the suggested new PM7 and PS5 criteria and the adjusted PVS1 caveat into account, we have classified the variants and show that the pathogenicity assessment of the variants of our included cases would significantly change (Table 1) and have become more conforming to our interpretation of the *ARID1B* variants.

The UK Association for Clinical Genomic Science (UK-ACGS) publishes an annual ACMG-based specification for variant interpretation [39]. Using these criteria, the initial interpretation of the *ARID1B* variants would change for Cases 1, 2 and 4–7 into likely pathogenic and the interpretation using our updated ACMG criteria would remain unchanged.

CONCLUSIONS

Due to the variable phenotype among *ARID1B* patients, it can be challenging to distinguish familial variants with variable expression from benign variants. We urge caution in the interpretation of variants in the 5' end of exon 1 and in-frame deletions and suggest to update the ACMG criteria PVS1 to also include caution with variants at the 5' end of the transcript. We demonstrate the diagnostic utility of DNA methylation signatures in such cases and show the variable expression of the *ARID1B* phenotype, by reporting the first affected parent with normal IQ. Furthermore, we confirmed that *ARID1B* patients aged < 10 years have a distinct facial phenotype and found that the facial phenotype of *ARID1B* patients may become less specific with age. For this reason, if facial analyses are used to aid interpretation, childhood pictures should be preferred. Finally, we suggest the addition of facial analysis and/or DNA methylation signature analysis to the ACMG guidelines for interpretation of sequence variants in determining the impact of a genetic change on a patient's diagnosis.

SUPPLEMENTARY MATERIALS

The following are available online at https://www.mdpi.com/article/10.3390/genes12081275/s1, Figure S1: (A) t-SNE plot depicting clustering of photographs of familial ARID1B cases (1 photograph per case was included) compared to confirmed ARID1B patients aged < 10 years photographs and age-, sex-, and ethnicity matched controls with intellectual disability. (B) t-SNE plot depicting clustering of photographs of familial ARID1B cases compared to confirmed ARID1B patients' photographs and matched controls, Figure S2: depicting similarity and ranking of facial photographs of

the familial *ARID1B* cases* in F2G. Table S1: overview of the clustering of photographs of *ARID1B* inherited cases in comparison to *ARID1B* outpatient clinic cases, Table S2: previously reported inherited loss-of-function *ARID1B* variants.

Supplementary Information (665.3KB, zip)

FUNDING

This research received no external funding.

DATA AVAILABILITY STATEMENT

Not applicable.

CONFLICTS OF INTEREST

Jill Rosenfeld: The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing completed at Baylor Genetics Laboratories. All other authors declare no conflict of interest.

REFERENCES

1. Hoyer J., Ekici A.B., Endele S., Popp B., Zweier C., Wiesener A., Wohlleber E., Dufke A., Rossier E., Petsch C., et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am. J. Hum. Genet. 2012;90:565–572. doi: 10.1016/j.ajhg.2012.02.007.
2. Wright C.F., Fitzgerald T.W., Jones W.D., Clayton S., McRae J.F., van Kogelenberg M., King D.A., Ambridge K., Barrett D.M., Bayzetinova T., et al. Genetic diagnosis of developmental disorders in the DDD study: A scalable analysis of genome-wide research data. Lancet. 2015;385:1305–1314. doi: 10.1016/S0140-6736(14)61705-0.
3. Santen G.W.E., Aten E., Vulto-van Silfhout A.T., Pottinger C., van Bon B.W., van Minderhout I.J., Snowdowne R., van der Lans C.A., Boogaard M., Linssen M.M., et al. Coffin-Siris syndrome and the BAF complex: Genotype-phenotype study in 63 patients. Hum. Mutat. 2013;34:1519–1528. doi: 10.1002/humu.22394.
4. Wieczorek D., Bogershausen N., Beleggia F., Steiner-Haldenstatt S., Pohl E., Li Y., Milz E., Martin M., Thiele H., Altmuller J., et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum. Mol. Genet. 2013;22:5121–5135. doi: 10.1093/hmg/ddt366.
5. Tsurusaki Y., Okamoto N., Ohashi H., Mizuno S., Matsumoto N., Makita Y., Fukuda M., Isidor B., Perrier J., Aggarwal S., et al. Coffin-Siris syndrome is a SWI/SNF complex disorder. Clin. Genet. 2014;85:548–554. doi: 10.1111/cge.12225.
6. Richards S., Aziz N., Bale S., Bick D., Das S., Gastier-Foster J., Grody W.W., Hegde M., Lyon E., Spector E., et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. 2015;17:405–424. doi: 10.1038/gim.2015.30.
7. Van der Sluijs P.J., Jansen S., Vergano S.A., Adachi-Fukuda M., Alanay Y., AlKindy A., Baban A., Bayat A., Beck-Wödl S., Berry K., et al. The ARID1B spectrum in 143 patients: From nonsyndromic intellectual disability to Coffin–Siris syndrome. Genet. Med. 2018 doi: 10.1038/s41436-018-0330-z.
8. Santen G.W., Clayton-Smith J., ARID1B-CSS Consortium The ARID1B phenotype: What we have learned so far. Am. J. Med. Genet. Part C Semin. Med. Genet. 2014;166C:276–289. doi: 10.1002/ajmg.c.31414.
9. Van der Donk R., Jansen S., Schuurs-Hoeijmakers J.H.M., Koolen D.A., Goltstein L.C.M.J., Hoischen A., Brunner H.G., Kemmeren P., Nellåker C., Vissers L.E.L.M., et al. Next-generation phenotyping using computer vision algorithms in rare genomic neurodevelopmental disorders. Genet. Med. 2019;21:1719–1725. doi: 10.1038/s41436-018-0404-y.
10. Aref-Eshghi E., Bend E.G., Hood R.L., Schenkel L.C., Carere D.A., Chakrabarti R., Nagamani S.C.S., Cheung S.W., Campeau P.M., Prasad C., et al. BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin-Siris and Nicolaides-Baraitser syndromes. Nat. Commun. 2018;9:4885. doi: 10.1038/s41467-018-07193-y.
11. Aref-Eshghi E., Kerkhof J., Pedro V.P., Groupe D.I.F., Barat-Houari M., Ruiz-Pallares N., Andrau J.C., Lacombe D., Van-Gils J., Fergelot P., et al. Evaluation of DNA Methylation Episignatures for Diagnosis and Phenotype Correlations in 42 Mendelian Neurodevelopmental Disorders. Am. J. Hum. Genet. 2020;106:356–370. doi: 10.1016/j.ajhg.2020.01.019.

12. Sadikovic B., Levy M.A., Kerkhof J., Aref-Eshghi E., Schenkel L., Stuart A., McConkey H., Henneman P., Venema A., Schwartz C.E., et al. Clinical epigenomics: Genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. Genet. Med. 2021;23:1065–1074. doi: 10.1038/s41436-020-01096-4.
13. Levy M.A., Kerkhof J., Belmonte F.R., Kaufman B.A., Bhai P., Brady L., Bursztyn L., Tarnopolsky M., Rupar T., Sadikovic B. Validation and clinical performance of a combined nuclear-mitochondrial next-generation sequencing and copy number variant analysis panel in a Canadian population. Am. J. Med. Genet. Part A. 2021;185:486–499. doi: 10.1002/ajmg.a.61998.
14. Dingemans A.J.M., Stremmelaar D.E., van der Donk R., Vissers L., Koolen D.A., Rump P., Hehir-Kwa J.Y., de Vries B.B.A. Quantitative facial phenotyping for Koolen-de Vries and 22q11.2 deletion syndrome. Eur. J. Hum. Genet. 2021 doi: 10.1038/s41431-021-00824-x.
15. Lord J., Gallone G., Short P.J., McRae J.F., Ironfield H., Wynn E.H., Gerety S.S., He L., Kerr B., Johnson D.S., et al. Pathogenicity and selective constraint on variation near splice sites. Genome Res. 2019;29:159–170. doi: 10.1101/gr.238444.118.
16. Ruzzo E.K., Perez-Cano L., Jung J.Y., Wang L.K., Kashef-Haghighi D., Hartl C., Singh C., Xu J., Hoekstra J.N., Leventhal O., et al. Inherited and De novo Genetic Risk for Autism Impacts Shared Networks. Cell. 2019;178:850–866.e826. doi: 10.1016/j.cell.2019.07.015.
17. Stanley K.E., Giordano J., Thorsten V., Buchovecky C., Thomas A., Ganapathi M., Liao J., Dharmadhikari A.V., Revah-Politi A., Ernst M., et al. Causal Genetic Variants in Stillbirth. N. Engl. J. Med. 2020;383:1107–1116. doi: 10.1056/NEJMoa1908753.
18. Al-Shamsi A., Hertecant J.L., Souid A.K., Al-Jasmi F.A. Whole exome sequencing diagnosis of inborn errors of metabolism and other disorders in United Arab Emirates. Orphanet J. Rare Dis. 2016;11:94. doi: 10.1186/s13023-016-0474-3.
19. Stessman H.A., Xiong B., Coe B.P., Wang T., Hoekzema K., Fenckova M., Kvarnung M., Gerdts J., Trinh S., Cosemans N., et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat. Genet. 2017;49:515–526. doi: 10.1038/ng.3792.
20. Sekiguchi F., Tsurusaki Y., Okamoto N., Teik K.W., Mizuno S., Suzumura H., Isidor B., Ong W.P., Haniffa M., White S.M., et al. Genetic abnormalities in a large cohort of Coffin-Siris syndrome patients. J. Hum. Genet. 2019 doi: 10.1038/s10038-019-0667-4.
21. Smith J.A., Holden K.R., Friez M.J., Jones J.R., Lyons M.J. A novel familial autosomal dominant mutation in ARID1B causing neurodevelopmental delays, short stature, and dysmorphic features. Am. J. Med. Genet. Part A. 2016;170:3313–3318. doi: 10.1002/ajmg.a.37945.
22. Mignot C., Moutard M.L., Rastetter A., Boutaud L., Heide S., Billette T., Doummar D., Garel C., Afenjar A., Jacquette A., et al. ARID1B mutations are the major genetic cause of corpus callosum anomalies in patients with intellectual disability. Brain. 2016;139:e64. doi: 10.1093/brain/aww181.
23. Yu Y., Yao R., Wang L., Fan Y., Huang X., Hirschhorn J., Dauber A., Shen Y. De novo mutations in ARID1B associated with both syndromic and non-syndromic short stature. BMC Genom. 2015;16:701. doi: 10.1186/s12864-015-1898-1.
24. Ben-Salem S., Sobreira N., Akawi N.A., Al-Shamsi A.M., John A., Pramathan T., Valle D., Ali B.R., Al-Gazali L. Gonadal mosaicism in ARID1B gene causes intellectual disability and dysmorphic features in three siblings. Am. J. Med. Genet. Part A. 2016;170A:156–161. doi: 10.1002/ajmg.a.37405.
25. Cherot E., Keren B., Dubourg C., Carre W., Fradin M., Lavillaureix A., Afenjar A., Burglen L., Whalen S., Charles P., et al. Using medical exome sequencing to identify the causes of neurodevelopmental disorders: Experience of 2 clinical units and 216 patients. Clin. Genet. 2018;93:567–576. doi: 10.1111/cge.13102.

26. Cheng S.S.W., Luk H.M., Mok M.T., Leung S.S., Lo I.F.M. Genotype and phenotype in 18 Chinese patients with Coffin-Siris syndrome. Am. J. Med. Genet. Part A. 2021 doi: 10.1002/ajmg.a.62187.
27. Min Z., Qian C., Ying D. Novel ARID1B variant inherited from somatogonadal mosaic mother in siblings with Coffin-Siris syndrome 1. Exp. Ther. Med. 2021;21:614. doi: 10.3892/etm.2021.10046.
28. Abou Tayoun A.N., Pesaran T., DiStefano M.T., Oza A., Rehm H.L., Biesecker L.G., Harrison S.M., ClinGen Sequence Variant Interpretation Working G. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum. Mutat. 2018;39:1517–1524. doi: 10.1002/humu.23626.
29. Kearse M.G., Wilusz J.E. Non-AUG translation: A new start for protein synthesis in eukaryotes. Genes Dev. 2017;31:1717–1731. doi: 10.1101/gad.305250.117.
30. Cummings B.B., Karczewski K.J., Kosmicki J.A., Seaby E.G., Watts N.A., Singer-Berk M., Mudge J.M., Karjalainen J., Satterstrom F.K., O'Donnell-Luria A.H., et al. Transcript expression-aware annotation improves rare variant interpretation. Nature. 2020;581:452–458. doi: 10.1038/s41586-020-2329-2.
31. Johnston J.J., Lewis K.L., Ng D., Singh L.N., Wynter J., Brewer C., Brooks B.P., Brownell I., Candotti F., Gonsalves S.G., et al. Individualized iterative phenotyping for genome-wide analysis of loss-of-function mutations. Am. J. Hum. Genet. 2015;96:913–925. doi: 10.1016/j.ajhg.2015.04.013.
32. Yuen R.K.C., Merico D., Bookman M., Howe J.L., Thiruvahindrapuram B., Patel R.V., Whitney J., Deflaux N., Bingham J., Wang Z., et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat. Neurosci. 2017;20:602–611. doi: 10.1038/nn.4524.
33. Pascolini G., Valiante M., Bottillo I., Laino L., Fleischer N., Ferraris A., Grammatico P. Striking phenotypic overlap between Nicolaides-Baraitser and Coffin-Siris syndromes in monozygotic twins with ARID1B intragenic deletion. Eur. J. Med. Genet. 2019:103739. doi: 10.1016/j.ejmg.2019.103739.
34. Van der Sluijs P.J., Santen G.W.E. Letter regarding the article: Striking phenotypic overlap between Nicolaides-Baraitser and Coffin-Siris syndromes in monozygotic twins with ARID1B intragenic deletion. Eur. J. Med. Genet. 2020;63:103813. doi: 10.1016/j.ejmg.2019.103813.
35. Fujita T., Ihara Y., Hayashi H., Ishii A., Ideguchi H., Inoue T., Imaizumi T., Yamamoto T., Hirose S. Coffin-Siris syndrome with bilateral macular dysplasia caused by a novel exonic deletion in ARID1B. Congenit. Anom. 2020;60:189–193. doi: 10.1111/cga.12383.
36. Cappuccio G., Sayou C., Tanno P.L., Tisserant E., Bruel A.L., Kennani S.E., Sa J., Low K.J., Dias C., Havlovicova M., et al. De novo SMARCA2 variants clustered outside the helicase domain cause a new recognizable syndrome with intellectual disability and blepharophimosis distinct from Nicolaides-Baraitser syndrome. Genet. Med. 2020;22:1838–1850. doi: 10.1038/s41436-020-0898-y.
37. Bend E.G., Aref-Eshghi E., Everman D.B., Rogers R.C., Cathey S.S., Prijoles E.J., Lyons M.J., Davis H., Clarkson K., Gripp K.W., et al. Gene domain-specific DNA methylation episignatures highlight distinct molecular entities of ADNP syndrome. Clin. Epigenetics. 2019;11:64. doi: 10.1186/s13148-019-0658-5.
38. Ellard S., Baple E.L., Berry I., Forrester N., Turnbull C., Owens M., Eccles D.M., Abbs S., Scott R., Deans Z.C., et al. ACGS Best Practice Guidelines for Variant Classification 2019: Association for Clinical Genetics Science (ACGS) [(accessed on 25 May 2021)]; Available online: https://www.acgs.uk.com/news/acgs-best-practice-guidelines-for-variant-classification-2019/
39. Ellard S., Baple E.L., Berry I., Forrester N., Turnbull C., Owens M., Eccles D.M., Abbs S., Scott R., Deans Z.C., et al. ACGS Best Practice Guidelines for Variant Classification 2020: Association for Clinical Genetics Science (ACGS) [(accessed on 25 May 2021)]; Available online: https://www.acgs.uk.com/quality/best-practice-guidelines/#VariantGuidelines.

Chapter 5

ARID1B-related disorder in 87 adults: Natural history and self-sustainability

Pleuntje J van der Sluijs, Marjolein Gösgens, Alexander J M Dingemans, Pasquale Striano, Antonella Riva, Cyril Mignot, Anne Faudet, Georgia Vasileiou, Mona Walther, Samantha A Schrier Vergano, Marielle Alders, Fowzan S Alkuraya, Ibrahim Alorainy, Hessa S Alsaif, Britt-Marie Anderlid, Iben Bache, Irma van Beek, Maud Blanluet, Bregje W van Bon, Theresa Brunet, Han Brunner, Miriam L Carriero, Perrine Charles, Nicolas Chatron, Emanuela Coccia, Cristèle Dubourg, Rachel K Earl, Evan E Eichler, Laurence Faivre, Nicola Foulds, Claudio Graziano, Anne-Marie Guerrot, Mais Omar Hashem, Solveig Heide, Delphine Heron, Scott E Hickey, Saskia M J Hopman, Anja Kattentidt-Mouravieva, Jennifer Kerkhof, Jolien S Klein Wassink-Ruiter, Evangeline C Kurtz-Nelson, Katarína Kušíková, Malin Kvarnung, Francois Lecoquierre, Gloria S Leszinski, Lorenzo Loberti, Pilar L Magoulas, Francesca Mari, Isabelle Maystadt, Giuseppe Merla, Jeff M Milunsky, Stephanie Moortgat, Gael Nicolas, Melanie O' Leary, Sylvie Odent, Jillian R Ozmore, Kareesma Parbhoo, Rolph Pfundt, Maria Piccione, Anna Maria Pinto, Bernt Popp, Audrey Putoux, Heidi L Rehm, Andre Reis, Alessandra Renieri, Jill A Rosenfeld, Massimiliano Rossi, Emanuela Salzano, Pascale Saugier-Veber, Marco Seri, Giulia Severi, Fatma M Sonmez, Gertrud Strobl-Wildemann, Kyra E Stuurman, Eyyüp Uctepe, Hilde Van Esch, Giulia Vitetta, Bert BA de Vries, Dagmar Wahl, Tianyun Wang, Pia Zacher, Karin R Heitink, Fabienne G Ropers, Duco Steenbeek, Tamara Rybak, Gijs W E Santen

Published: Genet Med Open. 2024 Jul 23;2:101873. doi: 10.1016/j.gimo.2024.101873

ABSTRACT

Purpose: *ARID1B* is one of the most frequently mutated genes in intellectual disability cohorts. Thus, far few adult-aged patients with *ARID1B*-related disorder have been described, which limits our understanding of the disease's natural history and our ability to counsel patients and their families.

Methods: Data on patients aged 18+ years with *ARID1B*-related disorder were collected through an online questionnaire completed by clinicians and parents.

Results: Eighty-seven adult patients with *ARID1B* were included. Cognitive functioning ranged from borderline to severe intellectual disability. Patients identified through the genetic workup of their child were either mosaic or had a variant in exon 1. New clinical features identified in this population are loss of skill (16/64, 25%) and recurrent patella luxation (12/45, 32%). Self-sustainability data showed that 88% (45/51) could eat independently, and 16% (7/45) could travel alone by public transport. Facial photo analysis showed that patients' photographs taken at different ages clustered consistently, separate from matched controls.

Conclusion: The *ARID1B* spectrum is broad, and as patients age, there is a significant shift in the medical aspects requiring attention. To address the changing medical needs with increasing age, we have formulated recommendations to promote timely intervention in an attempt to mitigate disease progression.

Keywords**:** Adult, *ARID1B*, Coffin–Siris syndrome, Developmental delay, Intellectual disability

INTRODUCTION

ARID1B (HUGO Gene Nomenclature Committee: 18040) is one of the most frequently mutated genes in intellectual disability (ID) and neurodevelopmental delay (NDD) cohorts at around 1%.1, 2, 3, 4 The phenotypic spectrum of *ARID1B*-related disorder is broad, ranging from severe ID in Coffin–Siris syndrome (CSS) patients (OMIM 135900) to normal IQ scores in patients with developmental delay.[5]

Although many challenges faced by patients during childhood are well-documented (such as feeding difficulties, failure to thrive, and seizures)5, 6, 7, 8 there remains a significant knowledge gap regarding the obstacles they encounter later in adult life. This knowledge gap is common in genetic ID and NDD, as most published patients are minors.

Although there are sporadic case reports of adult patients with *ARID1B*-related disorder,7, 8, 9, 10, 11 it is unclear whether these represent a biased subset of the phenotype. Thus, studies documenting the development, functioning, and challenges faced by adult-aged patients in large cohorts are needed to improve the counseling of parents about the diagnosis, provide appropriate screening and guidance to diagnosed patients, and facilitate the transition of patients from pediatric to adult care.[12] To provide a more complete overview of adult patients with *ARID1B*-related disorder, we acquired and analyzed data from 87 adult-aged patients, investigated the natural history of this disorder, assessed functional (in)dependence, identified potential comorbidities associated with aging, and aggregated these data into screening recommendations.

MATERIALS AND METHODS

Patient collection

Patients with a heterozygous pathogenic variant in *ARID1B* and aged 18 years and above were identified through the following diverse sources: national and international colleagues approaching us with their patients; ClinVar; Leiden Open Variant Database; contacts from the Baylor Genetics Laboratories; physician referrals for second opinions; and the Facebook group named "Coffin-Siris Syndrome Group". Additionally, recruitment took place through the outpatient CSS expertise center at Leiden University Medical Center, Leiden, the Netherlands.

Data collection

Data were collected through an online questionnaire which was completed by the referring physician. If possible, parents were also asked to complete an online questionnaire (languages available were English, Dutch, French, German, Italian and Japanese). The clinical questionnaire focused on medical history and physical examination. Although several questions overlapped with those in the parental questionnaire, the parental questionnaire also included additional inquiries about activities of daily living and overall functioning. Variable labels of all questions are included in Supplemental Excel 1. When 2 patients with the same pathogenic variant were included, their phenotypes were compared. If overlapping features were observed, contributing clinicians were contacted to verify whether these were different patients. In this manner, we were able to identify 2 duplicate entries.

Data assessment

Alamut Visual Plus version 1.6.1 was used to translate the genetic variant or derive the missing genomic location if a genetic variant was reported on a different *ARID1B* transcript, if another genomic build was used, or if the genomic location was not reported. If the standard deviation score (SDS) was not reported, and raw data on weight, height, or occipital–frontal circumference (OFC) were available, the SDS was determined using published growth charts.[13,14] Furthermore, we contacted clinicians of published patients with *ARID1B*-related disorders (now aged 18+) by email to request an update. All analyses were executed using SPSS version 25. R version 4.2.1 was used to create graphs. Control data for survival were derived from the Human Mortality database as reported by Beltrán-Sánchez et al[15] and Verguet et al[16], data on developmental milestones were derived from Sheldrick et al[17] and data on toilet training from Schum et al.[18]

PhenoScore

To assess whether the individuals in this study had a facial gestalt distinguishable from other NDD patients, the facial module of PhenoScore was utilized. PhenoScore is a next-generation phenomics framework that uses artificial intelligence to assess whether the phenotype of a specific group is different from that of age-, sex- and ethnicity-matched controls with neurodevelopmental disorders.[19] In this case, we used QMagFace[20] as the facial feature extraction method in PhenoScore and then trained it to see whether the facial features of the investigated individuals were different from those of matched controls with NDD. Matched NDD controls were derived from the in-house database of the Radboud University Medical Center with over 1,200 individuals seen at their outpatient clinic.[19] The individuals included in this study comprise a sample of patients with Neurodevelopmental Disorders (NDD) seen at the Radboud University Medical

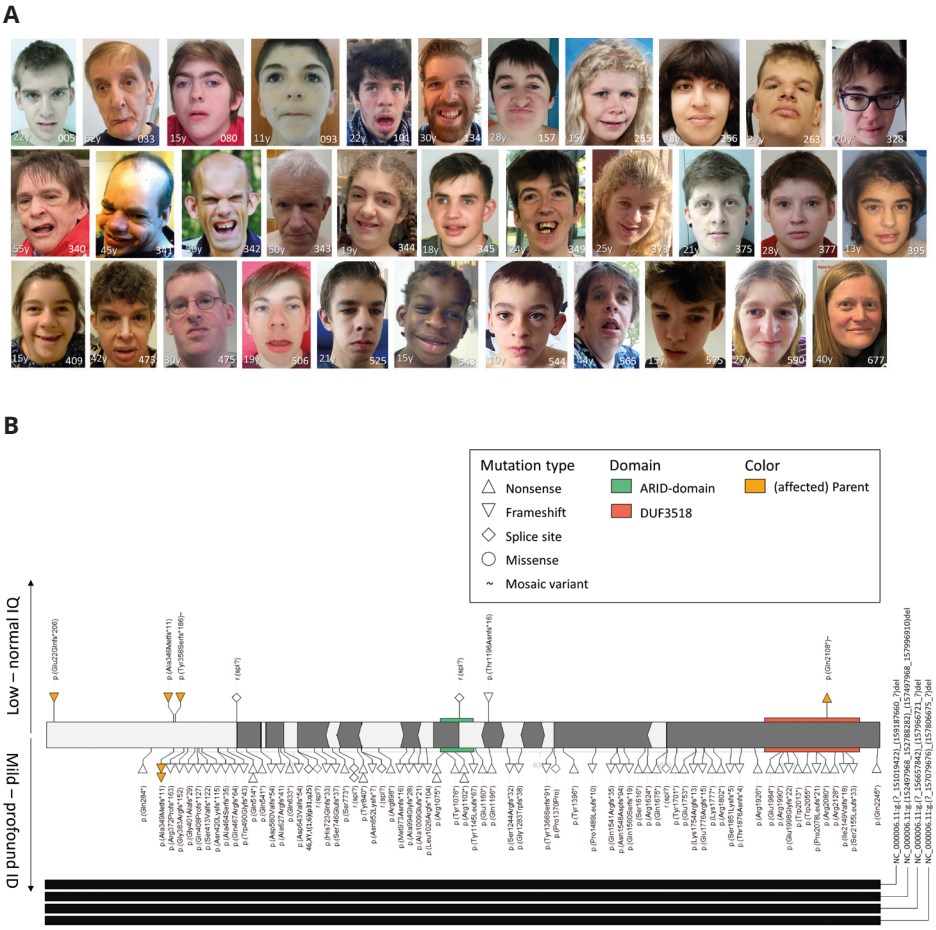
Center. This cohort encompasses both individuals with known genetic causes of NDD and those with unknown causes. Most of the individuals in this group have undergone exome sequencing to identify potential genetic factors contributing to their condition. This sample is not biased or overrepresented by specific genetic disorders; rather, it represents a random and non-selected subset of the NDD population. In other words, it reflects the diversity and heterogeneity typically observed within the NDD population. For the full methodology of PhenoScore, please see Dingemans et al.[19] These analyzes were performed for the whole study group and different subgroups based on age.

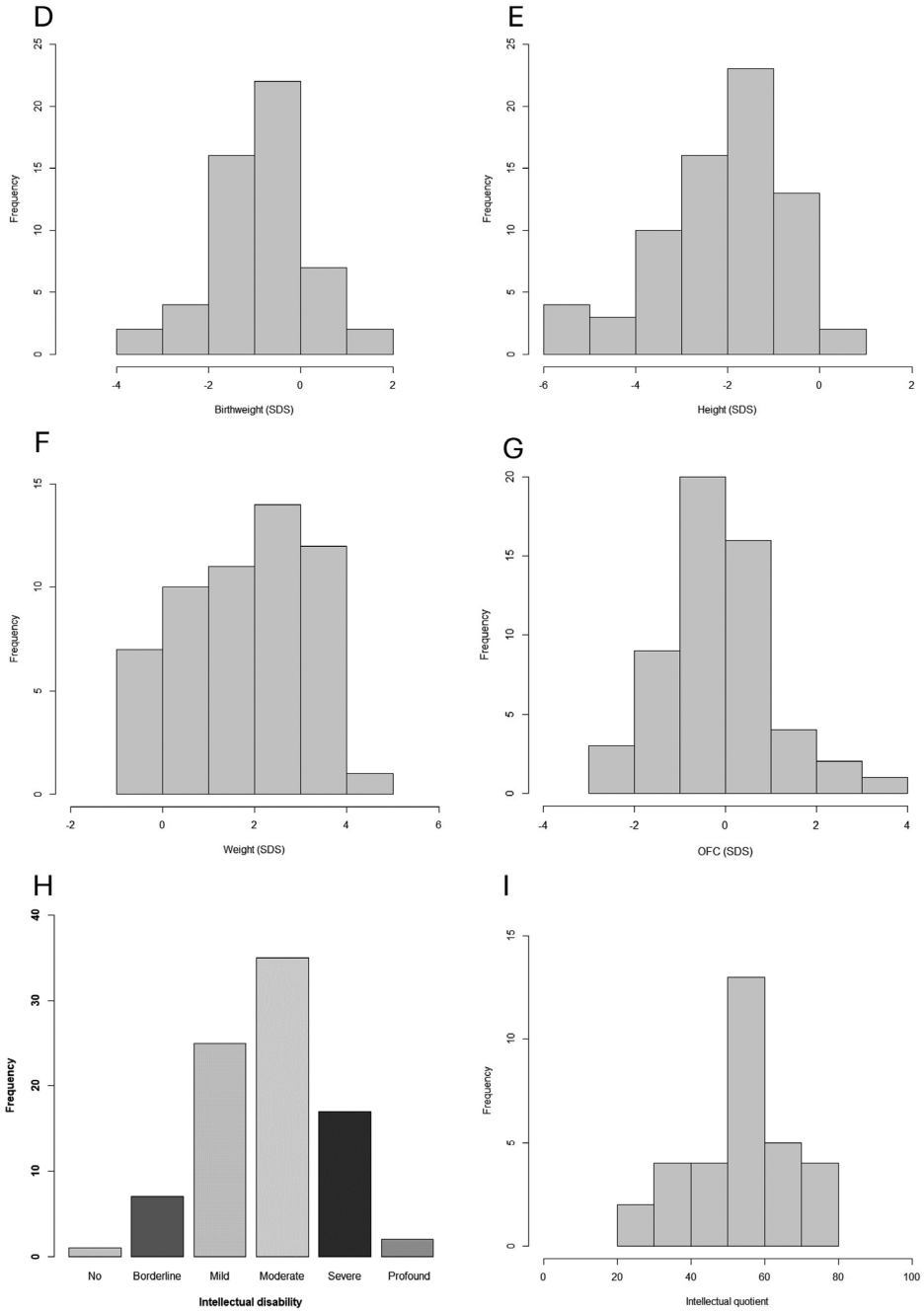
RESULTS

Eighty-seven adult-aged patients with pathogenic variants in *ARID1B* were included. For all patients, an online clinical survey was completed by the clinician. In 53 cases (61%) parents also completed an online questionnaire about their child. Forty-five patients have previously been published, and in 38 of them (84%) clinical data were updated (Supplemental Table 1). Facial photographs were available for 55 patients, and parents or caretakers provided consent for publication for 38 patients (Figure 1A and Supplemental Figure 1).

Genotype

Figure 1B and Supplemental Table 1 give an overview of the *ARID1B* variants in our cohort. In 6 cases, the variant was passed on to 1 or more children (currently < 18 years of age, therefore their data was not included in this manuscript). Two of these patients had a mosaic variant. In 62 patients the pathogenic variant was *de novo*. In the remaining 19 patients, inheritance could not be determined or only 1 parent was available for testing. One inherited variant (ie., NM_020732.3:c. c.63_73del p.(Glu22Glnfs∗206)) was identified through exome sequencing and was initially classified as a variant of uncertain significance due to its location early in the transcript. Additional DNA methylation analysis on DNA derived from blood showed a BAFopathy episignature (Supplemental Figure 2), after which this variant was reclassified as pathogenic. All patients have variants predicted to lead to haploinsufficiency. A prior suspicion of CSS was present in 22.8% (18/79) of the cases. Figure 1B shows the distribution of pathogenic variants and degree of ID. Aside from the pathogenic variants in *ARID1B,* no other pathogenic deletions, duplications, or single nucleotide variants were identified in our patient cohort.


A
B
Mutation type
Nonsense
Frameshift
Splice site
Missense
Mosaic variant
Domain
ARID-domain
DUF3518
Color
(affected) Parent
Low – normal IQ
Mild – profound ID
C
All ARID1B patients, n=87
Control population*
Survival probability
1.00
0.75
0.50
0.25
0.00
20
30
40
50
60
70
Time (years)


D
Frequency
Birthweight (SDS)
E
Frequency
Height (SDS)
F
Frequency
Weight (SDS)
G
Frequency
OFC (SDS)
H
Frequency
No
Borderline
Mild
Moderate
Severe
Profound
Intellectual disability
I
Frequency
Intellectual quotient

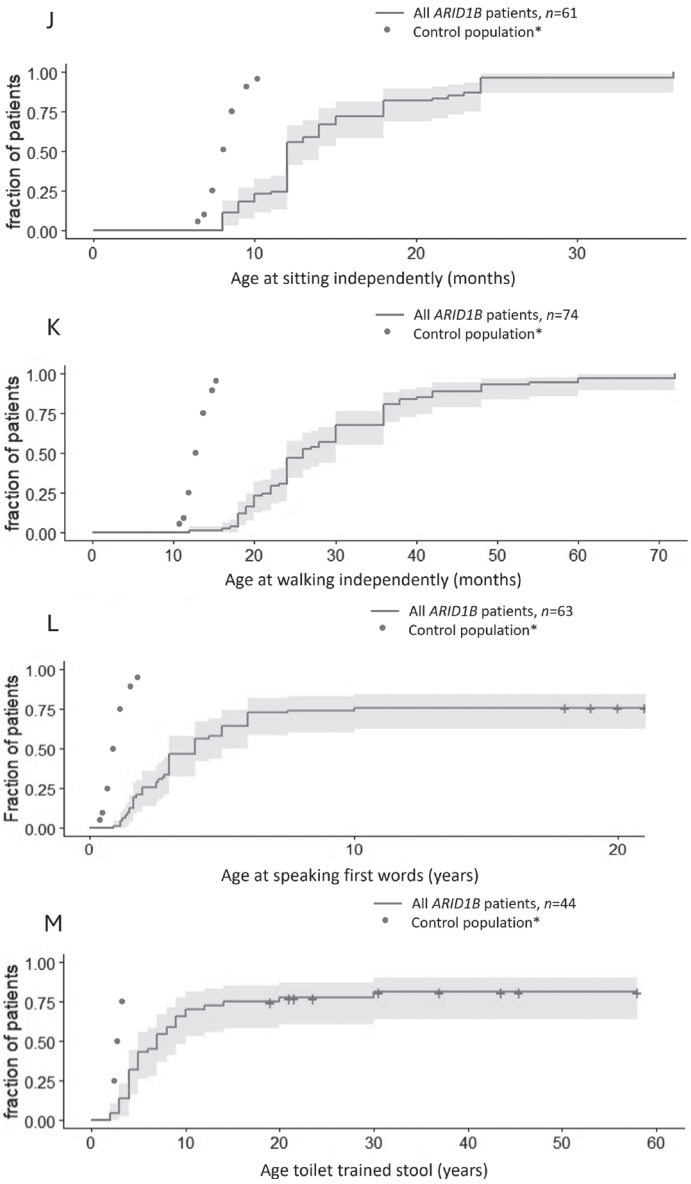


Figure 1. General characteristics of the adult *ARID1B* population.

(**A**) Facial photos of *ARID1B* patients. (**B**) *ARID1B* transcript with pathogenic variants identified in our patients. (**C**) Survival plot for death, *n* = 87. Histogram of the SDS of (**D**) Birthweight, *n* = 53; (**E**) Height, *n* = 69; (**F**) weight for length, *n* = 55; (**G**) OFC, *n* = 55. (**H**) Distribution of ID severity, *n* = 87. (**I**) IQ score, *n* = 32. Cumulative distribution of developmental milestones (**J**) sitting, *n* = 61; (**K**) walking, *n* = 74; (**L**) first words, *n* = 63; (**M**) toilet trained stool, *n* =44. ∗Percentiles of mortality are based on data derived from the Human Mortality database as reported by Beltrán-Sánchez et al[15] and Verguet et al.[16] Percentiles of developmental milestones are based on normative data of children without developmental delay.[17,18] ID, intellectual disability; IQ, intellectual quotient; SDS, standard deviation score.

Table 1. Clinical characteristics of *ARID1B* patients

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Age (nr, min-max)	85	(18-69)	0.64	T	16	(18,5-48,5)	69	(18-69)	2	(35-37)
Sex (female)	85	61%	0.78	Chi	16	56%	69	62%	2	50%
Died	85	2%	1.00	F	16	0%	69	3%	2	0%
Growth parameters & development										
Gestational age, weeks (mean)	73	224	0.19	T	11	275	62	213	1	-
Birthweight (<-2 SDS)	53	11%	0.58	F	7	0%	46	13%	0	-
Height at birth (<-2 SDS)	16	31%	1.00	F	2	0%	14	36%	0	-
OFC at birth (<-2 SDS)	18	6%	1.00	F	3	0%	15	7%	0	-
Age last measurements, years (nr, min-max)	83	(4,2-69)	0.59	T	16	(11,8-40)	67	(4,2-69)	2	(35-37)
Weight (<-2 SDS)	55	0%	-	-	6	0%	49	0%	0	-
BMI >25 kg/m² or overweight	70	56%	1.00	F	10	60%	60	52%	1	100%
Length (<-2 SDS)	69	54%	0.49	F	9	67%	60	52%	2	50%
OFC (<-2 SDS)	55	5%	1.00	F	9	0%	46	7%	1	0%
Motor delay	78	92%	0.01	F	14	71%	64	97%	2	0%
Motor skills gross, delayed	78	76%	0.31	F	14	64%	64	78%	2	0%
Motor skills fine, delayed	78	65%	0.07	F	14	43%	64	70%	2	0%
Speech, delayed	81	98%	0.35	F	15	93%	66	98%	2	0%
Sleeping problems	59	36%	0.24	F	8	13%	51	39%	1	0%
Obstructive sleep apnea	85	1%	1.00	F	16	0%	69	1%	2	0%
Laryngomalacia	53	9%	1.00	F	10	10%	43	9%	2	0%
Feeding difficulties	81	65%	0.03	Chi	15	40%	66	71%	2	0%

Table 1 continues on next page

Table 1. Continued

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Duration of feeding problems	42		0.17	F	5		37		0	
Brief		40%				60%		38%		-
Several years		38%				0%		43%		-
Ongoing		21%				40%		19%		-
Recurrent infections	71	39%	0.75	F	12	33%	59	41%	1	0%
Upper airway tract	71	4%	1.00	F	12	0%	59	5%	1	0%
Lower airway tract	71	3%	1.00	F	12	0%	59	3%	1	0%
ENT infections	71	17%	0.68	F	12	8%	59	19%	1	0%
Otitis media	71	14%	0.67	F	12	17%	59	14%	1	0%
Urinary tract	71	8%	1.00	F	12	8%	59	8%	1	0%
Neurological features										
IQ (nr, min-max)	32	(20-80)	0.48	T	4	(54-65)	28	(20-80)	0	-
Intellectual disability	85	93%	0.08	F	16	81%	69	96%	2	0%
Borderline	0	7%				19%		4%		50%
Mild	0	28%				38%		26%		0%
Mild-moderate	-	-	-	-	-	-	-	-	-	-
Moderate	0	42%				31%		45%		0%
Moderate-severe	-	-	-	-	-	-	-	-	-	-
Severe	0	20%				13%		22%		0%
Profound		0%				0%		3%		0%
Hypotonia	74	77%	0.13	F	12	58%	62	81%	2	0%

Table 1. Continued

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Seizures	81	47%	0.67	F	14	36%	67	49%	2	0%
No seizures, but abnormal EEG		7%				7%		7%		0%
Still experiencing seizures	20	15%	0.15	F	1	100%	19	11%	0	-
Loss of skill	64	25%	0.45	F	10	0%	54	30%	2	0%
Motoric	0	11%				0%		13%		0%
Speech	0	8%				0%		9%		0%
Unspecified	0	8%				0%		9%		0%
Agenesis of the corpus callosum	58	47%	0.94	F	8	50%	50	46%	2	0%
Partial/hypoplasia		31%				38%		30%		0%
Brain abnormality	67	55%	1.00	F	10	60%	57	54%	2	50%
MRI performed	78	85%	0.05	Chi	15	67%	63	89%	1	100%
Vision and hearing impairments										
Vision impaired	82	83%	1.00	F	15	87%	67	82%	2	0%
Myopia	60	78%	1.00	F	12	83%	48	77%	0	-
Hypermetropia	53	26%	0.09	F	8	0%	45	31%	0	-
Cataract	43	7%	1.00	F	7	0%	36	8%	1	0%
Hearing loss	79	28%	1.00	F	15	27%	64	28%	2	0%
Hearing loss, conductive	79	14%	1.00	F	15	13%	64	14%	2	0%
Hearing loss, perceptive	79	8%	0.59	F	15	0%	64	9%	2	0%
Eartubes	56	41%	1.00	F	11	36%	45	42%	2	0%
Hearing aid	17	35%	1.00	F	3	33%	14	36%	0	-
Musculoskeletal anomalies										
Orthopedic anomalies (scoliosis+patella+pes pedes)	85	61%	0.16	Chi	16	44%	69	65%	2	50%

Table 1 continues on next page

Table 1. Continued

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Scoliosis	82	30%	1.00	F	15	27%	67	31%	2	50%
Degree scoliosis	3	(37-75)	-	-	0	-	3	(37-75)	0	-
Operation scoliosis needed	22	32%	0.52	F	3	0%	19	37%	1	0%
Pes planus	46	67%	0.65	F	6	83%	40	65%	2	50%
Patella luxation	45	27%	0.84	F	8	13%	37	30%	1	0%
Recurrent	12	67%			1	100%	11	64%		
Pectus, excavatum	80	3%	1.00	F	15	0%	65	3%	1	0%
Primary dentition, delayed	47	23%	0.66	F	9	11%	38	26%	2	0%
Permanent dentition, delayed	47	40%	0.28	F	9	22%	38	45%	2	0%
Widely spaced teeth	47	15%	1.00	F	9	11%	38	16%	2	0%
Abnormal dentition	36	69%	1.00	F	4	75%	32	69%	2	50%
Dental surgeon operation/treated by a dental surgeon	34	50%	0.38	F	5	20%	29	55%	0	-
Joint laxity	46	50%	0.24	F	8	25%	38	55%	2	0%
Early arthritis	33	6%	1.00	F	7	0%	26	8%	1	0%
Clinodactyly	65	12%	0.63	F	12	17%	53	11%	2	0%
Brachydactyly fifth finger	65	25%	1.00	F	12	25%	53	25%	2	0%
Small nails	66	38%	1.00	Chi	13	46%	53	36%	2	0%
Which nails, 5th finger, and/or toe	66	29%	0.17	Chi	13	46%	53	25%	2	0%
Intestinal										
Inguinal hernia	53	8%	0.54	F	9	11%	44	7%	1	0%
Intestinal problems	73	42%	0.54	F	13	31%	60	45%	1	0%
Constipation	73	27%	0.10	F	13	8%	60	32%	1	0%

Table 1. Continued

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Gastroesophageal reflux	73	11%	0,63	F	13	15%	60	10%	1	0%
Diarrhea	73	0%	-	-	13	0%	60	0%	1	0%
Pyloric Stenosis	73	0%	-	-	13	0%	60	0%	1	0%
Umbilical hernia	73	4%	0.08	F	13	15%	60	2%	1	0%
Cardiac & urogenital anomalies										
Cardiac anomalies	65	14%	1.00	F	12	8%	53	15%	1	0%
ASD	65	6%	1.00	F	12	0%	53	8%	1	0%
VSD	65	0%	-	-	12	0%	53	0%	1	0%
Aortic valve abnormality	65	3%	0.34	F	12	8%	53	2%	1	0%
Mitralis insufficiency	65	2%	1.00	F	12	0%	53	2%	1	0%
Renal anomalies	42	43%	0.01	F	6	0%	36	50%	0	-
Hydronephrotic kidney	42	10%	1.00	F	6	0%	36	11%	0	-
Nephrolithiasis	42	21%	0.31	F	6	0%	36	25%	0	-
Renal sonography, abnormal	43	42%			6	0%	37	49%	0	-
Age identification of first renal stone (nr, min-max)	5	(7-59)	-	-	0	-	5	(7-59)	0	-
Cryptorchidism	30	60%	0.66	F	6	50%	24	63%	1	0%
Endocrinological abnormalities										
Diabetes mellitus	54	11%	1.00	F	8	13%	46	11%	0	-
Type 2 diabetes mellitus	54	11%	1.00	F	8	13%	46	11%	0	-
Hypothyroidism	54	15%	1.00	F	8	13%	46	15%	0	-
Growth hormone deficiency	54	2%	1.00	F	8	0%	46	2%	0	-

Table 1 continues on next page

Table 1. Continued

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Other										
Anemia	54	6%	1.00	F	8	0%	46	7%	0	-
Elevated cholesterol	54	7%	1.00	F	8	0%	46	9%	0	-
Hypertension	35	17%	0.56	F	5	0%	30	20%	0	-
Behavioral abnormalities	80	85%	0.40	F	13	77%	67	87%	2	50%
Hyperactivity	75	7%	0.59	F	13	0%	62	8%	2	0%
High pain threshold	53	64%	0.26	F	9	44%	44	68%	1	0%
Psychiatric disorders										
ADHD	80	9%	1.00	F	13	8%	67	9%	2	50%
Autistic traits	80	26%	0.50	F	13	15%	67	28%	2	0%
Autism	80	31%	0.75	F	13	23%	67	33%	2	0%
Age autism diagnosis (nr, min-max)	21	(0-25)	0.89	MW	3	(3-12)	18	(0-25)	0	-
Auto-mutilation	80	19%	0.11	F	13	0%	67	22%	2	0%
Malignancies	73	1%	1.00	F	12	0%	61	2%	2	0%
Lifestyle										
Daycare	49	65%	0.15	F	7	57%	42	67%	0	-
Regular	0	22%				43%		19%		
Special	0	43%				14%		48%		
Primary education	59	100%	0.15	F	8	100%	51	100%	0	-
Regular		2%				13%		0%		
Special		69%				88%		67%		

Table 1. Continued

Patient Groups:	**18+ (LoF)**				**LoF Exon 1**		**LoF > Exon 1**		**Mosaic**	
Clinical Features [+]	***n* = 85**	**%**	***P Value*** [a]	***Test*** [a]	***n* = 16**	**%**	***n* = 69**	**%**	***n* = 2**	**%**
Secondary education	45	73%	0.38	F	7	86%	38	71%	0	-
Regular	0	67%				14%		5%		
Special	0	0%				71%		66%		
Living situation	57		0.22	F	8	0%	49	0%	0	-
At home/with parents		67%				75%		65%	0	-
Independently guided/assisted living		9%				13%		8%	0	-
Residential group (>residents/caretaker)		19%				0%		22%	0	-
Residential group (1 on 1 guidance)		5%				13%		4%	0	-
Medication	65	74%	0.69	F	9	67%	56	75%	2	100%
Anti-epileptics	65	23%	1.00	F	9	22%	56	23%	2	0%
Anti-depressants	65	12%	0.31	F	9	22%	56	11%	2	50%
Anti-psychotics	65	12%	0.59	F	9	0%	56	14%	2	0%
Diuretics/Anti-hypertensives	65	12%	1.00	F	9	11%	56	13%	2	50%
Amphetamines	65	5%	1.00	F	9	0%	56	5%	2	0%
Anti-diabetics	65	8%	1.00	F	9	0%	56	9%	2	0%
Hypo-/hyperthyroidism medication	65	8%	0.14	F	9	22%	56	5%	2	50%
Laxatives	65	15%	0.33	F	9	0%	56	18%	2	0%
PPI	65	11%	0.58	F	9	0%	56	13%	2	0%
Other	65	43%	0.07	F	9	11%	56	48%	2	50%

A, ANOVA; ADHD, attention deficit hyperactivity disorder; ASD, Atrial Septal Defect; BMI, body mass index; Chi, Chi-square; EEG, electro encephalography; ENT, ear nose throat; F, Fisher's exact; KW, Kruskal-Wallis; LoF, loss of function variants; MW, Mann-Whitney U; OFC, occipital–frontal circumference; SDS, standard deviation score; T, T-test; VSD, ventricular septal defect.

[+] the total number of a feature can differ from the sum of subcategories because in some cases it was possible to answer with more than 1 option or to report the existence of a feature without specifying.

[a] Groups compared are patients with a pathogenic variant in exon 1 versus patients with an exon 2-20 variant or a deletion in *ARID1B*.

Phenotype

An overview of patients' characteristics is given in Table 1 and Supplemental Table 2. Ages ranged from 18 to 69 years with a median of 23.3 years. One patient died at the age of 24 years due to renal abscesses, and 1 patient died at the age of 47.2 years by asphyxiation due to choking on food (see also Figure 1C).

Frequencies reported henceforth concern the 85 non-mosaic patients.

Congenital anomalies

Frequently reported congenital anomalies are agenesis of the corpus callosum (27/58, 47%), cardiac anomalies (9/65, 14%), renal abnormalities (18/42, 43%), and cryptorchidism (18/30, 60%).

Growth

Birthweight below 2 SDS was observed in 11% (6/53) of patients (Figure 1D), and feeding difficulties were reported in 65% (53/81). For the majority, feeding issues started at birth (74%, 34/46) and were transient in 79% of cases (brief: 40%, 17/42; several years: 38%, 16/42), with 21% (9/42) experiencing ongoing difficulties.

Histograms of the SDS of height, weight, and occipital-frontal circumference (OFC) are shown in Figure 1E-G. The majority of patients have a height below 0 SDS (97%, 67/69); 56% (39/70) have a body mass index above 25 kg/m² or are reported to be overweight; OFC is distributed normally around 0 SDS.

Development

Ninety-three percent of patients have ID (Table 1, Figure 1H-I), with total IQ scores (n = 32) ranging from 20-80 (Figure 1I). Eight patients had borderline or normal intelligence (i.e. an estimated normal intelligence or an IQ score of 80 or higher). IQ values were available for only 3 patients with an estimated borderline or normal IQ score. One patient had a total IQ of 80 measured at the age of 14 years, another patient had a verbal IQ of 92 and a performance IQ of 70 at the age of 18 years, and the last patient had a verbal IQ of 83 and a performance IQ of 70 at an unknown age. Of the 8 patients with a borderline or normal IQ, 5 had behavioral anomalies: attention deficit hyperactivity disorder in 4 patients and autistic features in 2 patients. Self-mutilation was not reported in this group.

Hypotonia was observed in 77% (57/74). Figures 1J-L show developmental milestones. Motor and speech are delayed in most patients. While almost all patients eventually walk, approximately one-third of patients do not develop speech. Seventy-five percent

of patients are toilet trained (Figure 1M, Supplemental Figure 3A). Fifty-five patients can read, and 33 patients can write. The age of puberty onset (*n* = 39) varied between 9 and 21 years (Supplemental Figure 3B).

Seizures occurred in 47% (38/81) of patients, with an additional 7% (6/81) with an abnormal electroencephalogram (EEG). The age of onset (*n* = 80) varied between 0 and 41 years (Figure 2A). Loss of skills was noted in 25% (16/64) of patients. No specific triggering event was reported. The age of onset was documented in 4 cases. This age ranged from 31 to 54 years (Figure 2B). Loss of motor skills, particularly in walking ability (with increased tripping and decreased balance) was the most common (7/16), followed by loss of speech (5/16). Some patients required the use of a wheelchair (5/62).

Vision and hearing

Most patients had impaired vision 83% (68/82), usually due to myopia (78%, 47/60). Myopia severity (*n* = 27) ranged from -1 to -25 (Supplemental Figure 3C), with 14 patients having a severity below -10. Hearing loss was present in 28% (22/79) of patients; with 6 patients needing a hearing aid. Both conductive and perceptive hearing impairment are reported, with conductive hearing impairment being slightly more prevalent (Supplemental Table 2).

Other features

Diabetes mellitus type 2 was reported in 11% (6/54) of patients (Figure 2C), hypothyroidism in 15% (8/54) of patients (Figure 2D), and nephrolithiasis in 21% (9/42) (Figure 2E). Recurrent infections were present in 39% (28/71) of patients. Anemia was identified in 6% (3/54), elevated cholesterol in 7% (4/54), and hypertension in 17% (6/35) of patients. Sleeping problems were reported in 36% (21/59).

Behavioral anomalies were reported by clinicians in 85% (68/80). Seventy-five percent (36/48) of parents report the behavior of their child as being problematic ("sometimes" 25/48 [52%] or "often" 11/48 [23%]). Auto-mutilation was reported in 19% (15/80) of patients.

Abnormal dentition was noted in 69% (25/36) of patients, and 50% (17/34) received treatment from a dental surgeon. Although the specific reasons for dental treatment were not explicitly asked, it was reported that several patients required the extraction of multiple teeth.

Approximately one-third (30%, 25/82) of patients developed scoliosis, with 32% of those (7/22) requiring surgery. Pes planovalgus was reported for 67% (31/46), and 60% (25/42)

of patients used support insoles. Talocalaneonavicular dislocation was reported in the surveys' open fields. Patella luxation (Figure 2F) was reported in 12/45 patients, and in 8 cases these were recurrent. Patella luxation occurred at ages between 4 and 49 years with a median of 27.9 years. At least 3 patients were reported to have undergone surgery for recurrent patella luxation, as noted in the free text, although no specific question regarding this surgery was included.

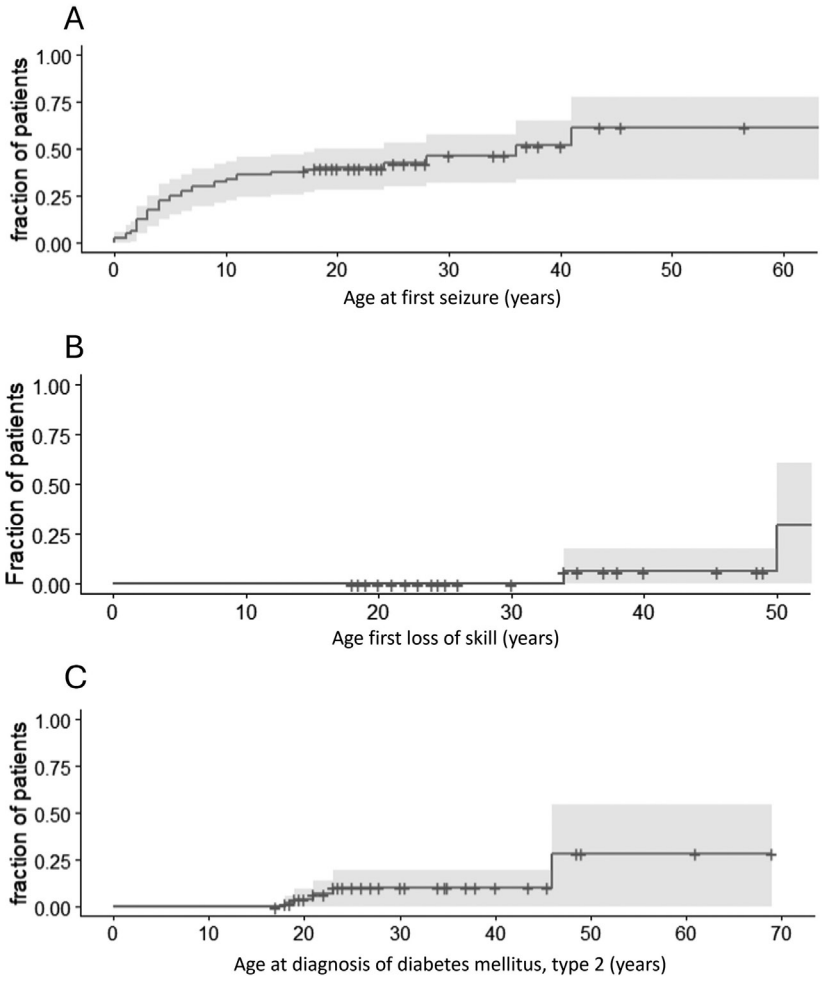
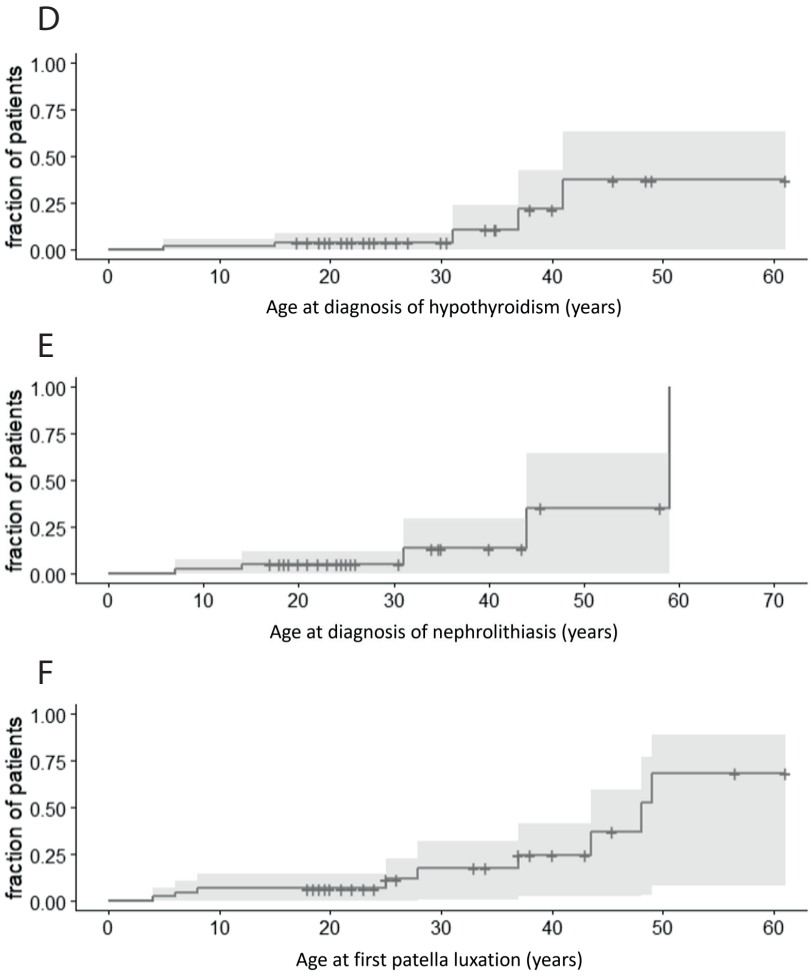


Figure 2. Survival analysis of features developing later in life in adult *ARID1B* patients. Kaplan-Meier plot for the age of (**A**) onset first seizure, *n* = 80; (**B**) loss of skill, *n* = 54; (**C**) diabetes mellitus, type 2, *n* = 53; (**D**) hypothyroidism, *n* = 51; (**E**) nephrolithiasis, *n* = 38; (**F**) first patella luxation, *n* = 43.

One patient in our cohort was reported to have a lung tumor, but based on radiological examination and consultation with a pulmonologist, it was determined to be likely benign. Therefore, a biopsy was deemed to be overly burdensome in this patient with severe ID.



Functioning

Table 2 and Supplemental Table 3 show to what extent the included adult patients with *ARID1B*-related disorder were able to perform activities of daily living based on our parental questionnaire. Many (45/51, 88%) patients were able to eat and drink independently, 73% (37/51) were able to dress without help, 30% (15/50) independently shop for groceries and 8% (4/49) could prepare dinner independently. Sixty-five percent of the patients (30/46) could stay home alone for 30 minutes, and 16% (7/45) could travel alone by public transport. Most patients lived at home with parents (38/57, 67%), 9% of patients (5/57) lived in assisted living, and the remaining patients lived in groups or had 1 on 1 guidance.

Medication use

Medication is utilized by 74% (48/65) of patients in the cohort. Among the reported medications, anti-epileptic drugs (23%, 15/65) are the most frequently used, followed by laxatives (15%, 10/65), anti-depressants (12%, 8/65), and anti-psychotics (12%, 8/65) (Table 1). A combination of antidepressant or antipsychotic medication is used by 20% (13/65) of patients. Other medications include antihypertensive drugs (12%, 8/65), anti-diabetics (8%, 5/65), and medication for hypo/hyperthyroidism (8%, 5/65). Response to medication was assessed only for seizure medication. Frequently prescribed anti-convulsive medications were valproic acid (n = 5), lamotrigine (n = 3), carbamazepine (n = 3), and, levetiracetam (n = 3). Ninety percent of patients (18/20) responded well to anti-convulsive therapy.

Variants in exon 1 lead to a milder phenotype than deletions or variants in exon 2-20

All pathogenic *ARID1B* variants inherited from non-mosaic parents are located in exon 1 (Figure 1A). Among patients with exon 1 variants, 19% (3/16) exhibit borderline to no ID; in comparison, only 4% (3/69) of individuals in the exon 2-20 group are described as having borderline to no ID (P = .07). In addition, exon 1 patients tend to have less fine motor delay (P = .01), fewer feeding difficulties (P = .03) and no reported renal anomalies (P = .01) (Table 1, Supplemental Table S2, Supplemental Figure 4). On all activities of daily living mentioned in Table 2, a higher proportion of patients with the exon 1 variant score as 'independent', indicating that they have a higher level of self-sustainability. For example, 86% of patients with exon 1 variants can take a bath or shower independently compared to 34% in the exon 2-20 and whole gene deletion group (P = .05).

Table 2. Activities of daily living: parent-reported outcomes for their adult children

	18+ (LoF)				LoF patients exon 1		LoF patients >exon 1	
	n = 85	%	*P* Value[a]	Test	*n* = 16	%	*n* = 69	%
Can your child make his/her own bed?	49		0.48	F	7		42	
With help		53%				57%		52%
Independently		31%				43%		29%
Can your child clean up, and do light house-work?	51		0.73	F	7		44	
With help		59%				71%		57%
Independently		27%				29%		27%
Can your child do the groceries?	50		0.06	F	7		43	
With help		34%				43%		33%
Independently		30%				57%		26%
Can your child replace a lamp, or tighten a screw?	50		0.00	F	7		43	
With help		32%				71%		26%
Independently		14%				29%		12%
Can your child do the laundry?	49		0.03	F	7		42	
With help		47%				57%		45%
Independently		16%				43%		12%
Can your child take a bath or shower?	51		0.05	F	7		44	
With help		51%				14%		57%
Independently		41%				86%		34%
Can your child brush his/her teeth and comb his/her hair?	51		0.19	F	7		44	
With help		47%				29%		50%
Independently		37%				71%		32%
Can your child dress and undress him/herself?	51		0.26	F	7		44	
With help		22%				0%		25%
Independently		73%				100%		68%
Can your child go to the toilet?	51		0.74	F	7		44	
With help		16%				0%		18%
Independently		78%				100%		75%
Can your child make sandwiches?	51		0.16	F	7		44	
With help		29%				14%		32%
Independently		51%				86%		45%
Can your child fry an egg, make pancakes, or heat food in the microwave	50		0.25	F	7		43	
With help		48%				57%		47%
Independently		16%				29%		14%

Table 2 continues on next page

Table 2. Continued

	18+ (LoF)				LoF patients exon 1		LoF patients >exon 1	
	n **= 85**	**%**	***P*** **Value[a]**	**Test**	***n*** **= 16**	**%**	***n*** **= 69**	**%**
Can your child prepare dinner?	49		0.02	F	7		42	
With help		37%				57%		33%
Independently		8%				29%		5%
Can your child set and clear the table?	50		0.41	F	7		43	
With help		30%				14%		33%
Independently		58%				86%		53%
Can your child drink from a cup?	51		1.00	F	7		44	
With help		0%				0%		0%
Independently		96%				100%		95%
Can your child eat from a plate?	51		1.00	F	7		44	
With help		8%				0%		9%
Independently		90%				100%		89%
Can your child do the dishes or load the dishwasher?	47		0.64	F	7		40	
With help		30%				29%		30%
Independently		55%				71%		53%
Can your child handle money, and pay in the store?	46		0.21	F	7		39	
With help		35%				43%		33%
Independently		13%				29%		10%
Can your child stay home alone for 30 minutes?	46		0.39	F	7		39	
Yes		65%				86%		62%
Can your child travel alone by public transport?	45		0.30	F	7		38	
Yes		16%				29%		13%

F, Fisher's exact; LoF: loss of Function variants.
[a]Groups compared are patients with a pathogenic variant in exon 1 versus patients with an exon 2-20 variant or a deletion in *ARID1B*.

Facial features

Using the most recent facial photographs, patients with *ARID1B*-related disorders (*n* = 48) were distinguishable from age and sex-matched NDD controls (analysis 1: $P < .01$) (Supplemental Table 4, Supplemental Figure 5). This distinction held true when stratified by age groups (0-4, 5-10, 11-17, 18-25, and 25+ years) (analysis 2: $P < .01$). Notably, the age group 11-17 years displayed the lowest Brier score and the highest area under the curve (AUC). Using 14 facial photos of patients aged below 11 years and the same 14

patients aged above 25 years, the photos of these patients aged below 11 years were more significantly different from controls (analysis 3a: $P < .01$) compared to photos of these patients aged above 25 years (analysis 3b: $P = .05$). Facial photos of 8 patients with variants in exon 1 were not distinguishable from NDD controls (analysis 4a: $P = .65$), while the 43 photos of patients with variants outside exon 1 or whole gene deletions differed from NDD controls (analysis 4b: $P < .01$). Additionally, when comparing the 7 photos (one photo could not be age-matched) of patients with exon 1 variants to those with variants in other exons, no significant difference was observed (analysis 4c: $P = .84$).

Facial photos of patient 066, who had a mosaic *ARID1B* pathogenic variant, were analyzed at different ages (0.6, 4, 10, 18, and 40 years) and were found to cluster with photos of patients with non-mosaic pathogenic variants (Supplemental Table 5). When comparing each photo to the photos of patients and controls in the corresponding age group, those taken at ages 11 and 17 years exhibited the most consistent clustering patterns.

DISCUSSION

We report the first adult-aged cohort of 87 patients with pathogenic *ARID1B* variants. We confirmed our previous hypothesis[11] that patients with variants predicted to lead to haploinsufficiency in exon 1 tend to have a milder ID phenotype. In addition, we determined that (adult-aged) patients with *ARID1B*-related disorder have a risk of recurrent patella luxation, loss of skills, and auto-mutilation. Additionally, we confirmed our previous findings that patients with *ARID1B*-related disorder have a risk of seizures, myopia, nephrolithiasis, hypothyroidism, diabetes mellitus type 2, and scoliosis.

Genotype

All patients in our cohort have variants predicted to lead to haploinsufficiency. Some pathogenic missense variants have been reported in literature,[8,21, 22, 23] but they are much less common than predicted loss-of-function variants. Further research is necessary to study these missense variants and determine if patients carrying such variants differ from those with predicted loss-of-function variants, which have been shown to lead to nonsense-mediated decay on several occasions.[24, 25, 26] Further studies are needed to confirm that nonsense-mediated decay is happening with most predicted loss-of-function variants versus the formation of truncated proteins.

Most of the variants occurred *de novo*, but several variants were inherited. All variants inherited from nonmosaic parents were located in exon 1. In retrospect, these parents

have several features fitting with *ARID1B*-related disorder, indicating a full penetrance of these variants.[27]

Genotype-phenotype

As shown in Figure 1B, patients with pathogenic variants in exon 1 of *ARID1B* tend to have milder ID. It is hypothesized that variants at the start of exon 1 may not be pathogenic.[5,11] Based on our current data, we conclude that predicted loss-of-function variants in exon 1 are pathogenic, but tend to lead to a milder phenotype compared to variants located further on the transcript.

This phenomenon might be explained by a partial rescue because of an alternative start site.[28] For example, in the GTEx portal[29] there are several transcripts starting later than the exon 1 start site of NM_020732.3, and there are also several transcripts starting before exon 1 start site of NM_020732.3.

Phenotype

This study provides the first comprehensive assessment of the self-sustainability of adult-aged patients with pathogenic *ARID1B* variants. Parents of newly diagnosed individuals often inquire about the extent to which this patient group is toilet trained, able to read and/or write, has received education, and can perform activities of daily living, and those questions can now be answered.

Our study confirms the wide spectrum of individuals with *ARID1B*-related disorders. This spectrum includes individuals without ID who live independently, typically associated with very early exon 1 variants or mosaic variants. On the other end of the spectrum are individuals with severe ID requiring 24-hour care. Between these extremes, we found patients with varying levels of independence (Table 2).

We also confirmed that height in *ARID1B* patients is lower than that of the general population with an average SDS of -2.0. We used published growth charts to impute missing SDS. These growth charts are based on a predominantly White population. It could, therefore, be possible that in the 7 cases with a mixed or non-White ancestry SDS for height was overestimated. However, based on these growth charts, 3 of 7 patients had a height below -2 SDS, which is similar to the distribution of the group as a whole. Myopia (78%) and hypermetropia (26%) are more prevalent in our population, compared with the general population where myopia is present in 4.9% to 18.2% and hypermetropia in 2.2% to 14.3%.[30] Age of puberty onset is at an average of 14.7 years, with several outliers starting at the age of 17 to 21 years. These numbers are based on both clinician and parent reports.

New features identified in this cohort are recurrent patella dislocation and loss of skill. Recurrent patella luxation can have several causes, including weakness of the thigh muscles, or excess pronation of the feet. Both factors can play a role in our population as many patients (77%) experience hypotonia, 67% have pes planovalgus and in 1 patient, a talocalaneonavicular dislocation has been reported. The loss of skills observed in 25% of patients is important to consider when caring for a patient with a pathogenic *ARID1B* variant. This number may be biased, as loss of skills may be the reason to refer an adult to a genetic center for diagnostic evaluation. Further investigation is required to assess whether timely interventions can, for example, improve functioning or aid in restoring motor function. In instances where the onset of loss of skills can be linked to a specific event, eye movement desensitization and reprocessing therapy may offer valuable support.

Interestingly, although feeding problems are frequent in childhood, we observe that over half of our adult cohort is overweight (body mass index > 25 kg/m², or as reported by a parent). People with intellectual disabilities are generally less physically active,[31] and research has demonstrated the benefits of exercise in improving cardiorespiratory and muscular fitness.[32] Given the correlation of obesity and type 2 diabetes (4/5 individuals with this disease had obesity in our cohort), on indication, diabetes should be tested. This also goes for other features that develop with age (Supplemental Tables 6-7), and therefore, we have formulated screening recommendations (Table 3). Furthermore, it is noteworthy that aside from laxatives, anti-epileptic drugs, anti-depressants, and anti-psychotics are the most prescribed medications in our patient group, reflecting the impact of behavior on daily living and patient management.

We also noted some differences with current frequency estimates, compared to our previous work. For example, in one publication[5] 75% of patients had developed speech by the age of 5 years. In our cohort, only 60% of patients have developed speech by the age of 5 years, although this estimate rises toward 75% at the age of 7 years (Figure 1L). This difference may be caused by random variation, but may also point to a more severely affected cohort because of either ascertainment bias (genetic testing is more often done in more severely affected adult-aged patients) or a reduced quality of care in this older cohort. Similarly, Figure 2A indicates that 60% will develop epilepsy whereas previously this estimate was 35%.[5] This difference seems to be caused by a substantial proportion having their first seizure after the age of 20 years, which we previously missed because of reduced follow-up.

Table 3. Recommendations surveillance guidelines *ARID1B* patients

Evaluation (Inquire/Perform physical examination)	**After diagnosis**	**Age category**			**(para) medic**[a]
		Children (0-18 years)	***Adults (18+)***	**Frequency**	
Congenital abnormalities (heart, kidneys and cryptorchidism)	Yes	Upon indication		Upon indication	Treating physician
Feeding difficulties	Yes	Yes		At every visit	Pediatrician
Constipation	Yes	Yes	Yes	At every visit	Treating physician
Growth and weight	Yes	Yes	Yes	At every visit, to avoid excessive weight gain	Pediatrician, dietitian
Seizures	Yes	Yes	Yes	At every visit	Treating physician
Endocrine/hormonal	Upon indication	Upon indication	At every visit[b]	Upon indication or every 3 years if there are risk factors (i.e. overweight)	Treating physician
Vision	Yes	Yes	Yes	Every 2 years, the frequency can be adjusted if the patient is able to report on eyesight	Ophthalmologist
Hearing	Yes	Yes	Upon indication	At every visit ask for signs of hearing loss and refer to an audiologist if suspected	Treating physician
Nephrolithiasis	Yes	Yes	Upon indication	At every visit	Treating physician/urologist
Scoliosis	Yes	Yes	Upon indication	Periodically, until length growth is complete	Orthopedic surgeon/ physiatrist/ pediatrician
Patella luxation	Yes	Upon indication	Upon indication	Upon indication	Orthopedic surgeon/ physiatrist/ pediatrician
Pedes (plano)valgi	Yes	Yes	Upon indication	Periodically	Orthopedic surgeon/ physiatrist/ pediatrician
Dentition/dental health	Yes	Yes	Yes	Twice a year	Dentist
Motor development and/or loss of skill	Yes	Yes	Upon indication	At every visit	Pediatrician/ physiatrist/ physical therapist/ treating physician
Cognitive development	If aged <16	Yes		At every visit	Pediatrician/ psychologist

Table 3. Continued

Evaluation (Inquire/Perform physical examination)	After diagnosis	Age category			(para) medic[a]
		Children (0-18 years)	***Adults (18+)***	**Frequency**	
Communication/language development and/or loss of skill	Yes	Yes	Upon indication	Screening at age 2 and 3 years, then on indication	Pediatrician/ physiatrist/ speech therapist/ treating physician
Behavioral and social development/impairments	Yes	Yes	Yes	At every visit	Treating physician
Sexual development	When applicable	Yes	Yes	When applicable	Treating physician
Transition to adult care	When applicable	When applicable		From the age of 16 years	Physiatrist/ pediatrician

[a] treating physician (eg, general practitioner, pediatrician).
[b] especially, glucose and thyroid lab.

Somatic variants in *ARID1B* have been associated with several types of cancer.[33] In the literature only occasional cases of *ARID1B* patients with cancer are reported.[5,34] We did not identify an increased cancer risk in our cohort. Based on our cohort and literature, there is no indication that germline variants in *ARID1B* give an increased cancer risk at pediatric age. Although there is currently no evidence to suggest an elevated cancer risk in adult-aged patients, it is important to acknowledge that our study had a limited representation of patients over the age of 50 years, and further longitudinal research is needed to confirm this.

Pediatric versus adult cohort

Compared to the previously published predominantly pediatric cohort[5] (Supplemental Figure 6), our adult cohort exhibits a notable shift in phenotype (Supplemental Table 6-7). One example is the previously mentioned shift from feeding difficulties in children to overweight in adults. Similarly, recurrent infections are significantly less prevalent in our adult cohort (39% compared to 57%,[5] $P = .03$). In children with *ARID1B*-related disorder, the primary emphasis often revolves around development and the acquisition of new skills. However, as these patients transition into adulthood, the focus shifts towards the preservation of current skills or the prevention of loss of skills and maintenance of muscular and cardiorespiratory fitness. This divergence in focus underscores the evolving needs and priorities of individuals with this condition as they grow older.

Facial photograph analyses

Our study demonstrated that *ARID1B* patients are distinguishable from matched NDD controls based on facial features in infancy. However, in our earlier analysis, there was an indication these distinctive facial features may become less specific as individuals with *ARID1B*-related disorders age,[11] whereas in the current analysis, all age groups clustered separately from matched NDD controls. When using fewer photographs, we did see less significant clustering results for photos of patients aged 25+ years (Supplemental Table 5), indicating the difference in results from our previous analysis may be explained by the then limited number of available photos.

Nonetheless, if facial analysis is implemented to assist with interpretation, we recommend using childhood photographs, especially since most available photos are still of patients aged below 18 years. This approach can provide a more reliable basis for accurate diagnosis and interpretation.

Limitations

Our study has several potential limitations that should be considered. One is that multiple clinicians contributed to data entry, which may have introduced variability and dataset inconsistencies. There is also a risk of overestimating the prevalence of certain features due to the unknown status regarding specific features in some patients as, in our experience, a clinician is more likely to tick the box "unknown" for a feature than "absent." In addition, there may be ascertainment bias in our study towards more severe cases as genetic diagnostics may be performed more often on more severely affected adult patients. We tried to limit this by including as many adult-aged patients as we could find.

CONCLUSION

The *ARID1B* spectrum is broad, and as patients age, there is a significant shift in the medical aspects that need attention. Several features warrant extra attention in adult patients, and screening and treatment of these features may prevent progression. Therefore, we have updated our screening recommendations[5,35] for all age groups to promote timely intervention in an attempt to potentially improve health outcomes (Table 3).

DATA AVAILABILITY

De-identified patient data will be made available on request to the corresponding author.

CONFLICT OF INTEREST

Jill A. Rosenfeld: The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing completed at Baylor Genetics Laboratories. Evan E. Eichler is a scientific advisory board (SAB) member of Variant Bio, Inc. All other authors declare no conflicts of interest.

ACKNOWLEDGMENTS

We would like to thank Johan den Dunnen for his assistance in adding patient data to the Leiden Open Variant Database (LOVD).

FUNDING

This work was supported, in part, by US National Institutes of Health (NIH) grant MH101221 to E.E.E. (E.E.E. is an investigator of the Howard Hughes Medical Institute), by "the Fundamental Research Funds for the Central Universities" starting fund (BMU2022RCZX038) to T.W., by the Ministero dell'Istruzione, dell'Università e della Ricerca via PNRR-MUR-M4C2 PE0000006 Research Program "MNESYS"—A multiscale integrated approach to the study of the nervous system in health and disease (to Pa.S.). IRCCS 'G. Gaslini' is a member of ERN-Epicare, and by the Dutch Organisation for Health Research and Development: ZON-MW grants 912-12-109 (to B.B.A.d.V.) and Donders Junior researcher grant 2019 (B.B.A.d.V.).

Sequencing and analysis of individual 101 was provided by the Broad Institute Center for Mendelian Genomics funded by the National Human Genome Research Institute grants UM1HG008900 (with additional support from the National Eye Institute, and the National Heart, Lung and Blood Institute) and R01HG009141.

ETHICS DECLARATION

The research included in this report was conducted in a manner consistent with the principles of research ethics. The Leiden University Medical Center's Institutional Review Board granted approval waivers for using de-identified and aggregated data (no: G18.098) without requiring specific informed consent. Patient data was de-identified using assigned numbers, and, when feasible, informed consent was obtained through the referring clinician. Written consent was obtained and archived for all included patient photos.

SUPPLEMENTARY MATERIAL

Supplementary Tables and Figures

The supplementary material consists of 7 tables, 6 figures and 1 supplementary excel file that support the results reported in this article. All tables, figures and the excel are mentioned in the text. This supplementary material includes, but is not limited to, variable labels of the online questionnaire questions, genetic variants, facial photographs, extensive overviews of the clinical characteristics of the patients reported in this paper, PhenoScore analyses, and a comparison of our patient group with a previously published cohort of *ARID1B* patients.

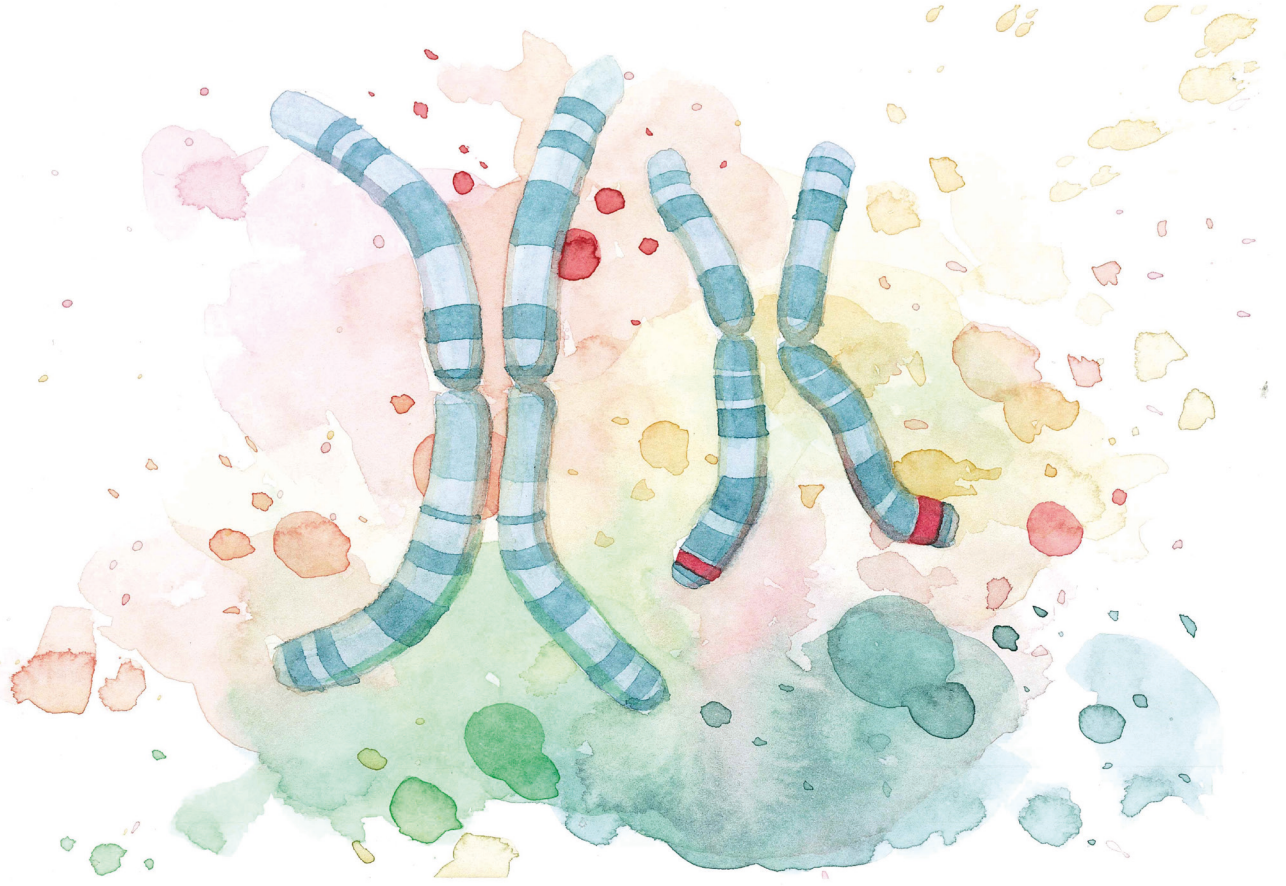
Supplementary Information (5.3MB, pdf)

Supplementary Data (47.8KB, xlsx)

REFERENCES

1. Hoyer J., Ekici A.B., Endele S., et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am J Hum Genet. 2012;90(3):565–572. doi: 10.1016/j.ajhg.2012.02.007.
2. Wright C.F., Fitzgerald T.W., Jones W.D., et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet. 2015;385(9975): 1305–1314. doi: 10.1016/S0140-6736(14)61705-0.
3. Gillentine M.A., Wang T., Eichler E.E. Estimating the prevalence of de novo monogenic neurodevelopmental disorders from large cohort studies. Biomedicines. 2022;10(11) doi: 10.3390/biomedicines10112865.
4. Valencia A.M., Sankar A., van der Sluijs P.J., et al. Landscape of mSWI/SNF chromatin remodeling complex perturbations in neurodevelopmental disorders. Nat Genet. 2023;55(8):1400–1412. doi: 10.1038/s41588-023-01451-6.
5. van der Sluijs P.J., Jansen S., Vergano S.A., et al. The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin-Siris syndrome. Genet Med. 2019;21(6):1295–1307. doi: 10.1038/s41436-018-0330-z.
6. Schrier Vergano S., Santen G.W.E, Wieczorek D., Wollnik B., Matsumoto N., Deardorff M.A. In: GeneReviews®. Adam M.P., Feldman J., Mirzaa G.M., et al., editors. University of Washington; 1993-2024. Coffin-Siris syndrome.https://www.ncbi.nlm.nih.gov/books/NBK131811/
7. Smith J.A., Holden K.R., Friez M.J., Jones J.R., Lyons M.J. A novel familial autosomal dominant mutation in ARID1B causing neurodevelopmental delays, short stature, and dysmorphic features. Am J Med Genet A. 2016;170(12):3313–3318. doi: 10.1002/ajmg.a.37945.
8. Mignot C., Moutard M.L., Rastetter A., et al. ARID1B mutations are the major genetic cause of corpus callosum anomalies in patients with intellectual disability. Brain. 2016;139(11) doi: 10.1093/brain/aww181.
9. Yu Y., Yao R., Wang L., et al. De novo mutations in ARID1B associated with both syndromic and non-syndromic short stature. BMC Genomics. 2015;16(1):701. doi: 10.1186/s12864-015-1898-1.
10. Schipper L.J., Hoekstra-van Duijn M.L.D. Coffin-Siris Syndroom, een overzicht aan de hand van twee casus [Coffin-Siris syndrome, an overview based on two case studies] Tijdschr Artsen Verstandelijk Gehandicapten (TAVG) 2019;37(3):116–121.
11. van der Sluijs P.J., Alders M., Dingemans A.J.M., et al. A case series of familial ARID1B variants illustrating variable expression and suggestions to update the ACMG criteria. Genes. 2021;12(8):1275. doi: 10.3390/genes12081275.
12. van Gameren-Oosterom H.M.B., Fekkes M., Oudesluys-Murphy A.M., van Wouwe J.P. Zelfredzaamheid van jongeren met Down-syndroom [Young people with Down syndrome: independence and social functioning] Ned Tijdschr Geneeskd. 2014;158
13. Fenton T.R., Kim J.H. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. BMC Pediatr. 2013;13:59. doi: 10.1186/1471-2431-13-59.
14. de Onis M., Onyango A.W., Borghi E., Siyam A., Nishida C., Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ. 2007;85(9):660–667. doi: 10.2471/blt.07.043497.
15. Beltrán-Sánchez H., Subramanian S.V. Period and cohort-specific trends in life expectancy at different ages: analysis of survival in high-income countries. SSM Popul Health. 2019;8 doi: 10.1016/j.ssmph.2019.100422.
16. Verguet S., Niwa M., Bolongaita S. Top and bottom longevity of nations: a retrospective analysis of the age-at-death distribution across 18 OECD countries. Eur J Public Health. 2023;33(1):114–120. doi: 10.1093/eurpub/ckac134.
17. Sheldrick R.C., Schlichting L.E., Berger B., et al. Establishing new norms for developmental milestones. Pediatrics. 2019;144(6) doi: 10.1542/peds.2019-0374.

18. Schum T.R., Kolb T.M., McAuliffe T.L., Simms M.D., Underhill R.L., Lewis M. Sequential acquisition of toilet-training skills: a descriptive study of gender and age differences in normal children. Pediatrics. 2002;109(3) doi: 10.1542/peds.109.3.e48.
19. Dingemans A.J.M., Hinne M., Truijen K.M.G., et al. PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. Nat Genet. 2023;55(9):1598–1607. doi: 10.1038/s41588-023-01469-w.
20. Terhörst P, Ihlefeld M, Huber M, et al. QMagFace: simple and accurate quality-aware face recognition. Paper presented at: 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); January 2-7, 2023; Waikoloa, HI:3473-3483. 10.1109/WACV56688.2023.00348
21. Chevarin M., Duffourd Y., Barnard R.A., et al. Excess of de novo variants in genes involved in chromatin remodelling in patients with marfanoid habitus and intellectual disability. J Med Genet. 2020;57(7):466–474. doi: 10.1136/jmedgenet-2019-106425.
22. Yan H., Shi Z., Wu Y., et al. Targeted next generation sequencing in 112 Chinese patients with intellectual disability/developmental delay: novel mutations and candidate gene. BMC Med Genet. 2019;20(1):80. doi: 10.1186/s12881-019-0794-y.
23. Miyamoto S., Kato M., Hiraide T., et al. Comprehensive genetic analysis confers high diagnostic yield in 16 Japanese patients with corpus callosum anomalies. J Hum Genet. 2021;66(11):1061–1068. doi: 10.1038/s10038-021-00932-y.
24. Nord A.S., Roeb W., Dickel D.E., et al. Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. Eur J Hum Genet. 2011;19(6):727–731. doi: 10.1038/ejhg.2011.24.
25. Tsurusaki Y., Okamoto N., Ohashi H., et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet. 2012;44(4):376–378. doi: 10.1038/ng.2219.
26. Sim J.C.H., White S.M., Lockhart P.J. ARID1B-mediated disorders: mutations and possible mechanisms. Intractable Rare Dis Res. 2015;4(1):17–23. doi: 10.5582/irdr.2014.01021.
27. Sekiguchi F., Tsurusaki Y., Okamoto N., et al. Genetic abnormalities in a large cohort of Coffin-Siris syndrome patients. J Hum Genet. 2019;64(12):1173–1186. doi: 10.1038/s10038-019-0667-4.
28. Kearse M.G., Wilusz J.E. Non-AUG translation: a new start for protein synthesis in eukaryotes. Genes Dev. 2017;31(17):1717–1731. doi: 10.1101/gad.305250.117.
29. Exon expression for ARID1B (ENSG00000049618.22). GTEx Portal. https://www.gtexportal.org/home/gene/ENSG00000049618#gene-transcript-browser-block
30. Hashemi H., Fotouhi A., Yekta A., Pakzad R., Ostadimoghaddam H., Khabazkhoob M. Global and regional estimates of prevalence of refractive errors: systematic review and meta-analysis. J Curr Ophthalmol. 2018;30(1):3–22. doi: 10.1016/j.joco.2017.08.009.
31. Hilgenkamp T.I.M., Reis D., van Wijck R., Evenhuis H.M. Physical activity levels in older adults with intellectual disabilities are extremely low. Res Dev Disabil. 2012;33(2):477–483. doi: 10.1016/j.ridd.2011.10.011.
32. Bouzas S., Martínez-Lemos R.I., Ayán C. Effects of exercise on the physical fitness level of adults with intellectual disability: a systematic review. Disabil Rehabil. 2019;41(26):3118–3140. doi: 10.1080/09638288.2018.1491646.
33. Kadoch C., Hargreaves D.C., Hodges C., et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat Genet. 2013;45(6):592–601. doi: 10.1038/ng.2628.
34. Vengoechea J., Carpenter L., Zárate Y.A. Papillary thyroid cancer in a patient with interstitial 6q25 deletion including ARID1B. Am J Med Genet A. 2014;164A(7):1857–1859. doi: 10.1002/ajmg.a.36515.
35. Vergano S.A., van der Sluijs P.J., Santen G. In: GeneReviews®. Adam M.P., Feldman J., Mirzaa G.M., et al., editors. University of Washington; 1993-2024. ARID1B-related disorder.https://www.ncbi.nlm.nih.gov/books/NBK541502/



Chapter 6

Microduplications of *ARID1A* and *ARID1B* cause a novel clinical and epigenetic distinct BAFopathy

Pleuntje J van der Sluijs, Sébastien Moutton, Alexander J M Dingemans, Denisa Weis, Michael A Levy, Kym M Boycott, Claudia Arberas, Margherita Baldassarri, Claire Beneteau, Alfredo Brusco, Charles Coutton, Tabib Dabir, Maria L Dentici, Koenraad Devriendt, Laurence Faivre, Mieke M van Haelst, Khadije Jizi, Marlies J Kempers, Jennifer Kerkhof, Mira Kharbanda, Katherine Lachlan, Nathalie Marle, Haley McConkey, Maria A Mencarelli, David Mowat, Marcello Niceta, Claire Nicolas, Antonio Novelli, Valeria Orlando, Olivier Pichon, Julia Rankin, Raissa Relator, Fabienne G Ropers, Jill A Rosenfeld, Rani Sachdev, Sarah A Sandaradura, Elena Shukarova-Angelovska, Duco Steenbeek, Marco Tartaglia, Matthew A Tedder, Slavica Trajkova, Norbert Winer, Jeremy Woods, Bert B A de Vries, Bekim Sadikovic, Marielle Alders, Gijs W E Santen

Published: Genet Med. 2025 Jan;27(1):101283. doi: 10.1016/j.gim.2024.101283.

ABSTRACT

Purpose: *ARID1A/ARID1B* haploinsufficiency leads to Coffin-Siris syndrome, duplications of *ARID1A* lead to a distinct clinical syndrome, whilst *ARID1B* duplications have not yet been linked to a phenotype.

Methods: We collected patients with duplications encompassing *ARID1A* and *ARID1B* duplications.

Results: 16 *ARID1A* and 13 *ARID1B* duplication cases were included with duplication sizes ranging from 0.1 to 1.2 Mb (1-44 genes) for *ARID1A* and 0.9 to 10.3 Mb (2-101 genes) for *ARID1B*. Both groups shared features, with *ARID1A* patients having more severe intellectual disability, growth delay, and congenital anomalies. DNA methylation analysis showed that *ARID1A* patients had a specific methylation pattern in blood, which differed from controls and from patients with *ARID1A* or *ARID1B* loss-of-function variants. *ARID1B* patients appeared to have a distinct methylation pattern, similar to *ARID1A* duplication patients, but further research is needed to validate these results. Five cases with duplications including *ARID1A* or *ARID1B* initially annotated as duplications of uncertain significance were evaluated using PhenoScore and DNA methylation reanalysis, resulting in the reclassification of 2 *ARID1A* and 2 *ARID1B* duplications as pathogenic.

Conclusion: Our findings reveal that *ARID1B* duplications manifest a clinical phenotype, and *ARID1A* duplications have a distinct episignature that overlaps with that of *ARID1B* duplications, providing further evidence for a distinct and emerging BAFopathy caused by whole-gene duplication rather than haploinsufficiency.

Keywords: *ARID1A*; *ARID1B*; BAF complex; Duplication; Intellectual Disability.

INTRODUCTION

In the last decade, pathogenic germline variants in BAF complex genes have been frequently identified in patients with intellectual disability (ID),[1,2] making it one of the most common causes of ID. *ARID1B* is the most frequently mutated gene in the BAF complex.[3, 4, 5, 6, 7] *ARID1A* (HGNC:11110, NM_006015.6) and *ARID1B* (HGNC:18040, NM_020732.3) are mutually exclusive parts of the complex with a similar structure and function. Loss-of-function variants or deletions in either gene cause nonsyndromic ID or Coffin-Siris syndrome (CSS)[1,3,5] (OMIM 13500 and 614607). These patients all have a BAFopathy episignature present in their blood.[8]

The *ARID1A* and *ARID1B* genes are both triple sensitive genes.[9] Deletions or truncating variants in *ARID1A*, as well as *ARID1B* result in CSS, whereas *ARID1A* whole-gene duplications lead to a distinct clinical syndrome marked by intellectual disability, microcephaly and growth delay in most patients.[10, 11, 12] This suggests that *ARID1B* whole-gene duplications may lead to a similar phenotype; however, no patients with *ARID1B* duplications have been reported so far.

To investigate whether an *ARID1B* duplication phenotype exists, and how this relates to the *ARID1A* phenotype, we included patients with an *ARID1A* or *ARID1B* duplication, compared their (facial) phenotypes,[13] and determined DNA methylation pattern[8] in a subset of patients.

MATERIALS AND METHODS

Patient collection

Patients were identified through various sources, including international contacts, physician referrals for second opinions, the DECIPHER[14] database, and the Coffin-Siris International Facebook group. Furthermore, patients were also recruited through our outpatient CSS expertise center located at Leiden University Medical Center in the Netherlands.

Data collection

Data were collected via an online questionnaire, which was completed by the referring physician.

DNA methylation

DNA samples derived from peripheral blood were available from 15 patients, and DNA methylation data were generated using bisulfited-converted DNA and Illumina

Infinium EPIC arrays. Analysis of the DNA methylation array data was performed at the Verspeeten Clinical Genome Centre in London, Canada, using the established previously described EpiSign episignature development pipeline.[15, 16, 17] Data were normalized, beta values were generated, and cell composition was estimated using the R package sesame (version 1.18.4).[18] Y chromosome probes, probes that contained SNPs at the CpG interrogation or single-nucleotide extension sites, and probes which are known to cross-react with other genomic locations were removed.[19] For each analysis a set of control samples was chosen from the EpiSign Knowledge Database, consisting of disorder-specific samples[17,20] and unaffected controls. Controls included samples from the case batches and from other batches and were matched for sex and age using the R package MatchIt (version 4.5.4). Beta values were converted to M values for linear modeling. Using the R package limma (version 3.56.2), a linear regression model with estimated blood cell composition as a covariate was fitted, and the empirical Bayes method was used to calculate moderated t-statistics and corresponding *P* values. *P* values were adjusted using the Benjamini-Hochberg method to control for false discovery rates. Microarray probes with a mean difference between cases and controls of less than 5% were removed. Probe selection was performed using a combination of effect size and *P* values (absolute methylation difference multiplied by −log(*P* value)), receiver operating characteristic curve analysis, and removal of highly correlated probes, with parameter values selected that maximized the differences between cases and controls as assessed using hierarchical clustering and multidimensional scaling. A support vector machine classifier was created using the R package e1071 (version 1.7-13). To increase specificity for the ARID1 classifiers, the models were trained not just against the matched controls but also against additional controls and disorder-specific samples from the EpiSign Knowledge Database. Support vector machine decision values were converted to probability scores according to the Platt's scaling method. Leave-one-out cross-validation was performed for each case sample and evaluated using hierarchical clustering, multidimensional scaling, and methylation variant pathogenicity scores.

PhenoScore

Quantitative phenotypic similarity analysis was performed with PhenoScore. PhenoScore is a machine learning framework that combines facial recognition technology with Human Phenotype Ontology (HPO)-based phenotype analyses to quantify phenotypic similarity.[21] These analyses were performed as previously described using clinical data of 14 patients with *ARID1A* (9 with facial photographs) and 10 patients *ARID1B* whole-gene duplications (4 with facial photographs). The following comparisons were performed: (1) whether patients were distinguishable from age-, sex-, and ethnicity-matched neurodevelopmental (NDD) controls; (2) whether *ARID1A* duplication patients

were distinguishable from *ARID1B* duplication patients; and (3) whether ARID1 whole-gene duplication patients were recognizable from *ARID1A* and *ARID1B* patients with loss-of-function variants. NDD-matched controls were selected from the Radboud University Medical Center inhouse control database, which contains phenotypic data (both facial and other clinical features in HPO) of 1600 individuals with NDD. *P* values <.05 were considered significant. For each duplication of *ARID1A* or *ARID1B* initially interpreted as a duplication of uncertain significance several PhenoScore analyses were performed. It was determined whether the patient data (ie, HPO-terms) and, when available, facial photos clustered with HPO-terms and photos of ARID1 duplication patients, controls, or patients with loss-of-function variants in *ARID1A* and *ARID1B*.

Interpretation of duplications of uncertain significance

After DNA methylation and PhenoScore analyses of variants with whole-gene duplications, the variants initially classified as a duplication of uncertain significance either because of the size of the duplication or based on the phenotypic assessment of the clinician, were interpreted using the American College of Medical Genetics and Genomics criteria for copy-number variants.[22] American College of Medical Genetics and Genomics-criteria 4L "Statistically significant increase among observations in cases (with a consistent, specific, well-defined phenotype) compared with controls" was allocated when a specific DNA methylation pattern was identified in a patient with a duplication of uncertain significance.

Genotype-phenotype description

Only data of clearly pathogenic variants were included in the phenotype overview and PhenoScore analysis. All analyses were executed using SPSS version 25. R version 4.1.0 was used to create graphs.

RESULTS

Genotype

We identified 14 patients with a whole-gene duplication of *ARID1A* and 2 patients with a duplication of uncertain significance in *ARID1A* either due to the size of the duplication or based on the estimation of the clinician based on the presenting phenotype. In addition, 11 patients with a whole-gene duplication of *ARID1B* and 5 patients (including 2 parents) with a duplication of uncertain significance in *ARID1B* were identified (Supplemental Table 1). The size of the duplications ranged from 0.1 to 1.2 Mb with 1 to 44 genes for *ARID1A* and 0.9 to 10.3 Mb with 2 to 101 genes for *ARID1B* (Figure 1). Inheritance was tested in 24 cases. All whole-gene duplications, except 3 (all *ARID1B*), occurred de

novo. In case 5.1B the duplication was the result of a maternal balanced translocation, whereas in the other cases, V1i_1.1B and V2_1.1B, the duplication was inherited from a potentially affected parent. We also included 3 cases with partial duplications of uncertain significance.

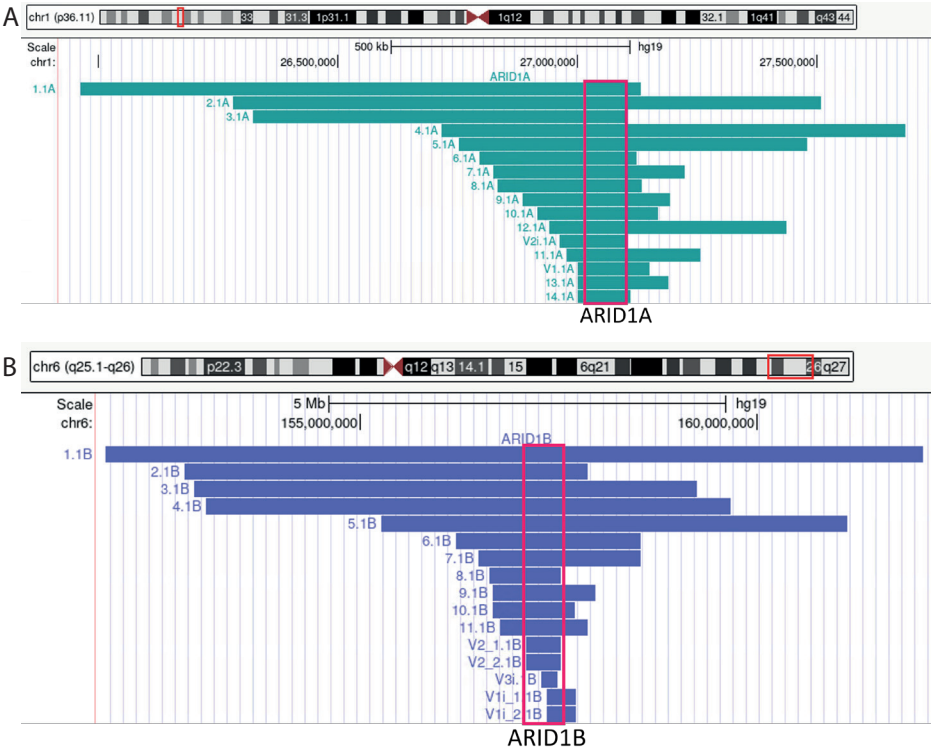


Figure 1. Genomic rearrangements in patients with *ARID1A* and *ARID1B* duplications.
Alignment in the University of California, Santa Cruz (UCSC) genome browser of the genomic rearrangements in our patients with gene duplications of (**A**) *ARID1A* and (**B**) *ARID1B*. The red square indicates the length of the transcript of the gene.

To investigate whether there were any other potential disease-causing genes included in the duplicated regions, we searched the maximal duplicated region in the DECIPHER database[14] and ordered the genes in this region in descending order of pTriplo. For the chromosome 1p36, region *ARID1A* was the top result, and only 1 other gene (ie, *WDTC1* [HGNC:29175]) had a score >0.95. This gene has not been linked to a human phenotype and is present in the duplication of patient 4.1A only and therefore is unlikely to influence our results.

For chromosome 6q25 *ARID1B* was the 2nd result after *SYNE1* (HGNC:17089), which is duplicated in 3 out of 13 patients (ie, 1.1B, 2.1B, and 3.1B). Homozygous *SYNE1* variants have been linked to arthrogryposis phenotype (OMIM 618484) and spinocerebellar ataxia (OMIM 610743), while heterozygous variants can cause Emery-Dreifuss muscular dystrophy (OMIM 612998). Duplications, however, have not been previously associated with human disease. The only other gene with a pTriplo score >0.95 was *SCAF8* (HGNC:20959), which is duplicated in 4/13 *ARID1B* patients (ie, 1.1B, 2.1B, 3.1B, and 4.1B) and has not been linked to a human phenotype. *SCAF8* and *SYNE1* are duplicated in only those *ARID1B* patients with the largest duplications. Our analysis of *ARID1B* duplication size showed no differences in severity or frequency of congenital anomalies, although patient numbers remain low. We therefore tentatively conclude that, based on current knowledge, *ARID1A* and *ARID1B* are the primary genes of interest in the duplicated regions of our patients.

In 3 patients other small duplications or a deletions were identified, of which 2 were deemed nonpathogenic. Patient 7.1A also carried 2 inherited deletions at 8p23.1 and 9p24.1. Both deletions were identified in nonaffected parents and/or siblings. Patient 2.1B also carried a 5q35.3 duplication. Based on previous findings this duplication is unlikely to be pathogenic.[23] Patient 7.1B carries a 3p26.3 duplication (NC_000003.11:g.(?_225645)_(1510822_?)dup) including *CHL1* (HGNC:1939) and *CNTN6* (HGNC:2176). Sporadic cases of a 3p26.3 duplications have been reported[24, 25, 26]; therefore, this variant may have contributed to this patient's phenotype.

DNA methylation

We performed methylation analysis in DNA samples (all isolated from peripheral blood) of 15 cases. In 5 patients with a whole-gene duplication of *ARID1A*, the DNA methylation pattern detected was found to be distinct from controls (Figure 2A and B, Supplemental Figure 1A, Supplemental Table 2). This distinct pattern was opposite to the pattern observed in patients with more common loss-of-function variants of *ARID1A* or *ARID1B* (Figure 2, Supplemental Figure 1D, G, and F). Two cases with variant of uncertain significance (V1.1A and V2i.1A) clustered with cases with *ARID1A* whole-gene duplications, supporting their pathogenicity. DNA methylation analyses in patients with whole-gene duplication of *ARID1B* showed that these cases clustered together and separate from controls (Figure 2C and D, Supplemental Table 2). Unfortunately, cross-validation failed for this potential episignature (Supplemental Figure 1B). The *ARID1B* duplication of uncertain significance of exon 6-20 (V1i_1.1B) clustered with cases and not with controls. The intragenic quadruplication of exon 6-10 of *ARID1B* (V3i.1B) clustered, as hypothesized beforehand, with BAFopathy samples, supporting its pathogenicity as a loss-of-function variant.

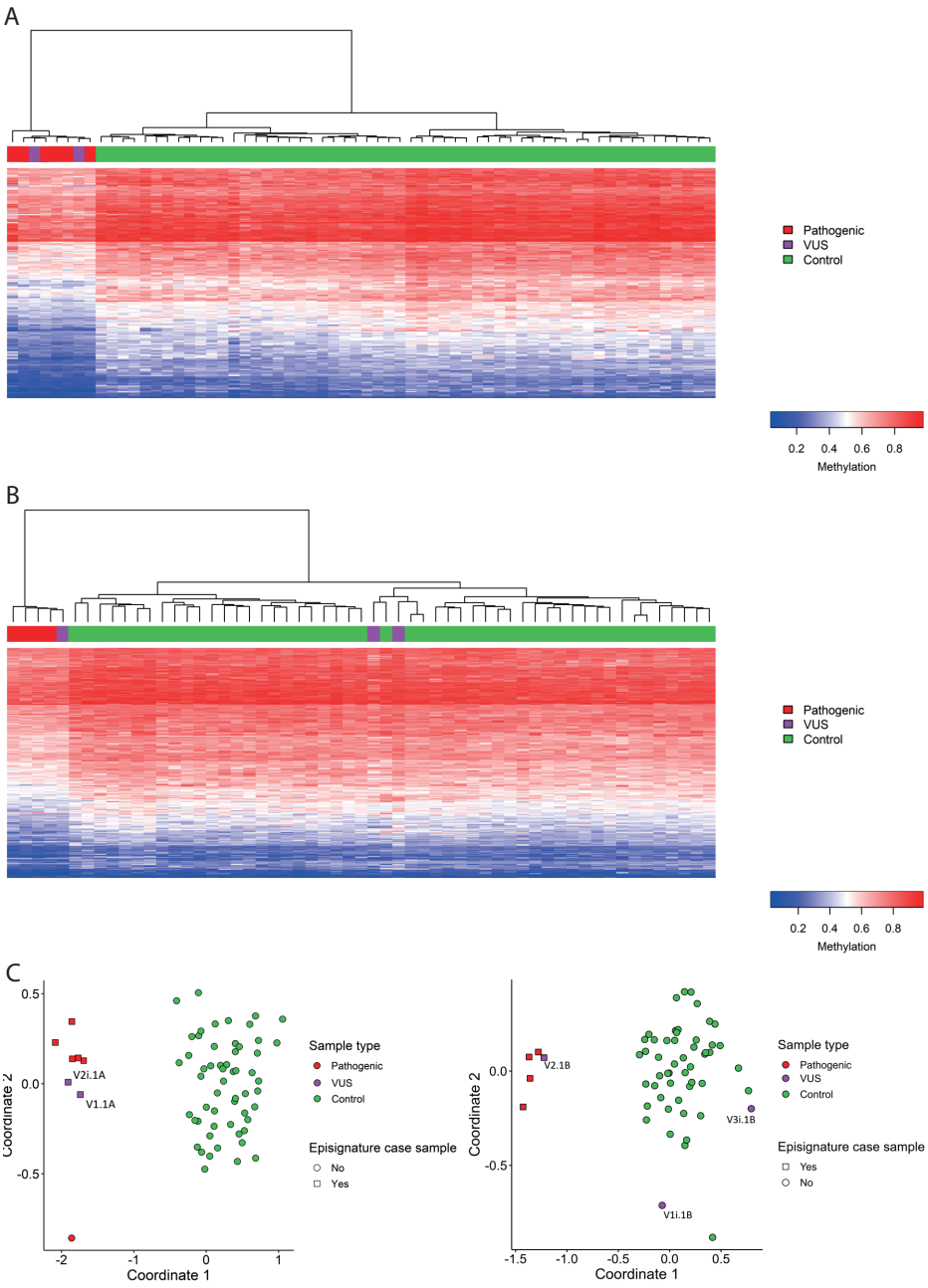


Figure 2. Hierarchical clustering with heatmap of methylation data.

Clustering analysis was performed using Ward's method on Euclidian distance. Rows represent the CpG probes, and columns represent the participants. The color scale from dark blue to dark red represents the range of the methylation levels (beta values) between 0 and 1. The legend represents the pathogenicity interpretation before DNA methylation analysis, i.e., pathogenic

(red), variant of uncertain significance (VUS) (purple), or control (green). Hierarchical clustering of 200 probes differentially methylated between *ARID1A* whole-gene duplication patients and controls generates 2 clusters: one composed of all *ARID1A* patients with pathogenic and the other composed of controls (**A**). A similar pattern is observed using the 219 probes that were identified for the *ARID1B* whole-gene duplication samples; in this case 1 VUS (patient V3i.1B) clusters with controls (**B**). MDS plots showing a similar the differentiation as in the heatmap between cases and controls for *ARID1A* (**C**) and *ARID1B* (**D**) cases.

When clustering patient samples using the *ARID1A* duplication probes, *ARID1A* duplication cases cluster separately from controls and form a separate cluster within the BAFopathy samples. *ARID1B* duplication samples cluster scattered through the controls (except for V3i.1B with a BAFopathy episignature) (Supplemental Figure 1C and D). Interestingly, when using the *ARID1B* duplication probes all *ARID1A* and *ARID1B* duplication cases (except for V3i.1B) cluster together, separate from controls and BAFopathy samples (Supplemental Figure 1E and F). Using the BAFopathy probes, all duplication samples (except V3i.1B) clustered within controls (Supplemental Figure 1G and H).

PhenoScore analysis of variants of uncertain significance

Our study included several duplications of uncertain significance. We used PhenoScore to analyze whether they clustered with ARID1 duplication cases, CSS patients, or NDD control cases based on their HPO terms and/or facial photographs, if available. Case V1i_1.1B, with a partial *ARID1B* duplication of exon 6-20, clustered with CSS cases. Case V2i.1A, with a partial *ARID1A* duplication of exon from halfway exon 1-20, clustered with CSS cases, indicating that the phenotype of this individual is more similar to the phenotypes of those patient groups than of NDD controls. Interestingly, case V3i.1B clustered with all patient groups and not with controls, indicating that the phenotype is more similar to both CSS and ARID1 dup cases compared with NDD controls. Case V1.1.A with an *ARID1A* whole-gene duplication has a high phenotypic similarity score (0.71) only to individuals with *ARID1A* duplications (Supplemental Table 4). Patient V1_1.1B and her father (V1_2.1B) both with an *ARID1B* whole-gene duplication did not cluster with CSS cases, whereas V1_2.1B was seen as more phenotypically similar to *ARID1B* duplication cases. Adding facial photographs did not substantially change these results (Supplemental Table 5).

Interpretation of duplications of uncertain significance

Both *ARID1A* duplications of uncertain significance (ie, V1.1A, V2i.1A) were reclassified as pathogenic *ARID1A* duplications. Two of the 3 unique *ARID1B* duplications of uncertain significance were classified as pathogenic. Patients V2_1.1B and V2_2.1B were regarded as carrying a pathogenic *ARID1B* whole-gene duplication, whereas case V3i.1B (with a

de novo intragenic triplication of *ARID1B*) was considered to have CSS. The inherited duplication of exon 6-20 of *ARID1B* in patients V1i_1.1B and V1i_2.1B was considered to remain a duplication of uncertain significance (Supplemental Table 1).

Phenotype

Clinical data were obtained for 16 cases with a pathogenic *ARID1A* duplication and 13 cases with a pathogenic *ARID1B* duplication (Table 1, Supplemental Figure 2), detailed information was available of 14 *ARID1A* and 11 *ARID1B* cases. In case 3.1B with intrauterine growth restriction (IUGR), severe polyhydramnios, and an *ARID1B* duplication of 6.3 Mb, the pregnancy was terminated. For all other cases, the age at last examination ranged from 1 to 46 years for patients with *ARID1A* duplications and from 1.5 to 31 years for patients with *ARID1B* duplications.

Table 1. Clinical characteristics of patients with *ARID1A* or *ARID1B* whole-gene duplication

Clinical and Developmental Findings	*ARID1A*, *n* = 16[a]			*ARID1B*, *n* = 13[b]			*P* Value	Test
	Affected	Total	%	Affected	Total	%		
Sex, female	7	16	44%	5	13	38%	1.00	χ^2
Neurodevelopmental delay	13	13	100%	7	7	100%		
Intellectual disability	14	14	100%	10	11	91%	.03	Fisher's
Mild	1	14	7%	6	11			
Moderate	6	14	43%	3	11			
Severe	6	14	43%	1	11			
Profound	1	14	7%	0	11			
Language delay	15	15	100%	9	9	100%		
No words	4	12	33%	0	8	0%	.12	Fisher's
Brain anomaly[c]	4	11	36%	1	4	25%	1.00	Fisher's
Behavior anomaly	9	13	69%	6	7	86%	.61	Fisher's
Short attention span	4	13	31%	2	7	29%	1.00	Fisher's
Autistic traits	2	13	15%	3	7	43%	.29	Fisher's
Stereotyped behavior	3	13	23%	1	7	14%	1.00	Fisher's
Epilepsy/seizures	1	11	9%	2	7	29%	.72	Fisher's
Vision impaired	6	11	55%	5	7	71%	.64	Fisher's
Myopia	1	4	25%	3	4	75%	.49	Fisher's
Hypermetropia	2	4	50%	2	3	67%	1.00	Fisher's
Hearing impaired	3	9	33%	1	5	20%	1.00	Fisher's
Pregnancy anomaly[c]	7	13	54%	4	10	40%	.68	Fisher's

Table 1. Continued

Clinical and Developmental Findings	*ARID1A*, *n* = 16[a]			*ARID1B*, *n* = 13[b]			*P* Value	Test
	Affected	Total	%	Affected	Total	%		
Growth delay	10	14	71%	4	10	40%	.21	Fisher's
Birthweight (SDS)	mean −1.4, SDS 1.1		*n* = 12	mean -0.07, SDS 1.1		*n* = 7	.02	*t* test
BW <−2SDS	4	12	33%	0	7	0%	.25	Fisher's
OFC at birth (SDS)	mean −1.6, SDS 0.4		*n* = 3	mean 0.3, SDS 1,6		*n* = 3	.11	*t* test
Height (SDS)	mean −2.2, SDS 1.3		*n* = 12	mean -0.1, SDS 2.1		*n* = 6	<.01	*t* test
Height (<−2 SDS)	8	12	67%	1	6	17%	.13	Fisher's
Weight (SDS − weight for length)	mean 0.8, SDS 2.6		*n* = 11	mean 1.4, SDS 1.6		*n* = 5	.32	*t* test
Weight (>2 SDS − weight for length)	5	11	45%	2	5	40%	1.00	Fisher's
OFC (SDS)	mean −2,7, SDS 1.2		*n* = 12	mean -1.0, SDS 2.4		*n* = 6	.06	*t* test
OFC (<−2 SDS)	11	15	73%	2	6	33%	.15	Fisher's
Ectodermal anomalies								
Nail anomalies	3	12	25%	1	5	20%	1.00	Fisher's
Abnormal dentition	3	9	33%	3	6	50%	.62	Fisher's
Facial characteristics								
Prominent/low columella	3	8	38%	1	4	25%	1.00	Fisher's
Short philtrum	5	10	50%	0	6	0%	.09	Fisher's
Thin upper lip	9	11	82%	3	5	60%	.48	Fisher's
Cleft palate	1	8	13%	2	9	22%	1.00	Fisher's
Small mouth	4	6	67%	0	4	0%	.08	Fisher's
Ear anomalies	5	12	42%	2	6	33%	1.00	Fisher's
Skeletal anomalies								
Hand and foot anomalies	12	12	100%	3	5	60%	.07	Fisher's
Hand anomaly	9	10	90%	2	5	40%	.08	Fisher's
Foot anomaly	7	11	64%	2	5	40%	.60	Fisher's
Pes planus	6	8	75%	1	5	20%	.10	Fisher's
Spinal malformations	2	13	15%	1	9	11%	1.00	Fisher's
Joint laxity	3	7	43%	3	5	60%	1.00	Fisher's
Cardiovascular anomalies[c]	3	12	25%	1	6	17%	1.00	Fisher's
Gastrointestinal anomalies	10	13	77%	3	7	43%	.17	Fisher's
Feeding problems	8	12	67%	1	5	20%	.13	Fisher's
GERD	5	13	38%	0	7	0%	.11	Fisher's
Constipation	6	13	46%	1	7	14%	.33	Fisher's

Table 1 continues on next page

Table 1. Continued

Clinical and Developmental Findings	*ARID1A*, *n* = 16[a]			*ARID1B*, *n* = 13[b]			*P* Value	Test
	Affected	Total	%	Affected	Total	%		
Recurrent airway infections	5	12	42%	0	5	0%	.25	Fisher's
Cryptorchidism	5	6	83%	3	5	60%	.55	Fisher's

IQ, intellectual quotient; *GERD*, gastroesophageal reflux disease; *OFC*, occipital head circumference; *SDS*, standard deviation score.
a 1 patient DECIPHER data only.
b 1 case was a pregnancy termination.
c Reported pregnancy anomalies: polyhydramnios (1× *ARID1A*, 1× *ARID1B*), intrauterine growth retardation (3× *ARID1A*, 2× *ARID1B*), cardiac defect (1× *ARID1A*, 1× *ARID1B*). Reported brain anomalies: T2 hyperintensity and T1 hypointensity in the right thalamus, prominent cerebral cortical sulci, ventricles and extra axial cerebral spinal fluid spaces, cavum septum pellucidum (1× *ARID1A*), ventriculomegaly (3× *ARID1A*), delayed myelination (1× *ARID1A*), hypoplasia of the corpus callosum (1× *ARID1A*), dysmorphic brain particularly affecting the left hippocampus (1× *ARID1B*). Reported cardiovascular anomalies: pulmonary stenosis (1× *ARID1A*). An *ARID1A* patient had a atrium septum defect, muscular small ventricular septum defect (VSD), and mild-moderate valvular pulmonary stenosis (PS), concluded as a combination of small cardiac abnormalities without hemodynamic consequences. He also had an additional anatomical variation in division of arch vessels (left arch joint origin of brachiocephalic trunk and left carotid artery). The VSD closed spontaneously. PS was not important. Other reported anomalies in patients were VSD (2× *ARID1A*), and cardiac rhythm anomaly (1× *ARID1B*).

ARID1A and ARID1B

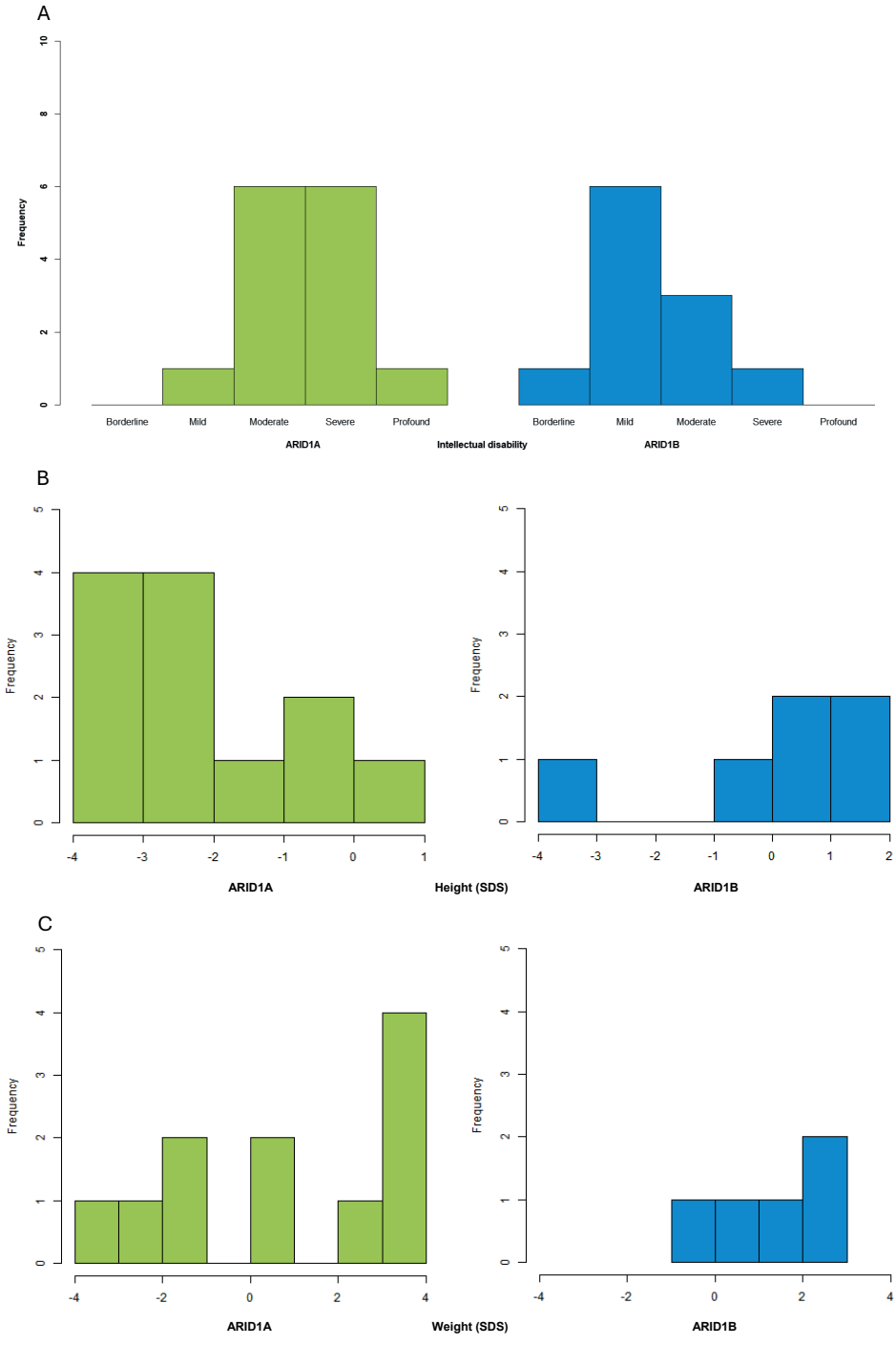
ARID1A duplication patients exhibited more severe developmental delays and intellectual disability (ID) compared with *ARID1B* duplication patients, as evidenced by their developmental milestones and IQ scores (Figure 3, Table 1). Head circumference tended to be smaller in *ARID1A* duplication patients compared with *ARID1B* duplication patients (Figure 3D, Table 2). Feeding difficulties were more common among *ARID1A* patients, with a prevalence of 67% (8/12) compared with 20% (1/5) in *ARID1B* patients.

Duplication size

After correcting for the affected gene, the analyses comparing patients with smaller duplications (*ARID1A*: *n* = 8, *ARID1B*: *n* = 6) to those with larger duplications (more than 5 genes) revealed that there is no difference between patients with duplications including more than 5 genes compared with patients with 5 or less genes in the duplicated region in motor and speech development, congenital anomalies or seizures (Supplemental Figure 3).

Shared features

Patients with an *ARID1A* or *ARID1B* duplication share several specific features, including intellectual disability (14/14 *ARID1A*, 10/11 *ARID1B*), growth delay (10/14 *ARID1A*, 4/10 *ARID1B*), cryptorchidism (*n* = 5 *ARID1A*, *n* = 3 *ARID1B*), and pes planus (6/8 *ARID1A*, 1/5 *ARID1B*) (Table 2).


A
Frequency
0
2
4
6
8
10
Borderline
Mild
Moderate
Severe
Profound
ARID1A
Intellectual disability
ARID1B
B
Frequency
0
1
2
3
4
5
-4
-3
-2
-1
0
1
2
ARID1A
Height (SDS)
ARID1B
C
Frequency
0
1
2
3
4
5
-4
-2
0
2
4
ARID1A
Weight (SDS)
ARID1B

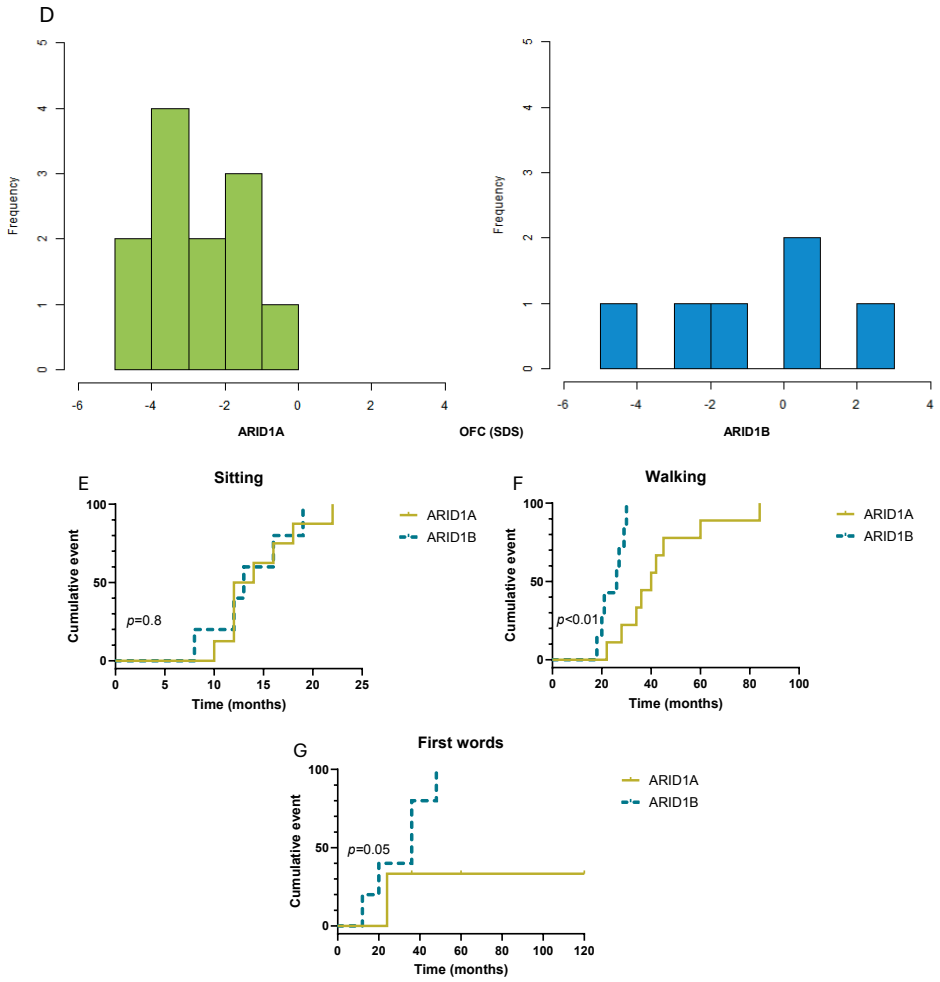


Figure 3. Intellectual disability degree and developmental milestones of the *ARID1A* and *ARID1B* duplication patients.

Histogram of the (**A**) distribution of intellectual disability severity, (**B**) height (SDS), (**C**) weight (SDS weight-for-height) and, (**D**) head circumference (SDS). Cumulative distribution of developmental milestones (**E**) sitting (*n* = 13), (**F**) walking (*n* = 15), and (**G**) first words (*n* = 11).

PhenoScore analysis

The PhenoScore analysis indicated that *ARID1A* duplication patients more often tend to have retractile testis, overweight, low birth weight, short stature, and microcephaly compared with controls (Supplemental Figure 4A and B). *ARID1B* duplication patients tend to have more motor delay, retractile testes, mild-moderate ID, and less joint hypermobility compared with controls (Supplemental Figure 4C and D). Combining the data from *ARID1A* and *ARID1B* duplication patients and comparing them with controls

produces results comparable to those obtained from separate analyses (Supplemental Figure 4E and F). Comparing *ARID1A* with *ARID1B* duplication patients, *ARID1A* duplication patients more frequently tend to have constipation, IUGR, reflux, and pes planus, whereas *ARID1B* patients more frequently tend to have mild-moderate ID, myopia, and retractile testis (Supplemental Figure 4G). The PhenoScore analysis also revealed that, based on HPO terms, *ARID1A* and *ARID1B* duplication patients clustered just as far from each other as from NDD controls (PhenoScore 0.50) or CSS patients (PhenoScore of 0.52) (Supplemental Table 3).

DISCUSSION

In total, we report data from 16 patients with a pathogenic *ARID1A* and 13 patients with a pathogenic *ARID1B* gene duplication.

Genotype-phenotype

Our study reveals a distinct clinical phenotype resulting from *ARID1B* whole-gene duplications, which shares features with the phenotype of *ARID1A* whole-gene duplications, including mild to severe intellectual disability, growth delay, and hand or feet anomalies, occasionally accompanied by a unique facial appearance.

Genes associated with the BAF complex have been linked to CSS, with each gene having its own genotype-phenotype association within the CSS-spectrum.[27,28] Patients with loss of function variants in *ARID1A* and *ARID1B* have clinical features most similar to each other within this spectrum, with *ARID1A* patients generally being more severely affected compared with those with pathogenic variants in *ARID1B*.[6,7]29, 30, 31 For duplications, we now also show that *ARID1A* duplications lead to a more severe phenotype than *ARID1B* duplications, considering that growth and ID-associated features differed significantly (Figure 3, Table 1). However, the PhenoScore analysis (Supplemental Figure 4, Supplemental Table 3), DNA methylation (Figure 2, Supplemental Figure 1), and facial photographs (Supplemental Figure 2) suggest that the duplications of *ARID1A* and *ARID1B* are not merely differing in severity, but should be considered different clinical entities.

CSS patients compared with whole-gene duplication patients

Loss-of-function variants, including deletions in *ARID1A* and *ARID1B*, result in CSS. CSS is characterized by ID, feeding difficulties, congenital anomalies, hypoplasia of the fifth digits and/or nails, hypertrichosis, and distinctive facial features. Nonspecific features, such as ID, feeding difficulties, and congenital anomalies, were also present

in our patients with whole-gene duplications. However, CSS-specific features, such as shirt fifth digit, hypoplasia of the nail of a fifth digit, hypertrichosis, and the distinct facial features of CSS, were not observed. This is reflected by the PhenoScore analysis showing that *ARID1A* and *ARID1B* patients with loss-of-function variants including deletions cluster separately from *ARID1A* and *ARID1B* whole-gene duplication patients (Supplemental Table 3).

DNA methylation

Over the last several years, researchers have discovered a growing number of unique DNA methylation patterns associated with specific disorders, referred to as episignatures.[20] These episignatures have become valuable biomarkers in the diagnostic testing for these disorders,[17,32] many of which are linked to chromatin remodeling genes, including genes that encode components of the BAF complex. In our study, we utilized DNA methylation analysis to explore the impact of *ARID1A* and *ARID1B* whole-gene duplications on these episignatures. Interestingly, we found a specific *ARID1A* duplication DNA episignature and the determined the presence of an *ARID1B* duplication episignature. It has been discussed that an intermediate clustering of a sample or episignature may reflect an intermediate phenotype.[33] It could therefore be that the inconclusive scoring of the DNA methylation data of *ARID1B* duplication patients is due to their milder phenotype. Using the *ARID1B* duplication probes the *ARID1A* and *ARID1B* duplication samples do appear to cluster together and separate from controls. The duplication episignatures differ from both controls and the episignatures of patients with other BAFopathies such as CSS/NCBRS, SMARCA2 BIS, SMARCA4, ARID2, SOX11, and others. All of this provides further evidence for the pathogenicity of *ARID1B* whole-gene duplications and suggests that *ARID1A* and *ARID1B* duplications cause a distinct BAFopathy.

The identification of a distinct DNA methylation pattern, or episignature, in patients with *ARID1A* whole-gene duplications provides clinicians with an additional tool to aid the interpretation of these duplications. For example, in case V2i.1A, the *ARID1A* exon 1 to exon 20 duplication, the *ARID1A* duplication DNA methylation pattern confirmed the pathogenicity of this variant. Whereas PhenoScore analysis showed an inconclusive clustering result. PhenoScore works for syndromic cases, a possible explanation could be that the features identified in our *ARID1A* duplication patients are too general (ID, growth delay, and microcephaly) or that there are too few patients in the discovery cohort. On the other hand, cases V1.1A and V2_2.1B, who had a *ARID1A* or *ARID1B* whole-gene duplication, who were classified as duplications of uncertain significance before DNA methylation results, tended to cluster with *ARID1A* or *ARID1B* duplication cases according to the PhenoScore analysis. After the final DNA methylation results

these samples were reclassified as (likely) pathogenic. In our first attempts to identify an ARID1 duplication episignature using DNA methylation data from 8 *ARID1A* and 5 *ARID1B* patients, several *ARID1B* cases clustered with *ARID1A* cases, showing the overlap between the 2 episignatures (Figure 2, Supplemental Figure 1). Later findings unveiled distinctions in the DNA methylation patterns of patients with *ARID1A* and *ARID1B* duplications. This underscores the potential uncertainty associated with episignatures. Therefore, it is crucial to acknowledge that, although the presence of an episignature strongly indicates pathogenicity, its absence, although suggestive, does not definitively exclude it.

Similarly, PhenoScore can be useful in interpreting genetic variants with uncertain significance by providing further confirmation of their pathogenicity or nonpathogenicity. If a variant has a low suspicion of pathogenicity and lacks a DNA methylation episignature or clustering with cases, PhenoScore can offer additional evidence that the variant is nonpathogenic. Alternatively, PhenoScore can provide evidence for its pathogenicity, if a patient with such a variant clusters with cases (eg, V1.1A). However, it should be noted that PhenoScore, just like the *ARID1B* duplication episignature, is currently only used for research purposes, and its reliability needs to be further evaluated in clinical studies.

Limitations

Our study has several limitations. First, our data were contributed by multiple clinicians and may not be complete in all cases. In addition, our cohort is predominantly composed of young patients, potentially leading to an underestimation of features that develop with age. Although we performed survival analyses to correct for this, further research including older adults is needed to gain a more complete understanding of the spectrum of features associated with advancing age.

Table 2. Recommendations surveillance guidelines *ARID1A/ARID1B* duplication patients

Evaluation (Inquire/Perform Physical Examination)	After Diagnosis	Age Category			(Para)Medic[a]
		In Children (0-18 Years)	In Adults (18+)	Frequency	
Congenital abnormalities (heart, kidneys, and cryptorchidism)	Yes	Upon indication		Upon indication	treating physician
Feeding difficulties	Yes	Yes		At every visit	pediatrician
Constipation	Yes	Yes	Yes	At every visit	treating physician
Growth and weight	Yes	Yes	Yes	At every visit, to avoid excessive weight gain	pediatrician
Seizures	Yes	Yes	Yes	At every visit	treating physician
Endocrine/hormonal	On indication	Upon indication	Upon indication	Upon indicationb	treating physician
Vision	Yes	Yes	Yes	Every 2 years during childhood, frequency can be adjusted if the patient is able to report on eyesight. In adults evaluation at age 30 and from age 40 every 5 years	ophthalmologist
Hearing	Yes	Yes	Upon indication	At every visit ask for signs of hearing loss and refer to audiologist if suspected	treating physician
Scoliosis	Yes	Yes		Periodically, until length growth is complete	orthopedic surgeon/ physiatrist/ pediatrician
Pedes (plano)valgi	Yes	Yes	On indication	Periodically	orthopedic surgeon/ physiatrist/ pediatrician
Dentition/dental health	Yes	Yes	Yes	Twice a year	dentist
Motor development	If aged <16	Yes		At every visit	pediatrician/ physiatrist/ physical therapist
Cognitive development	If aged <16	Yes		At every visit	pediatrician/ psychologist
Communication/language development	Yes	Yes		Screening at age 2 and 3 years, then on indication	pediatrician/ physiatrist/ speech therapist

Table 2. Continued

Evaluation (Inquire/Perform Physical Examination)	After Diagnosis	Age Category			(Para)Medic[a]
		In Children (0-18 Years)	In Adults (18+)	Frequency	
Behavioral and social development/ impairments	Yes	Yes	Yes	At every visit	treating physician
Sexual development	When applicable	Yes	Yes	When applicable	treating physician
Transition to adult care	When applicable	When applicable		At least from the age of 16 years	physiatrist/ pediatrician

a Treating physician (eg, general practitioner, pediatrician).
b Especially, glucose, and thyroid lab.

CONCLUSION

Our data reveal a distinct type of BAFopathy entity, characterized by a whole-gene duplication of *ARID1A* or *ARID1B*, rather than haploinsufficiency. The overlap in clinical phenotypes and episignatures between *ARID1B* and *ARID1A* whole-gene duplications supports the distinct nature of this BAFopathy.

Based on the clinical data of our cohort, we are suggesting the screening recommendations as described in Table 2.

DATA AVAILABILITY

Deidentified patient data will be made available on request to the corresponding author.

CONFLICT OF INTEREST

Bekim Sadikovic is a shareholder in EpiSign Inc. All other authors declare no conflicts of interest.

ACKNOWLEDGMENTS

The authors would like to thank Johan den Dunnen for his assistance in adding patient data to the Leiden Open Variant Database.

This study makes use of data generated by the DECIPHER community. A full list of centers who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. Funding for the DECIPHER project was provided by Wellcome.

Results from this article have been previously presented at the 55th European Society of Human Genetics (ESHG) Conference 2022.

FUNDING

The authors are grateful to the Dutch Organisation for Health Research and Development: ZON-MW grants 912-12-109 (to B.B.A.d.V.) and Donders Junior researcher grant 2019 (B.B.A.d.V.). Part of this research was also funded by the Ministero della Salute (Current Research Funds and PNRR-MR1-2022-12376811, to M.A.T.).

ETHICS DECLARATION

The Institutional Review Board of Leiden University Medical Center, Leiden, The Netherlands provided approval waivers (no: P17.301 and G21.129). Clinical data were deidentified by allocating patient numbers. Where possible, informed consent was obtained by the referring clinician.

Written consent was obtained and archived for all included patient photos.

ADDITIONAL INFORMATION

Supplementary Material.

Supplementary Information

Supplementary Table 1.
Supplementary Data

Supplementary Table 2.
Supplementary Data

REFERENCES

1. J. Hoyer, A.B. Ekici, S. Endele, *et al.* Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability.Am J Hum Genet, 90 (3) (2012), pp. 565-572, 10.1016/j.ajhg.2012.02.007
2. C.F. Wright, T.W. Fitzgerald, W.D. Jones, *et al.* Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet, 385 (9975) (2015), pp. 1305-1314, 10.1016/S0140-6736(14)61705-0
3. Y. Tsurusaki, N. Okamoto, H. Ohashi, *et al.* Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet, 44 (4) (2012), pp. 376-378, 10.1038/ng.2219
4. Y. Tsurusaki, E. Koshimizu, H. Ohashi, *et al.*De novo SOX11 mutations cause Coffin-Siris syndrome. Nat Commun, 5 (2014), p. 4011, 10.1038/ncomms5011
5. G.W.E. Santen, E. Aten, Y. Sun, *et al.*Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. Nat Genet, 44 (4) (2012), pp. 379-380, 10.1038/ng.2217
6. G.W.E. Santen, E. Aten, A.T. Vulto-van Silfhout, *et al.* Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum Mutat, 34 (11) (2013), pp. 1519-1528, 10.1002/humu.22394
7. D. Wieczorek, N. Bögershausen, F. Beleggia, *et al.* A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum Mol Genet, 22 (25) (2013), pp. 5121-5135, 10.1093/hmg/ddt366
8. E. Aref-Eshghi, E.G. Bend, R.L. Hood, *et al.* BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin-Siris and Nicolaides-Baraitser syndromes. Nat Commun, 9 (1) (2018), p. 4885, 10.1038/s41467-018-07193-y
9. R.L. Collins, J.T. Glessner, E. Porcu, *et al.* A cross-disorder dosage sensitivity map of the human genome. Cell, 185 (16) (2022), pp. 3041-3055.e25, 10.1016/j.cell.2022.06.036
10. M. Bidart, M. El Atifi, S. Miladi, *et al.* Microduplication of the ARID1A gene causes intellectual disability with recognizable syndromic features. Genet Med, 19 (6) (2017), pp. 701-710, 10.1038/gim.2016.180
11. C. Coutton, M. Bidart, J. Rendu, *et al.* 190-kb duplication in 1p36.11 including PIGV and ARID1A genes in a girl with intellectual disability and hexadactyly. Clin Genet, 84 (6) (2013), pp. 596-599, 10.1111/cge.12113
12. D.A. Dyment, A. O'Donnell-Luria, P.B. Agrawal, *et al.* Alternative genomic diagnoses for individuals with a clinical diagnosis of Dubowitz syndrome. Am J Med Genet A, 185 (1) (2021), pp. 119-133, 10.1002/ajmg.a.61926
13. R. van der Donk, S. Jansen, J.H.M. Schuurs-Hoeijmakers, *et al.* Next-generation phenotyping using computer vision algorithms in rare genomic neurodevelopmental disorders. Genet Med, 21 (8) (2019), pp. 1719-1725, 10.1038/s41436-018-0404-y
14. H.V. Firth, S.M. Richards, A.P. Bevan, *et al.* DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. Am J Hum Genet, 84 (4) (2009), pp. 524-533, 10.1016/j.ajhg.2009.03.010
15. E. Aref-Eshghi, J. Kerkhof, V.P. Pedro, *et al.* Evaluation of DNA methylation episignatures for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. Am J Hum Genet, 106 (3) (2020), pp. 356-370, 10.1016/j.ajhg.2020.01.019
16. B. Sadikovic, M.A. Levy, J. Kerkhof, *et al.* Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. Genet Med, 23 (6) (2021), pp. 1065-1074, 10.1038/s41436-020-01096-4

17. M.A. Levy, H. McConkey, J. Kerkhof, *et al.* Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. HGG Adv, 3 (1) (2022), Article 100075, 10.1016/j.xhgg.2021.100075
18. W. Zhou, T.J. Triche Jr., P.W. Laird, H. Shen. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. Nucleic Acids Res, 46 (20) (2018), Article e123, 10.1093/nar/gky691
19. R. Pidsley, E. Zotenko, T.J. Peters, *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol, 17 (1) (2016), p. 208, 10.1186/s13059-016-1066-1
20. M.A. Levy, R. Relator, H. McConkey, *et al.* Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders. Hum Mutat, 43 (11) (2022), pp. 1609-1628, 10.1002/humu.24446
21. A.J.M. Dingemans, M. Hinne, K.M.G. Truijen, *et al.* PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. Nat Genet, 55 (9) (2023), pp. 1598-1607, 10.1038/s41588-023-01469-w
22. E.R. Riggs, E.F. Andersen, A.M. Cherry, *et al.* Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med, 22 (2) (2020), pp. 245-257, 10.1038/s41436-019-0686-8
23. F. Yue, Y. Yu, Q. Xi, *et al.* Prenatal diagnosis of a 5q35.3 microduplication involving part of the ADAMTS2 locus: a likely benign variant without apparent phenotypic abnormality: case series. Medicine (Baltimore), 98 (49) (2019), Article e18258, 10.1097/MD.0000000000018258
24. L. Te Weehi, R. Maikoo, A. Mc Cormack, *et al.* Microduplication of 3p26.3 implicated in cognitive development. Case Rep Genet, 2014 (2014), Article 295359, 10.1155/2014/295359
25. C. Li, C. Liu, B. Zhou, C. Hu, X. Xu. Novel microduplication of CHL1 gene in a patient with autism spectrum disorder: a case report and a brief literature review. Mol Cytogenet, 9 (2016), p. 51, 10.1186/s13039-016-0261-9
26. K.V. Onate-Quiroz, B.U. Nwosu, P. Salemi. Novel duplication of the cell adhesion molecule L1-like gene in an individual with cognitive impairment, tall stature, and obesity: a case report. Front Neurol, 14 (2023), Article 1104649, 10.3389/fneur.2023.1104649
27. T. Kosho, N. Okamoto, Coffin-Siris Syndrome International Collaborators. Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. Am J Med Genet C Semin Med Genet, 166C (3) (2014), pp. 262-275, 10.1002/ajmg.c.31407
28. S. Schrier Vergano, G. Santen, D. Wieczorek, B. Wollnik, N. Matsumoto, M.A. Deardorff, R.A. Pagon, M.P. Adam, H.H. Ardinger, *et al.* (Eds.), Coffin-Siris syndrome, GeneReviews(R). University of Washington (1993), pp. 1-20 https://www.ncbi.nlm.nih.gov/books/NBK131811/, Accessed 12th Jan 2024

29. Y. Tsurusaki, N. Okamoto, H. Ohashi, *et al.*Coffin-Siris syndrome is a SWI/SNF complex disorder. Clin Genet, 85 (6) (2014), pp. 548-554, 10.1111/cge.12225
30. P.J. van der Sluijs, M. Alders, A.J.M. Dingemans, *et al.*A case series of familial ARID1B variants illustrating variable expression and suggestions to update the ACMG criteria. Genes, 12 (8) (2021), p. 1275, 10.3390/genes12081275
31. P.J. van der Sluijs, M. Joosten, C. Alby, *et al.*Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort. Genet Med, 24 (8) (2022), pp. 1753-1760, 10.1016/j.gim.2022.04.010
32. J. Kerkhof, C. Rastin, M.A. Levy, *et al.* Diagnostic utility and reporting recommendations for clinical DNA methylation episignature testing in genetically undiagnosed rare diseases. Genet Med, 26 (5) (2024), 10.1016/j.gim.2024.101075
33. Z. Awamleh, S. Goodman, S. Choufani, R. Weksberg. DNA methylation signatures for chromatinopathies: current challenges and future applications. Hum Genet, 143 (4) (2024), pp. 551-557, 10.1007/s00439-023-02544-2

Part II

Methods for studying mechanisms of disease and possible treatment

Chapter 7

Developing an *in vitro* model system to study the effects of Coffin-Siris Syndrome pathogenic variants on neuronal differentiation

Pleuntje J van der Sluijs[1], Kelly K D Vonk[2], Serkan Dogan[2], Maja Vukic[2], Lucy J Sinke[2], Christian Freund[3], Bastiaan T Heijmans[4], Gijs W E Santen[1], Lucia Daxinger[2]

Affiliations:
1. Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands
2. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands
3. LUMC hiPSC Hotel, Dept. Anatomy & Embryology, Leiden University Medical Center, Leiden, The Netherlands.
4. Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

Manuscript in preparation

ABSTRACT

Coffin-Siris syndrome (CSS) is caused by pathogenic variants in genes encoding the BAF chromatin remodelling -complex. The exact pathophysiological mechanisms underlying CSS are unknown. By using patient derived induced pluripotent stem cells (iPSCs) carrying pathogenic variants in *ARID1B* or *SMARCB1* and differentiation towards neuronal progenitor cells (NPCs), we established an in vitro model system to study variant effects on neuronal differentiation and on the iPSC and NPC epigenetic landscapes. Using RT-qPCR, IF-staining and DNA methylation analyses we show suitability of our cell model for studying CSS variant effects on neuronal differentiation.

INTRODUCTION

Coffin- Siris syndrome (CSS) (OMIM 135900) is characterized by neurodevelopmental delay (NDD), hypotonia, corpus callosum agenesis, distinctive facial features and aplasia or hypoplasia of the fifth digit and/or nails[1]. Pathogenic variants in genes encoding members of the canonical BAF chromatin remodelling complex (SWI/SNF complex), which consists of 15 subunits. Pathogenic variants in seven subunits have been identified as underlying the disease in most CSS patients[2-8]. Pathogenic variants in *ARID1B* are identified in 50-75% of CSS patients[4-6,8].. In addition, 7-12% of CSS patients carry pathogenic variants in *SMARCB1*[4-6,8]. Within the CSS spectrum, *ARID1B* patients generally exhibit a milder phenotype compared to those with pathogenic variants in *SMARCB1*[9]. Like many other developmental disorders[10], CSS patients exhibit a specific DNA methylation pattern in their blood, different from controls and suitable for diagnostics[10].

Understanding the pathophysiology behind the CSS phenotype is crucial for both comprehending the disease and identifying potential treatment targets. From studies performed in mice and human cell models, there are already some indications that specific neuronal processes are affected. For example, in ARID1B haploinsufficient mice[11], fewer inhibitory GABAergic interneurons and increased apoptosis and decreased proliferation were reported. More recently, through the use of an organoid model, it was shown that *ARID1B* haploinsufficient variants reduce chromatin accessibility at specific transcription factor binding sites and downregulate transcriptional programs of the corpus callosum[12]. Furthermore, Pagliaroli *et al.*[13], used an iPSC to neural crest cell differentiation model and found that *ARID1B* is required for exit from pluripotency and that this process is impaired in patient-derived iPSCs. However, despite the publication of several in vivo[11,14-19] and in vitro studies[13,20-23], the exact pathophysiological mechanisms underlying CSS remain enigmatic.

Here, to investigate the pathophysiology of CSS we created a cell model system that focuses on genetic variant effects in neuronal differentiation using two *ARID1B* (mild end of CSS) and three *SMARCB1* (severe end of CSS) patient iPSC lines (Figure 1). This system is designed to explore the molecular aspects of CSS and aims to shed light on how CSS pathogenic variants influence iPSC to neuronal differentiation and the transcriptional and epigenetic processes required for this process.

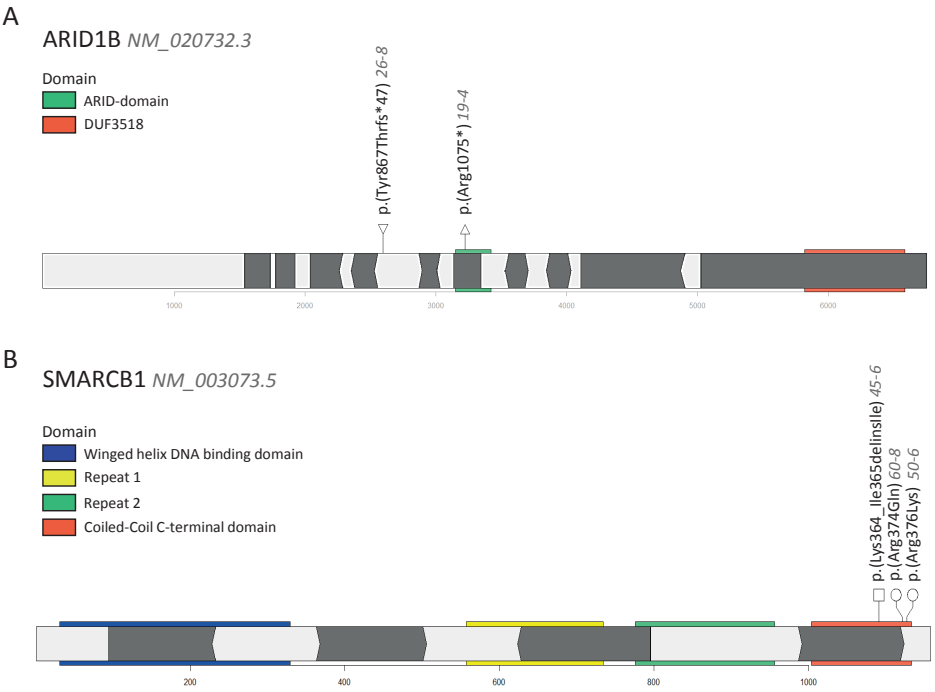


Figure 1: **a**-**b**) Illustration of the *ARID1B* loss of function variants and *SMARCB1* missense variants and inframe deletion of studied patients. The numbers in the gene model refer to transcripts NM_020732.3 (*ARID1B*) and NM_003073.5 (*SMARCB1*).

METHODS

Human iPSC generation and culture

Control iPSC lines (114-1 and 114-2) were derived from fibroblasts obtained from unaffected siblings of spinocerebellar ataxia patients[24]. Line 4-10 was generated from fibroblasts of a healthy 40-year-old control, and line 64-9 was derived from fibroblasts of a mosaic *ARID1A* patient. Skin fibroblasts were obtained from five CSS patients, two patients with a pathogenic variant in *ARID1B*, and three patients with a pathogenic variant in *SMARCB1* (Figure 1a-b). Line 4-10, 64-9, and all patient lines (*ARID1B* patients: 19-4, 26-8; *SMARCB1* patients: 50-6, 45-6, 60-8) were established at the LUMC hiPSC Hotel using the polycistronic lentiviral vector LV.RRL.PPT.SF.hOKSM.idTomato.-preFRT as described elsewhere[25,26].

Informed consent was obtained. The study was conducted in accordance with the criteria set by the Declaration of Helsinki. The protocol to make iPSCs was approved by the institutional review board (IRB) at LUMC.

For each line multiple clones were generated. Their pluripotency was assessed through immunofluorescence microscopy using antibodies against NANOG, OCT3/4, SSEA4, and Tra-1-81. After spontaneous differentiation into three germ layers antibodies against TUBB3, AFP and CD31 were used, as described by Warlich *et al*[26]. Clones exhibiting pluripotent characteristics were selected for downstream applications. Karyotyping by G binding was performed, and short tandem repeat (STR) profiling was carried out. Expansion of iPSC lines occurred in feeder-free, serum-free mTeSR™1 medium (STEMCELL Technologies), with passaging performed at a ratio of approximately 1:10 at 80% confluency using ReLeSR (STEMCELL Technologies). Small cell clusters (50–200 cells) were then plated on tissue culture dishes coated overnight with Corning® Matrigel® matrix (Corning) or Geltrex™ LDEV-Free hESC-qualified Reduced Growth Factor Basement Membrane Matrix (Fisher-Scientific).

Differentiation into medial ganglionic eminence (MGE) progenitors/neuronal progenitor cells (NPC)

iPSC lines were differentiated towards medial ganglionic eminence progenitors (MGE) or NPCs using an adaption of the protocol to produce GABAergic interneurons by Liu *et al*[27] (Figure 2). In brief, iPSCs were dissociated with accutase and suspended in STEMdiff neural induction medium (NIM), supplemented with SMADi and Y-27632. iPSCs were transferred to an AggreWell800 plate (all STEMCELL Technologies, Vancouver, Canada), with daily refreshing of NIM. After 7 days, embryoid bodies were moved to a poly-D-lysine (PDL)/laminin-coated 6-well plate. NIM was changed every other day until day 10, when neural rosettes appear. On day 10, 1.5 µM Pur was added to the medium, with refreshing of NIM+Pur every other day until day 16. On day 16, the differentiating colonies contain neural tube–like rosettes formed by multiple layers of columnar epithelia surrounded by a ring of flat cells.

Differentiation was performed twice. On day 12 DNA and RNA was harvested and cells were fixed for immunohistochemistry. During the second differentiation, neural tube-like rosettes were harvested on day 16 for DNA and RNA.

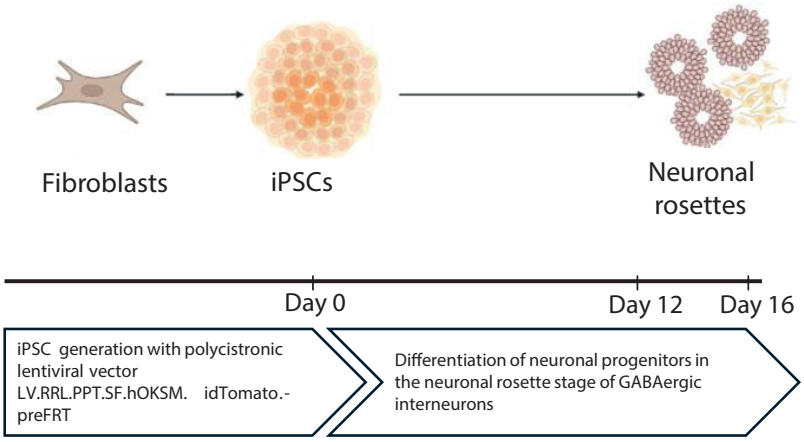


Figure 2: Schematic representation Neural differentiation of patient-derived induced pluripotent stem cells (iPSCs) to neural progenitor cells

Immunofluorenscence (IF)

Cells were fixed with 4% paraformaldehyde (PFA) for 30 min at room temperature. Cells were pre-incubated for 10 minutes at room temperature in 0.1-1.0% Triton X-100 in Dulbecco's phosphate-buffered saline (dPBS) and subsequently blocked for 30 minutes with 1% bovine serum albimun (BSA) in dPBS. Primary antibodies were diluted in 1% BSA in dPBS and samples were incubated in a humidified chamber overnight at 4°C. Secondary antibodies were diluted in 1%BSA in dPBS and samples were incubated for 2 hours at room temperature. Finally, samples were covered with Everbrite hardset containing 4', 6-diamidino-2-phenylindole (DAPI) (Biotum, USA). Antibodies are reported in Supplementary Table 1.

Real-time quantitative polymerase chain reaction (RT-qPCR)

Total RNA was isolated using QIAzol (5346994; Qiagen), with approximately 1 mg of total RNA utilized for reverse transcription via the RevertAid H Minus First Strand cDNA Synthesis Kit (K1632; Thermo). RT-qPCR was performed in triplicate on a C1000TM Thermal cycler (Bio-Rad) with SYBR Green (170-8887; Bio-Rad). Expression data were normalized to GUS. Primer sequences are reported in Supplementary Table 2.

DNA methylation analyses

Omixer[28] was used to randomize DNA samples to optimize sample distribution across batches. Raw DNA methylation data was obtained from Illumina HumanMethylationEPIC arrays (array type: EPICv2). The analyses were conducted in R version 4.3.2 using the following R packages: ChAMP, RPMM, Minfi, ggplot2, DNAmArray, irlba, pheatmap, wateRmelon, ggfortify, shinyMethyl, IlluminaHumanMethylation450kanno.ilmn12.hg19,

GenomicRanges package, ComplexHeatmap, and limma. Data processing and analysis were performed using Python version 3.8.5, and the following Python libraries: Pandas (version 1.2.4), NumPy (version 1.19.5), Scikit-learn (version 0.24.2), Matplotlib (version 3.4.3) and Seaborn (version 0.11.2).

RESULTS

Characterization of *ARID1B* and *SMARCB1* patient-derived iPSC lines

To create an *in vitro* model system focusing on variant effects in neuronal differentiation in CSS we generated iPSC lines from fibroblasts of two patients with pathogenic variants in *ARID1B*, three patients with pathogenic variants in *SMARCB1* (Figure 1a-b) and four control iPSC lines. All iPSC lines expressed the pluripotency marker genes *OCT4* and *NANOG* and there was no statistical significance in mRNA levels between controls and patients when measured by RT-qPCR (Figure 3a-b). Furthermore, control and patient-derived iPSCs showed a similar morphology. In accordance with the RT-qPCR data, immunofluorescence (IF)-staining showed expression of the pluripotency marker OCT4 (Figure 3c).

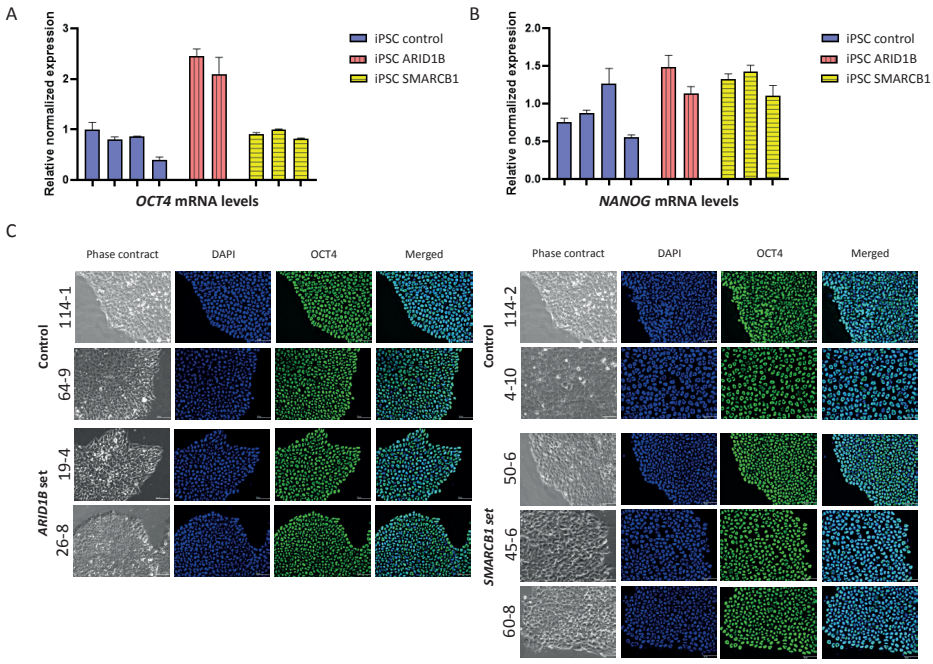


Figure 3: **a**-**b**) mRNA levels of *OCT4* and *NANOG*. **c**) IF-staining of OCT4 and DAPI.

Since increased apoptosis of neuronal progenitors was observed in ARID1B haploinsufficient mice[11,29], we investigated whether increased apoptosis is present in patient-derived iPSC using IF-staining of cleaved caspase-3. We found a similar cleaved caspase-3 staining in patient and control lines (Figure 4), suggesting that neither *ARID1B* nor *SMARCB1* CSS pathogenic variants lead to increased apoptosis of iPSCs.

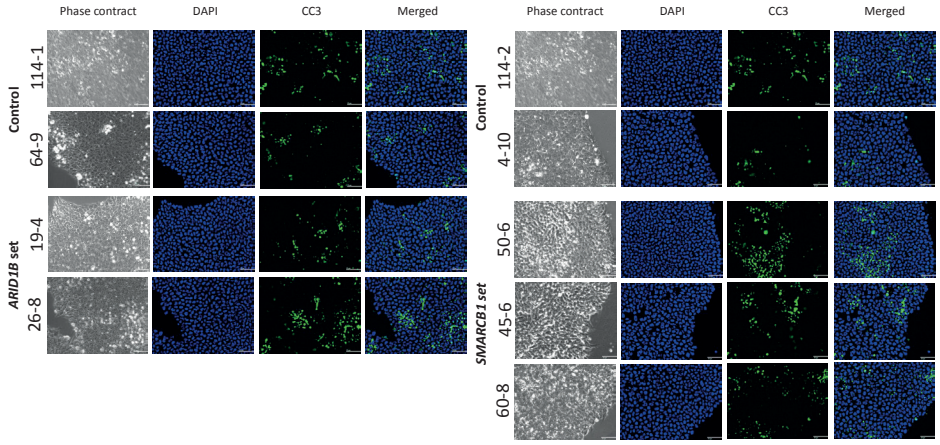


Figure 4: IF-staining of cleaved caspase-3 (CC3) and DAPI in iPSCs.

Next, we addressed the effect of the *ARID1B* and *SMARCB1* pathogenic variants on mRNA levels. The two *ARID1B* patient lines both have pathogenic variants that lead to premature stop codons (Figure 1a), and *ARID1B* haploinsufficiency[13]. Using RT-qPCR, we found that *ARID1B* mRNA levels in the *ARID1B* patient lines were slightly reduced, but not statistically significantly lower, when compared to control or *SMARCB1* patient lines (Figure 5a), most likely due to incomplete nonsense-mediated decay. The three *SMARCB1* pathogenic variants are all nontruncating variants (i.e. missense variants or in frame deletions) (Figure 1b). Consistent with this, *SMARCB1* mRNA levels were similar in *SMARCB1* patient when compared to control lines by RT-qPCR (Figure 5b). Together, these results show that iPSCs derived from patients with a pathogenic variant in either *ARID1B* or *SMARCB1* are pluripotent and phenotypically similar to iPSCs derived from healthy controls.

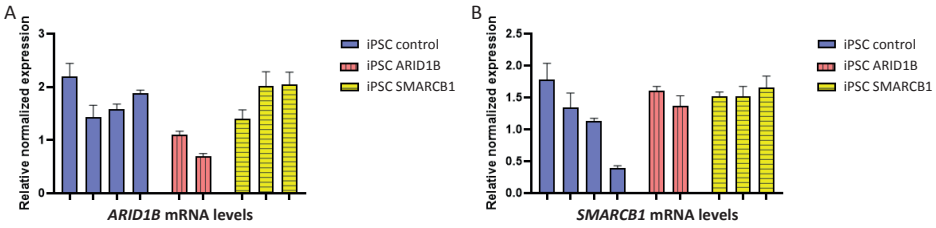


Figure 5: **a**-**b**) mRNA levels of *ARID1B* and *SMARCB1* in iPSCs.

A cellular model system to study CSS variant effects on neuronal differentiation

Next, we differentiated control and patient iPSC lines towards early neuronal progenitor cells (NPCs) and subsequently to neuronal rosettes, using an adapted protocol previously published by Liu *et al*[27] (Figure 2). Differentiation was performed twice.

In the first experiment, iPSCs were differentiated for 12 days towards NPCs. This is an intermediary timepoint, which was included to evaluate the differentiation process and to measure apoptosis levels in the various cell lines. The two *ARID1B* and three *SMARCB1* patient lines were differentiated separately, each set together with two control cell lines. Visual inspection of NPCs on day 12 revealed no obvious differences in cell morphology between control and CSS cell lines. We then measured mRNA levels of neuronal progenitor (*NESTIN*, *PAX6* and *SOX2*) and pluripotency marker (*OCT4*) genes by RT-qPCR. These data revealed no statistically significant differences in neuronal progenitor marker gene expression between patient and control NPCs (Figure 6-a-b), despite some variation between the individual cell lines. Furthermore, *OCT4* mRNA expression was only detected in an iPSC control but not in any of the NPC lines (Supplementary Figure 1a). This was verified by IF-staining of OCT4 in NPCs (Supplementary Figure 1b). In addition, we performed IF-staining for the neuronal markers SOX1 and MAP2 and observed similar signals in control and patient NPCs (Figure 6c-d). Finally, IF staining also showed no obvious differences in signal for the apoptosis marker cleaved caspase-3 between control and patient NPC lines (Figure 7). Altogether, these results suggest that *ARID1B* and *SMARCB1* patient-derived iPSCs can be differentiated to day 12 NPCs without showing increased levels of apoptosis.

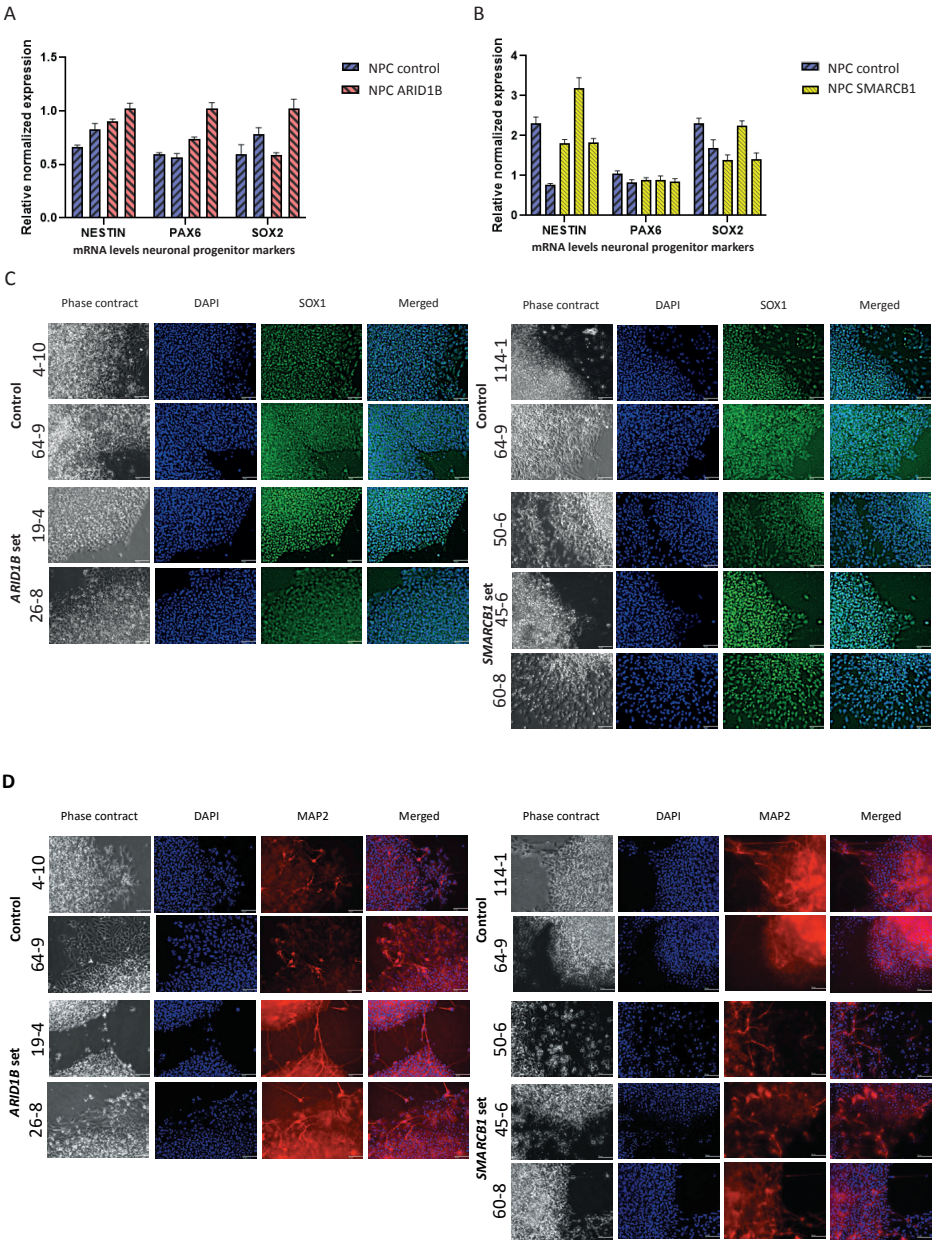


Figure 6: **a**-**b**) mRNA levels of *NESTIN*, *PAX6* and *SOX2* in day 12 differentiated cells. **c**-**d**) IF-staining of SOX1 and MAP2 on day 12 differentiated cells.

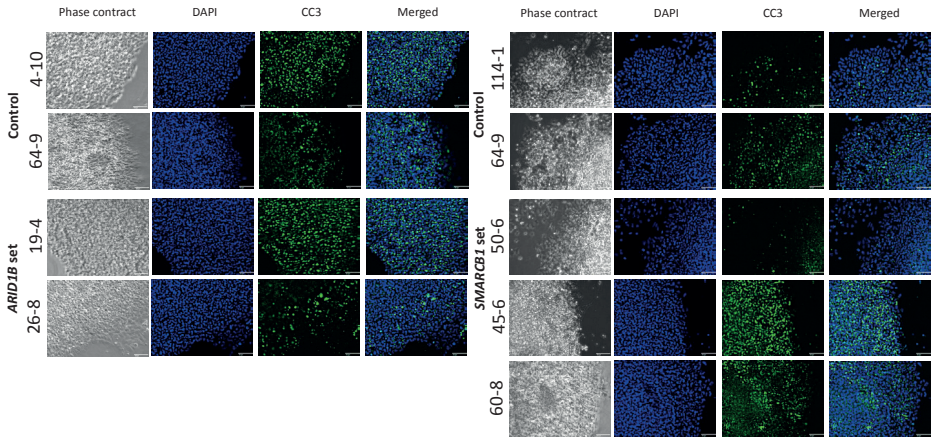


Figure 7: IF-staining of cleaved caspase-3 (CC3) in day 12 differentiated cells.

Next, we differentiated control and patient iPSC lines until day 16, when neuronal rosette formation can be observed[27]. Visual inspection at day 16 revealed that when compared to controls, neuronal rosettes of *ARID1B* patients appeared smaller while *SMARCB1* patient neuronal rosettes showed no obvious morphological differences (Figure 8a). When measured by RT-qPCR, we found that mRNA levels of neuronal (progenitor) markers (*NESTIN*, *PAX6*, *SOX2*, *MAP2*) did not differ significantly between control and *ARID1B* patient neuronal rosettes (Figure 8b). Performing these experiments in *SMARCB1* cell lines revealed largely similar results except for a small, yet statistically significant, reduction in *PAX6* mRNA levels (Figure 8c). We did not detect *OCT4* mRNA expression in *ARID1B* and *SMARCB1* patient neuronal rosettes (Supplementary Figure 2). This result is in agreement with our observations after 12 days of differentiation (Supplementary Figure 1a-b), and suggests that in our model system *ARID1B* haploinsufficiency does not impair exit from pluripotency. Combined, these results indicate that both *ARID1B* and *SMARCB1* patient-derived iPSCs can be successfully differentiated to day 12 neuronal progenitor cell populations and that day 16 neuronal rosettes express neuronal marker genes at levels largely similar to controls.

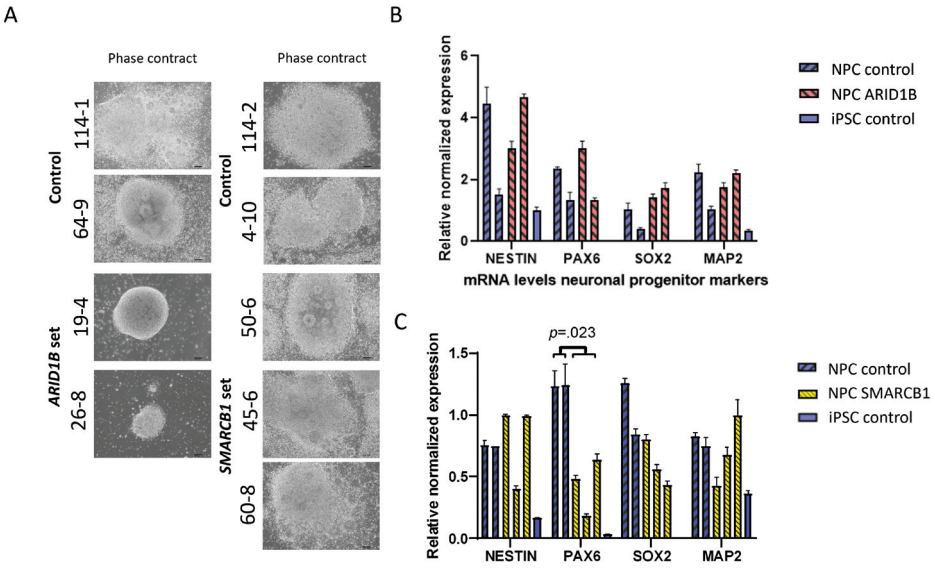


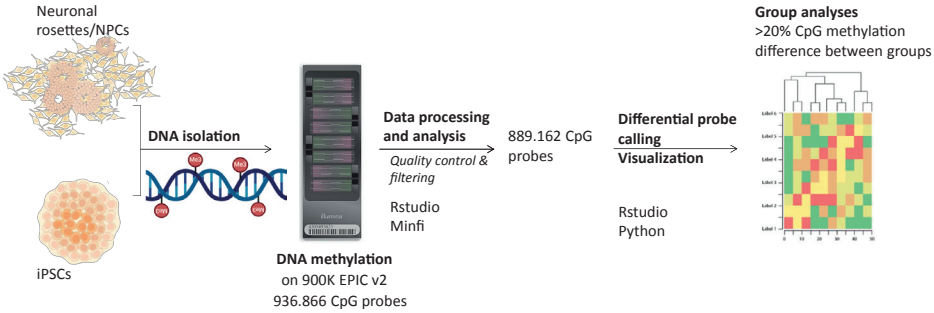
Figure 8: **a**) Phase contrast microscopy images taken at differentiation day 16 of control and patient lines, showing relatively small neuronal rosettes in *ARID1B* patient lines, and in *SMARCB1* patient lines neuronal rosette appear larger compared to control lines. **b**-**c**) mRNA levels of neuronal progenitor markers: *NESTIN*, *PAX6*, *SOX2*, *MAP2* in day 16 differentiated cells.

Genome-wide DNA methylation analysis in control and CSS patient-derived iPSCs and NPCs

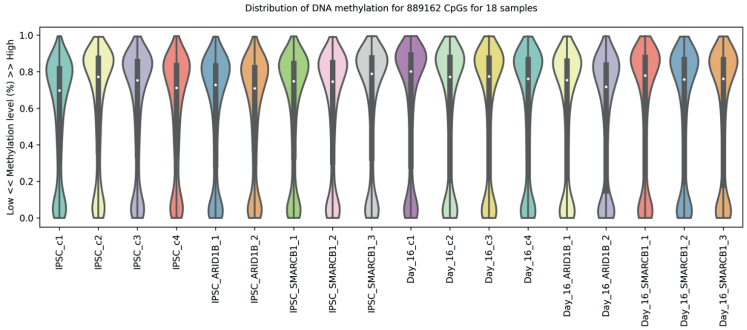
Previous studies have reported DNA methylation differences between *ARID1B* and *SMARCB1* patient samples and controls in blood[10], which can be used for CSS diagnostics. However, it is currently not known whether these CSS-associated methylation differences, termed BAFopathy episignatures, are also present in tissue and cell types other than blood. Therefore, we next asked whether the blood-BAFopathy episignature can also be detected in iPSCs and/or NPCs. While iPSCs represent the earliest stage where we may observe differences in DNA methylation levels between patient and control cell lines, NPCs represent a BAFopathy phenotype relevant cell type.

For the genome-wide DNA methylation profiling, we used the Illumina 900K bead chip array that enabled us to measure methylation levels of ~900K CpGs genome-wide in control and patient-derived iPSCs and NPCs (Figure 9a). After quality control and filtering, 889.162 probes were retained per sample for further analysis. We observed that control iPSCs had median b-values ranging between 0.697 and 0.772, while *ARID1B* patient-derived iPSCs showed median b-values between 0.709 and 0.712. For *SMARCB1* patient-derived iPSCs, we measured median b-values between 0.746 and 0.788 (Figure 9b, Supplementary Table 3). For the NPCs, median b-values were ranging

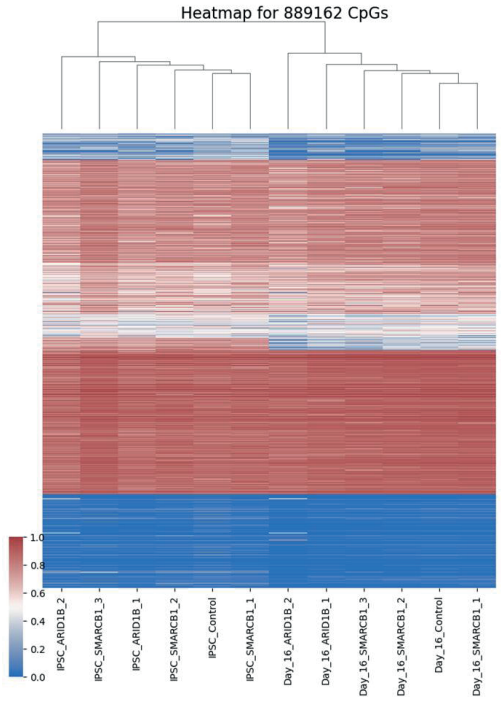
A


Neuronal rosettes/NPCs
iPSCs
DNA isolation
DNA methylation
on 900K EPIC v2
936.866 CpG probes
Data processing and analysis
Quality control & filtering
Rstudio
Minfi
889.162 CpG probes
Differential probe calling
Visualization
Rstudio
Python
Group analyses
>20% CpG methylation difference between groups

B


Distribution of DNA methylation for 889162 CpGs for 18 samples
Low << Methylation level (%) >> High
1.0
0.8
0.6
0.4
0.2
0.0
IPSC_c1
IPSC_c2
IPSC_c3
IPSC_c4
IPSC_ARID1B_1
IPSC_ARID1B_2
IPSC_SMARCB1_1
IPSC_SMARCB1_2
IPSC_SMARCB1_3
Day_16_c1
Day_16_c2
Day_16_c3
Day_16_c4
Day_16_ARID1B_1
Day_16_ARID1B_2
Day_16_SMARCB1_1
Day_16_SMARCB1_2
Day_16_SMARCB1_3

C


Heatmap for 889162 CpGs
1.0
0.8
0.6
0.4
0.2
0.0
IPSC_ARID1B_2
IPSC_SMARCB1_3
IPSC_ARID1B_1
IPSC_SMARCB1_2
IPSC_Control
IPSC_SMARCB1_1
Day_16_ARID1B_2
Day_16_ARID1B_1
Day_16_SMARCB1_3
Day_16_SMARCB1_2
Day_16_Control
Day_16_SMARCB1_1

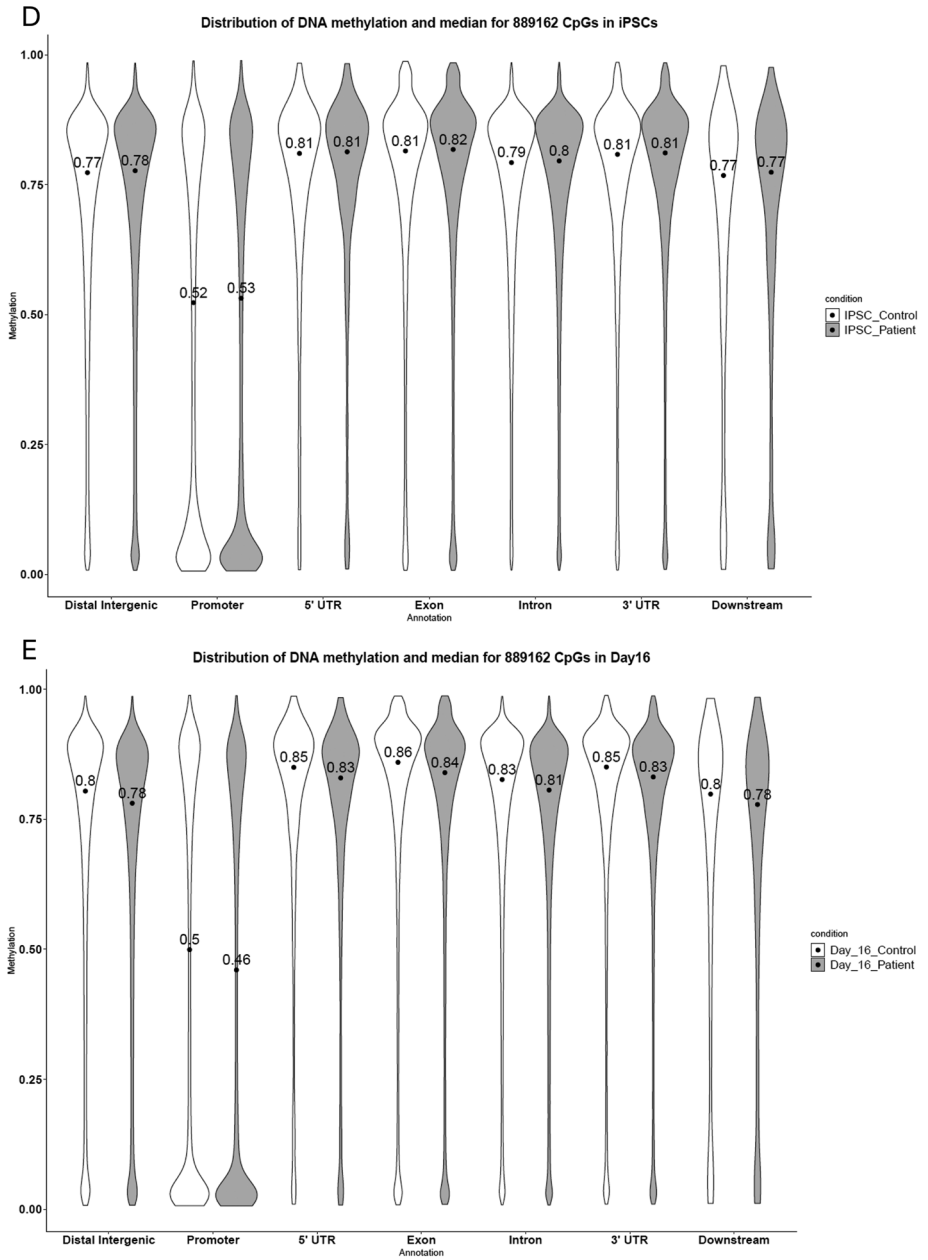


Figure 9: DNA methylation in iPSC and differentiation day 16. **a**) Representation of DNA methylation analysis. 900K EPIC v2 picture derived from https://emea.illumina.com/. **b**) Violin plots depicting individual samples' DNA methylation Beta values, white dots represent median values and black lines kernel density estimates. **c**) Heatmap showing DNA methylation for all 889162 CpGs in iPSC and NPC samples. Rows represent the CpG probes, and columns represent the samples. The color scale from dark blue to dark red represents the range of the methylation levels (beta values) between 0 and 1. **d**) Violin plots depicting DNA methylation Beta values of genomic regions of iPSCs per sample type. **e**) Violin plots depicting DNA methylation Beta values of genomic regions of NPCs per sample type.

between 0.761 and 0.801 for the controls, between 0.717 and 0.754 for the *ARID1B* samples and between 0.758 and 0.779 for the *SMARCB1* patient-derived NPCs (Figure 9b, Supplementary Table 3). Overall, this suggests that *ARID1B* haploinsufficiency and *SMARCB1* disruption, do not alter global DNA methylation levels in iPSCs and NPCs. Importantly, iPSCs clustered differently from NPCs (Figure 9c), which is expected because DNA methylation patterns are cell type-specific.

We next assessed DNA methylation levels for different genomic annotations including genic (promoters, exons, 5'and 3' UTRs, introns) and intergenic (distal intergenic downstream) regions. We found that in iPSCs, median b-values were similar between control and patient samples (Figure 9d). For the NPCs, we observed slightly lower (delta ~0.02%) median b-values in the patient when compared to the control samples for all genomic regions (Figure 9e).

Then, we asked whether the known blood BAFopathy episignature can be detected in iPSCs and/or NPCs. Therefore, we applied the described[10] BAFopathy probes to the iPSC and NPC control and CSS-patient samples. By using this set of CpGs (*n*=121), we found that the iPSC samples formed a cluster that is different from the NPCs (Figure 10a-b). However, within one cell type, patient and control samples clustered together and we could not find any evidence that the blood BAFopathy episignature is present in *ARID1B* or *SMARCB1* patient-derived iPSCs or NPCs. This suggests the BAFopathy pattern from blood is either generated further in the differentiation process or that the observed episignature in blood is tissue specific.

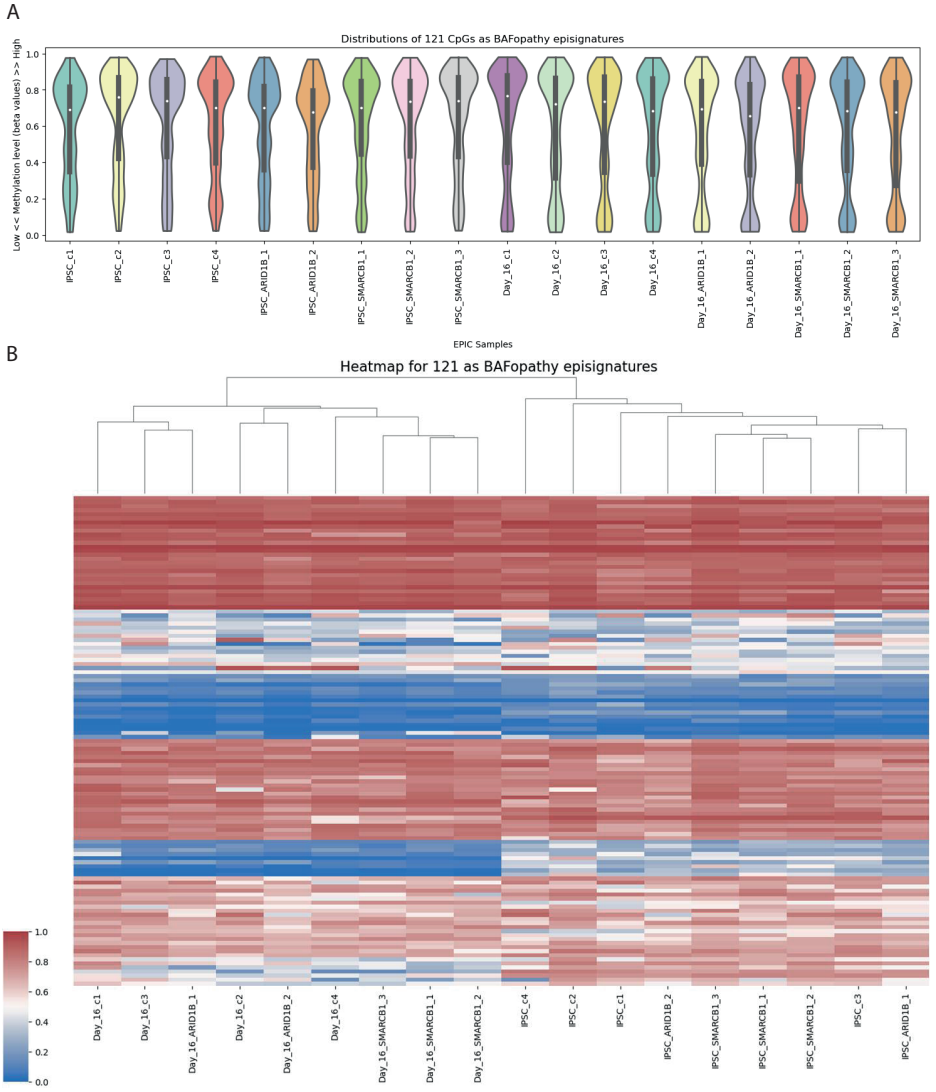


Figure 10: DNA methylation in iPSC and differentiation day 16. (**a**) Distribution of beta-values of the BAFopathy episignature probes per sample. (**b**) Clustering analysis using the BAFopathy episignature probes of iPSC control and NPC day 16 control samples. Rows represent the CpG probes, and columns represent the samples. The color scale from dark blue to dark red represents the range of the methylation levels (beta values) between 0 and 1.

DISCUSSION

Pathogenic variants in genes encoding subunits of the BAF-complex, including *ARID1B* and *SMARCB1*, lead to varying degrees of NDD in CSS patients. Understanding how these variants impact brain development remains challenging due to the difficulty of obtaining human brain samples, and findings from animal studies may not always translate to humans. Here, we employed an *in vitro* model that enables the study of neuronal differentiation differences between patients and controls, as well as within patient groups. The latter is of interest because CSS patients show a spectrum of disease severity that is linked to the underlying genetic defect, suggesting that there could be differences in cellular responses between CSS pathogenic variants.

In iPSCs, we did not observe any overt morphological differences between control and patient cell lines and when measuring expression of pluripotency markers. Furthermore, we showed that in both *ARID1B* and *SMARCB1* patient NPCs, expression of pluripotency markers was silenced, suggesting a successful exit from pluripotency. This differs from a previous observation by Pagliaroli *et al*[13], where they reported an inability of *ARID1B* patient-derived iPSCs to exit pluripotency using a 14-day differentiation paradigm towards neural crest cells. Therefore, it appears that the *ARID1B*-associated impaired exit from pluripotency is not detectable in all neuronal differentiation processes. In line with such a hypothesis, no defects in the exit from pluripotency in *ARID1B* patient-derived iPSCs were reported in a corpus callosum organoid model[12]. Combined, these observations suggest that the effects of CSS variants could be cell type-specific. Indeed, it has previously been reported that ARID1B can take on cell-type-specific roles in transcriptional regulation, where it can function in transcription activation in some cells[30] and transcriptional repression in other cell types[31].

Upon differentiation towards the neuronal lineage, we did observe morphological differences between patient and control cell lines. More specifically, on differentiation day 16, *ARID1B* patient lines exhibited smaller neuronal rosettes, while rosettes from *SMARCB1* patient lines appeared larger than those from control lines. This is interesting since CSS patients with *SMARCB1* mutations show a more severe phenotype. Although preliminary, our findings indicate that additional work will be required to characterize the observed morphological differences and the underlying molecular alterations in detail and it will be important to determine any influences on neuronal function. For instance, differentiation to mature GABAergic interneurons and investigating synaptic function using techniques such as microelectrode arrays[32] is expected to provide

deeper insights. Nevertheless, our observation could be suggestive of a different effect on neuronal differentiation of the pathogenic *ARID1B* compared to the *SMARCB1* variants.

While some of the findings presented in this chapter can be considered preliminary, they already emphasize that the generation and implementation of different *in vitro* model systems, brings unique strengths to the study of the effects of CSS variants on neuronal development and function. Since many different cell types are present in the brain, it will be important to study the (dys)function in a developing patient's brain and more elaborate models that allow differentiation into multiple cell types (e.g. organoids) and/or more mature neurons are likely to shed more light on this. For instance, a recent study investigated agenesis of the corpus callosum, the most common brain anomaly in *ARID1B*[33], *SMARCB1*[17] and CSS patients[34,35]. Martins-Costa *et al*[12] showed, using neural organoids generated from an isogenic *ARID1B* patient line, disrupted axonogenesis. They identified a transcriptional dysregulation of callosal projection neurons, less long-range projections and, less contralateral synapses. These findings underline the hypothesis that interneuron disfunction plays an important role in the pathogenesis of CSS. The *in vitro* system we are employing can be used to investigate the differentiation from iPSC to GABAergic interneurons[27]. Therefore, building on the findings of Martins-Costa *et al*[12] and incorporating ATAC- and RNA-seq in our model system could further our understanding of the role of *ARID1B* and other CSS genes in interneuron development.

At the molecular level, the BAF complex primarily functions in regulating chromatin accessibility but DNA methylation alterations between control and CSS patients have been reported in blood. This so called episignature can be used in diagnostics[10] but it remains unclear whether it is also present in other cell types. Taking advantage of our *in vitro* model, we observe that the blood episignature is not present in patient derived iPSCs or NPCs (Figure 10b). Possible explanations could be that the BAFopathy pattern from blood is generated further in the differentiation process or that the observed episignature in blood is tissue specific.

The technology used to generate episignatures plays a significant role in the quality and consistency of results. The Illumina EPIC array versions 1[36] and 2[37], while widely used, are limited by batch effects and differences in probe coverage, with v1 covering 800,000 probe locations and v2 expanding to 900,000. We used the Illumina EPIC v2 array in our study, while the BAFopathy episignature was originally detected using the EPIC v1 array. However, we do not think this impacted our investigation of the BAFopathy episignature in iPSC and NPC, as 121 out of 131[10] episignature probes were reliably measured across all our samples. To overcome the limitations of the Illumina EPIC array, future research could leverage Nanopore sequencing[38]. This technology eliminates batch effects and

offers the ability to analyze over 28 million CpG sites[39], providing a more comprehensive overview of DNA methylation.

In sum, we successfully established an *in vitro* model system to explore the effects of genetic variants on neuronal differentiation, utilizing two *ARID1B*, three *SMARCB1* patient lines and four control lines. This model could be used for further investigation of the effects of pathogenic variants in CSS genes during differentiation and on the transcriptomic and epigenomic states.

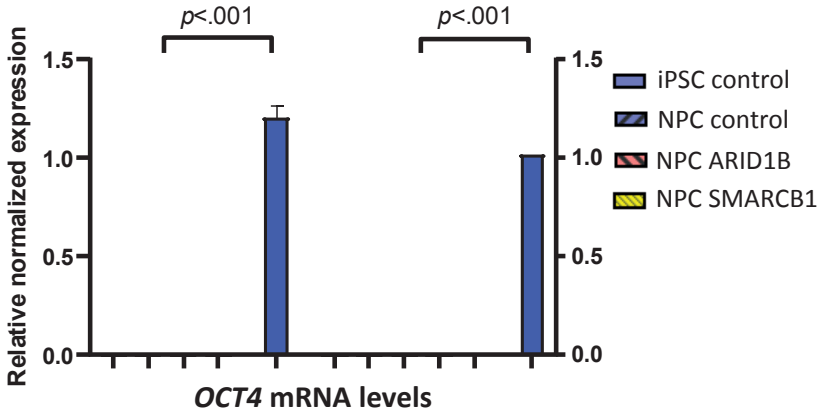
REFERENCES

1. Vergano SA, van der Sluijs PJ, Santen G. ARID1B-Related Disorder. In: Adam MP, Ardinger HH, Pagon RA, et al, eds. *GeneReviews((R))*. 1993.
2. Santen GW, Aten E, Sun Y, et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet*. Mar 18 2012;44(4):379-80. doi:10.1038/ng.2217
3. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet*. Mar 18 2012;44(4):376-8. doi:10.1038/ng.2219
4. Wieczorek D, Bogershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet*. Dec 20 2013;22(25):5121-35. doi:10.1093/hmg/ddt366
5. Santen GW, Aten E, Vulto-van Silfhout AT, et al. Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum Mutat*. Nov 2013;34(11):1519-28. doi:10.1002/humu.22394
6. Tsurusaki Y, Okamoto N, Ohashi H, et al. Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clin Genet*. Jun 2014;85(6):548-54. doi:10.1111/cge.12225
7. Kosho T, Okamoto N, Ohashi H, et al. Clinical correlations of mutations affecting six components of the SWI/SNF complex: detailed description of 21 patients and a review of the literature. *Am J Med Genet A*. Jun 2013;161A(6):1221-37. doi:10.1002/ajmg.a.35933
8. Sekiguchi F, Tsurusaki Y, Okamoto N, et al. Genetic abnormalities in a large cohort of Coffin-Siris syndrome patients. *J Hum Genet*. Sep 17 2019;doi:10.1038/s10038-019-0667-4
9. Schrier Vergano S, Santen G, Wieczorek D, Wollnik B, Matsumoto N, Deardorff MA. Coffin-Siris Syndrome. In: Pagon RA, Adam MP, Ardinger HH, et al, eds. *GeneReviews(R)*. University of Washington; 1993.
10. Aref-Eshghi E, Bend EG, Hood RL, et al. BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin-Siris and Nicolaides-Baraitser syndromes. *Nat Commun*. Nov 20 2018;9(1):4885. doi:10.1038/s41467-018-07193-y
11. Jung EM, Moffat JJ, Liu J, Dravid SM, Gurumurthy CB, Kim WY. Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. *Nat Neurosci*. Dec 2017;20(12):1694-1707. doi:10.1038/s41593-017-0013-0
12. Martins-Costa C, Wiegers A, Pham VA, et al. ARID1B controls transcriptional programs of axon projection in an organoid model of the human corpus callosum. *Cell Stem Cell*. May 6 2024;doi:10.1016/j.stem.2024.04.014
13. Pagliaroli L, Porazzi P, Curtis AT, et al. Inability to switch from ARID1A-BAF to ARID1B-BAF impairs exit from pluripotency and commitment towards neural crest formation in ARID1B-related neurodevelopmental disorders. *Nat Commun*. Nov 9 2021;12(1):6469. doi:10.1038/s41467-021-26810-x
14. Celen C, Chuang JC, Luo X, et al. Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. *Elife*. Jul 11 2017;6:e25730. doi:10.7554/eLife.25730
15. Shibutani M, Horii T, Shoji H, et al. Arid1b Haploinsufficiency Causes Abnormal Brain Gene Expression and Autism-Related Behaviors in Mice. *Int J Mol Sci*. Aug 30 2017;18(9)doi:10.3390/ijms18091872
16. Kim H, Kim D, Cho Y, et al. Early postnatal serotonin modulation prevents adult-stage deficits in Arid1b-deficient mice through synaptic transcriptional reprogramming. *Nat Commun*. Aug 27 2022;13(1):5051. doi:10.1038/s41467-022-32748-5

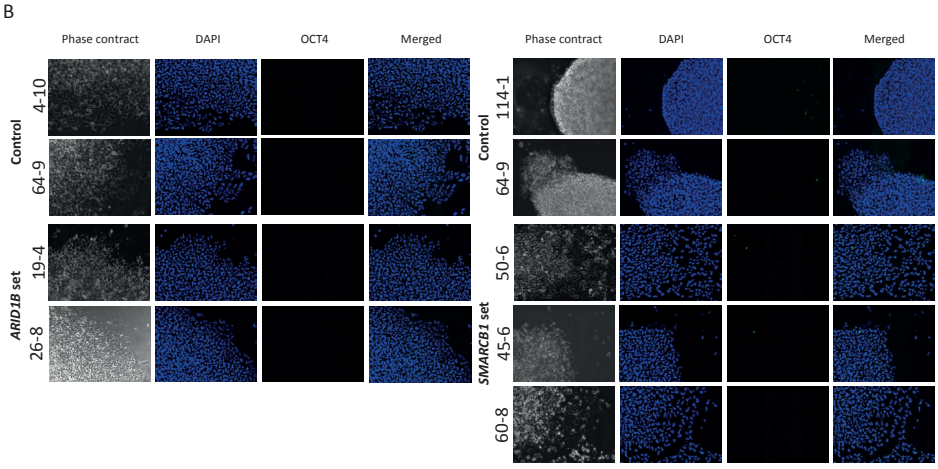
17. Filatova A, Rey LK, Lechler MB, et al. Mutations in SMARCB1 and in other Coffin-Siris syndrome genes lead to various brain midline defects. *Nat Commun*. Jul 4 2019;10(1):2966. doi:10.1038/s41467-019-10849-y
18. Ka M, Chopra DA, Dravid SM, Kim WY. Essential Roles for ARID1B in Dendritic Arborization and Spine Morphology of Developing Pyramidal Neurons. *J Neurosci*. Mar 02 2016;36(9):2723-42. doi:10.1523/JNEUROSCI.2321-15.2016
19. Smith AL, Jung EM, Jeon BT, Kim WY. Arid1b haploinsufficiency in parvalbumin- or somatostatin-expressing interneurons leads to distinct ASD-like and ID-like behavior. *Sci Rep*. May 12 2020;10(1):7834. doi:10.1038/s41598-020-64066-5
20. Valencia AM, Collings CK, Dao HT, et al. Recurrent SMARCB1 Mutations Reveal a Nucleosome Acidic Patch Interaction Site That Potentiates mSWI/SNF Complex Chromatin Remodeling. *Cell*. Nov 27 2019;179(6):1342-1356 e23. doi:10.1016/j.cell.2019.10.044
21. Kang E, Kang M, Ju Y, et al. Association between ARID2 and RAS-MAPK pathway in intellectual disability and short stature. *J Med Genet*. Nov 2021;58(11):767-777. doi:10.1136/jmedgenet-2020-107111
22. Li C, Fleck JS, Martins-Costa C, et al. Single-cell brain organoid screening identifies developmental defects in autism. *Nature*. Sep 2023;621(7978):373-380. doi:10.1038/s41586-023-06473-y
23. Crabtree G, Son E, Krokhotin A, Gourisankar S, Chang C-Y. ARID1B is a Dosage-sensitive Regulator of Polycomb Repressive Complex Distribution and HOX Gene Regulation in Patient-derived Neural Progenitors. Research Square; 2021.
24. Buijsen RAM, Gardiner SL, Bouma MJ, et al. Generation of 3 spinocerebellar ataxia type 1 (SCA1) patient-derived induced pluripotent stem cell lines LUMCi002-A, B, and C and 2 unaffected sibling control induced pluripotent stem cell lines LUMCi003-A and B. *Stem Cell Res*. May 2018;29:125-128. doi:10.1016/j.scr.2018.03.018
25. Chen J, Liu J, Chen Y, et al. Rational optimization of reprogramming culture conditions for the generation of induced pluripotent stem cells with ultra-high efficiency and fast kinetics. *Cell Res*. Jun 2011;21(6):884-94. doi:10.1038/cr.2011.51
26. Warlich E, Kuehle J, Cantz T, et al. Lentiviral vector design and imaging approaches to visualize the early stages of cellular reprogramming. *Mol Ther*. Apr 2011;19(4):782-9. doi:10.1038/mt.2010.314
27. Liu Y, Liu H, Sauvey C, Yao L, Zarnowska ED, Zhang SC. Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. *Nat Protoc*. Sep 2013;8(9):1670-9. doi:10.1038/nprot.2013.106
28. Sinke L, Cats D, Heijmans BT. Omixer: multivariate and reproducible sample randomization to proactively counter batch effects in omics studies. *Bioinformatics*. Sep 29 2021;37(18):3051-3052. doi:10.1093/bioinformatics/btab159
29. Moffat JJ, Jung EM, Ka M, Jeon BT, Lee H, Kim WY. Differential roles of ARID1B in excitatory and inhibitory neural progenitors in the developing cortex. *Sci Rep*. Feb 16 2021;11(1):3856. doi:10.1038/s41598-021-82974-y
30. Nagl NG, Jr., Wang X, Patsialou A, Van Scoy M, Moran E. Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J*. Feb 07 2007;26(3):752-63. doi:10.1038/sj.emboj.7601541
31. Raab JR, Resnick S, Magnuson T. Genome-Wide Transcriptional Regulation Mediated by Biochemically Distinct SWI/SNF Complexes. *PLoS Genet*. Dec 2015;11(12):e1005748. doi:10.1371/journal.pgen.1005748
32. Yuan X, Puvogel S, van Rhijn JR, et al. A human in vitro neuronal model for studying homeostatic plasticity at the network level. *Stem Cell Reports*. Nov 14 2023;18(11):2222-2239. doi:10.1016/j.stemcr.2023.09.011

33. van der Sluijs PJ, Jansen S, Vergano SA, et al. The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin–Siris syndrome. *Genetics in Medicine*. 2018/10/22 2018;doi:10.1038/s41436-018-0330-z
34. van der Sluijs PJ, Joosten M, Alby C, et al. Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort. *Genet Med*. May 17 2022;doi:10.1016/j.gim.2022.04.010
35. Kosho T, Okamoto N, Coffin-Siris Syndrome International C. Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *Am J Med Genet C Semin Med Genet*. Sep 2014;166C(3):262-75. doi:10.1002/ajmg.c.31407
36. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. Oct 7 2016;17(1):208. doi:10.1186/s13059-016-1066-1
37. Peters TJ, Meyer B, Ryan L, et al. Characterisation and reproducibility of the HumanMethylationEPIC v2.0 BeadChip for DNA methylation profiling. *BMC Genomics*. Mar 6 2024;25(1):251. doi:10.1186/s12864-024-10027-5
38. Rand AC, Jain M, Eizenga JM, et al. Mapping DNA methylation with high-throughput nanopore sequencing. *Nat Methods*. Apr 2017;14(4):411-413. doi:10.1038/nmeth.4189
39. Silva C, Machado M, Ferrao J, Sebastiao Rodrigues A, Vieira L. Whole human genome 5'-mC methylation analysis using long read nanopore sequencing. *Epigenetics*. Dec 2022;17(13):1961-1975. doi:10.1080/15592294.2022.2097473

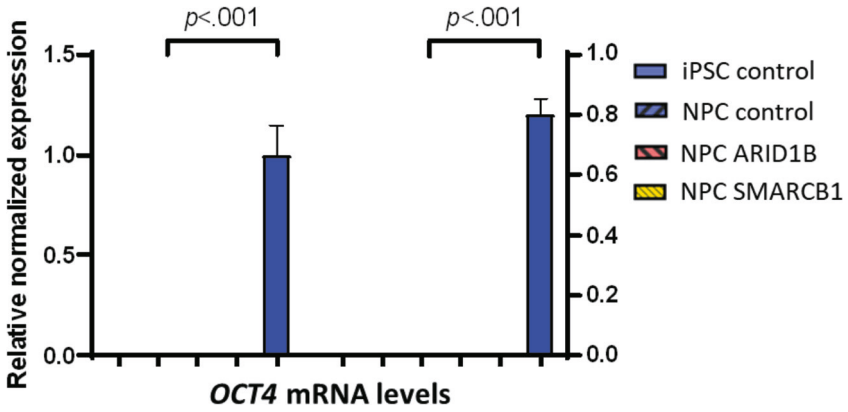
SUPPLEMENTARY TABLES AND FIGURES



Supplementary Figure 1a: mRNA levels *OCT4* at differentiation day 12



Supplementary Figure 1b: IF -staining of OCT4 on day 12 differentiated cells



Supplementary Figure 2: mRNA expression of *OCT4* at differentiation day 16

Supplementary Table 1: Antibodies used for immunofluorescence

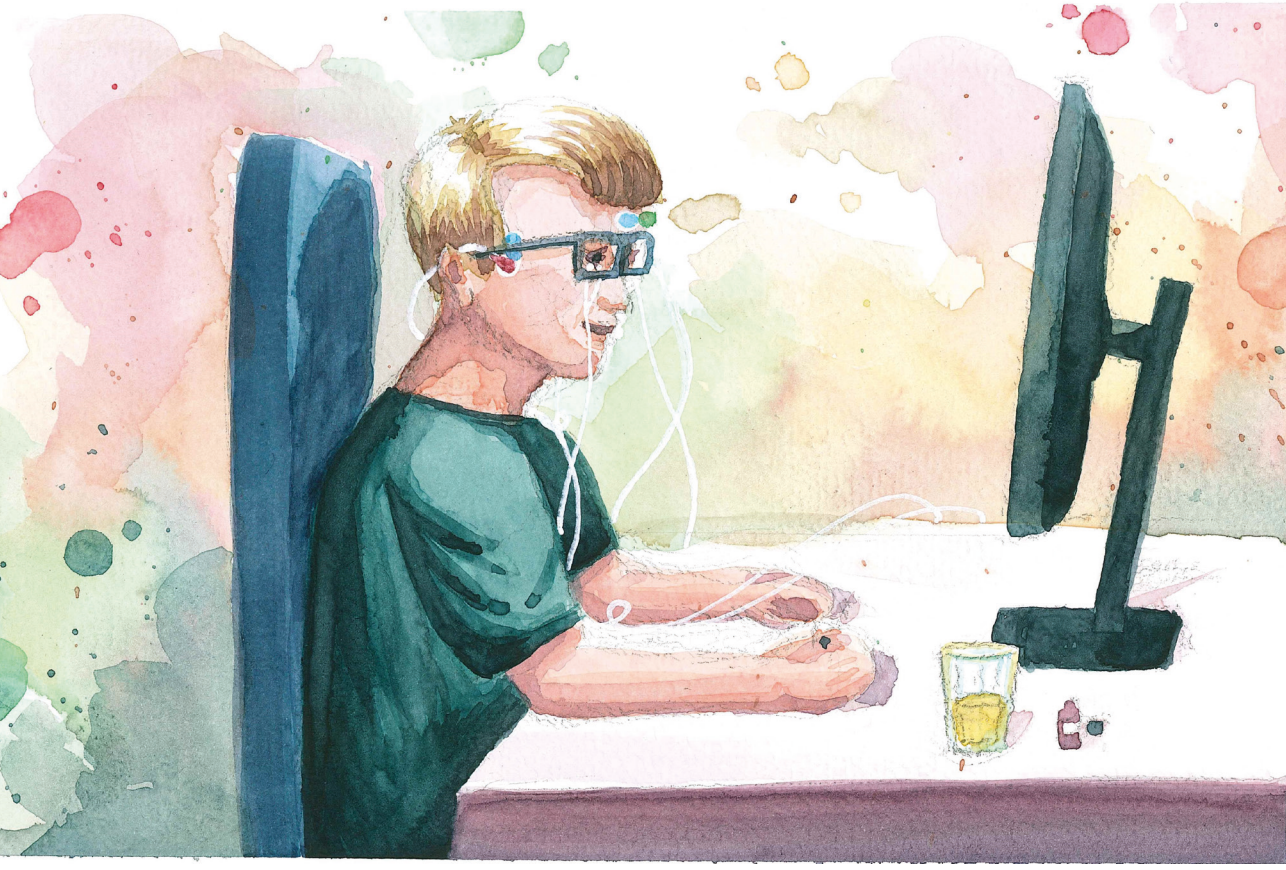
Antibody	Host	Article number	Dilution
OCT4	rabbit	ab19857	1:500
NANOG	rabbit	ab21624	1:500
SOX1	rabbit	Invitrogen PA5-23351	1:500
MAP2	mouse	MAB364, 3162484	1:500
Cleaved caspase-3	rabbit	rb9661S	1:500
anti-Rabbit 488	donkey	AlexaFluorTM 488 (A21206)	1:2000
anti-Mouse 594	donkey	AlexaFluorTM 594 (A21203)	1:2000
DAPI	-	R37605	

Supplementary Table 2: Primers used for RT-qPCR

	FW	**REV**	**Length**
Housekeeping gene			
GUS	CTCATTTGGAATTTTGCCGATT	CCGAGTGAAGATCCCCTTTTTA	81
Pluripotency markers			
OCT4	GACAGGGGGAGGGGAGGAGCTAGG	CTTCCCTCCAACCAGTTGCCCCAAAC	144
NANOG	AAGGTCCCGGTCAAGAAACA	TCTTCACCTGTTTGTAGCTGA	154
Neuronal (progenitor) marker			
NESTIN	CTCCAGAAACTCAAGCACC	TCCTGATTCTCCTCTTCCA	145
PAX6	GGTTGGTATCCGGGGACTT	TCCGTTGGAACTGATGGAGT	102
SOX2	AGCTCGCAGACCTACATGAA	TGGAGTGGGAGGAAGAGGTA	151
MAP2	AAGAGAATGGGATCAACGGAG	TGCTACAGCCTCAGCAGTGA	101
Genes			
ARID1B	GCGTGTGATGATGTCCCTTA	CCAGACAACTGGGAGAGATTG	123
SMARCB1	CTTCAGCGAGAACCCTCTGC	GCAAGACGCCTCATCCGC	156

Supplementary Table 3: median DNA methylation Beta values of individual samples

sample		median beta-value
iPSC	c1	0,697
iPSC	c2	0,772
iPSC	c3	0,753
iPSC	c4	0,712
iPSC	ARID1B_1	0,728
iPSC	ARID1B_2	0,709
iPSC	SMARCB1_1	0,748
iPSC	SMARCB1_2	0,746
iPSC	SMARCB1_3	0,788
Day 16	c1	0,801
Day 16	c2	0,722
Day 16	c3	0,775
Day 16	c4	0,761
Day 16	ARID1B_1	0,754
Day 16	ARID1B_2	0,717
Day 16	SMARCB1_1	0,779
Day 16	SMARCB1_2	0,758
Day 16	SMARCB1_3	0,761



Chapter 8

Clonazepam repurposing in ARID1B patients through conventional RCT and N-of-1 trials: an experimental strategy for orphan disease development

Pleuntje J van der Sluijs*, Koshar Safai Pour*, Cécile L Berends, Matthijs D Kruizinga, Annelieke R Müller, Agnies M van Eeghen, Mar Rodríguez-Girondo, Maria J Juachon, Duco Steenbeek, Adam F Cohen, Rob G J A Zuiker**, Gijs W E Santen**

*Contributed equally.
** Joint senior authors.

Published: J Med Genet. 2024 Dec 31;62(3):e109951. doi: 10.1136/jmg-2024-109951

ABSTRACT

Background: Clinical trials for rare disorders have unique challenges due to low prevalence, patient phenotype variability and high expectations. These challenges are highlighted by our study on clonazepam in *ARID1B* patients, a common cause of intellectual disability. Previous studies on Arid1b-haploinsufficient mice showed positive effects of clonazepam on various cognitive aspects.

Methods: This study used a randomised, double-blinded, placebo-controlled, two-way crossover study (RCT), followed by an N-of-1 design. In the crossover study, *ARID1B* patients received clonazepam (max 0.5 mg, two times per day) or a placebo for 22 days with a 3-week washout period. Assessments included safety, tolerability, pharmacokinetics, pharmacodynamics on neurocognitive tasks, behaviour and cognitive function. During phase I of the N-of-1 trial the optimal dosage and individual treatment goals were determined. Phase II evaluated the treatment effect. This phase was composed of three periods: an open-label period with placebo (4 weeks), followed by a double-blinded period (6 weeks), followed by an open-label period in which the patient received clonazepam (4 weeks).

Results: In the clonazepam group (*n*=16, 15 completing both periods), seven (44%) reported improvement on Clinician Global Impression of Improvement versus two (13%) on placebo. 13 (87%) showed 'no change' after placebo (two (13%) on clonazepam), while seven (44%) on clonazepam reported deterioration, often linked to side effects (*n*=6), suggesting potential benefit from lower dosing. Three N-of-1 trials with RCT responders saw two patients improve on clonazepam during double-blinding, but clinical evaluation deemed the improvements insufficient.

Conclusions: Our approach shows the feasibility and strength of combining conventional RCT and N-of-1 studies for therapeutic studies in populations with intellectual disabilities, distinguishing real treatment effects from expectation bias. Our findings suggest that clonazepam has no additional therapeutic value in *ARID1B* patients.

Trial registration number: EUCTR2019-003558-98, ISRCTN11225608.

Keywords: Mental Disorders, Therapeutics

WHAT IS ALREADY KNOWN ON THIS TOPIC

- Clinical trials for rare disorders have unique challenges due to the prevalence of the disorder, phenotype variability and (parental) expectations.
- Pathogenic variants in the *ARID1B* gene are among the most frequent causes of intellectual disability.
- Arid1b heterozygous mice show improvement on administration of clonazepam.

WHAT THIS STUDY ADDS

- The effect of clonazepam in *ARID1B* patients was investigated using a conventional crossover randomised controlled trial (RCT) design and an N-of-1 design.
- Effects of clonazepam were observed in the RCT, but the N-of-1s only replicated the effect on sleep, suggesting that clonazepam has no additional therapeutic value in *ARID1B* patients.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- This study demonstrates how the N-of-1 trial effectively employs and validates observed effects, differentiating between genuine treatment effects and expectation bias.

INTRODUCTION

ARID1B is one of the top mutated genes in heterogeneous intellectual disability, developmental delay and autism cohorts.[1-3] Most patients with Coffin-Siris syndrome (CSS, OMIM 13500)[4-7] have pathogenic variants in *ARID1B*. With the increasing availability and use of genome-wide diagnostics, the number of identified patients has increased. Several hundred cases have been reported in the literature between 2012 and 2022. Almost all *ARID1B* patients have developmental delay,[8] often with behavioural manifestations such as short attention span and hyperactivity.[8] In addition, expressive speech is usually severely affected.[8 9] As a result, individuals with a pathogenic variant in *ARID1B* can face numerous challenges in their daily lives.[8] These challenges can impact their ability to communicate, socialise or perform daily activities. Unfortunately, until now, no targeted established therapies exist to improve their functioning.

Three Arid1b-haploinsufficient (Arid1b$^{+/-}$) mice models have been generated and investigated.[10-13] These Arid1b$^{+/-}$ mice displayed increased anxiety, and reduced memory, learning and social interaction. Jung *et al* identified a significant reduction in GABAergic inhibitory interneurons in their model, particularly of the parvalbumin-positive subtype.[11] An investigation of the balance between inhibitory and excitatory synapses revealed a decrease in GABAergic inhibitory synapses, likely causing an inhibition-excitation imbalance. Administering a $GABA_A$ receptor positive allosteric modulator, such as clonazepam, may partly reverse this imbalance. Consequently, the investigators administered a single intraperitoneal dose of 0.0625 mg/kg clonazepam to adult mice. Clonazepam-treated Arid1b$^{+/-}$ mice, compared with saline-treated mice, performed better in object recognition, sociability novelty and demonstrated a decrease in anxiety-like behaviour 30–60 min post-treatment, while depression-like behaviour symptoms remained unaffected.[11]

Clonazepam is a Food and Drug Administration (FDA)-approved non-selective benzodiazepine with a half-life of 20–40 hours and a Tmax of 1–4 hours after oral dosing. It is currently used to treat seizures in children and adults and occasionally for anxiety. Given its promising effects in the Arid1b$^{+/-}$ mouse model, clonazepam is a potential low-cost treatment suitable for repurposing to improve the functioning of *ARID1B* patients.

In preparation for a randomised controlled clinical trial (RCT) investigating the effects of clonazepam in *ARID1B* patients, we conducted two preliminary studies. One study established the correlation between clonazepam concentrations in plasma and saliva in healthy volunteers, as plasma sampling (via blood draw) is invasive for *ARID1B* patients.[14]

Another study examined the neurocognitive phenotype of 12 *ARID1B* patients and 12 age-matched controls and identified potential biomarkers and outcome measures for treatment evaluation.[15] Important results of this study were that tests which previously have been evaluated in benzodiazepines,[16 17] such as for animal fluency, eye movements (saccadic peak velocity and smooth pursuit), body sway, finger tapping and adaptive tracking, could be performed adequately by *ARID1B* patients. Based on these studies, we designed this placebo-controlled study to assess the safety and tolerability, pharmacokinetics (PK) and pharmacodynamics (PD) of clonazepam treatment in *ARID1B* patients.

In the current study, we aim to investigate the effect of clonazepam in *ARID1B* patients and explore biomarkers known to be sensitive to clonazepam with the NeuroCart test procedures. Our research design involves a randomised, double-blinded, placebo-controlled, two-way crossover study, followed by an additional crossover in three patients, resulting in an N-of-1 trial to further delineate the observed treatment effect per patient. Furthermore, by addressing the challenges of validating biomarkers and managing parental expectations, our study may serve as an example of navigating the complexities of research in rare diseases and contribute to improving the care and outcomes of patients with *ARID1B*-related disorders.

METHODS

This study was conducted at the Centre of Human Drug Research (CHDR) in Leiden, the Netherlands, from September 2021 until May 2022. The study was registered at the EU Clinical Trials Register (2019-003558-98). The protocol is both available here and on the International Standard Randomised Controlled Trial Number registry (ISRCTN11225608). The study was conducted according to the Dutch Act on Medical Research Involving Human Subjects, the Dutch codes of conduct regarding medical research with minors and expression of objection by people with mental disabilities and in compliance with Good Clinical Practice.

Subjects

Inclusion criteria were the identification of a pathogenic variant in *ARID1B* and an age of at least 6 years. Exclusion criteria consisted of a history of severe respiratory problems or severe liver or renal insufficiency, or another medical or psychosocial history making the patient unsuitable for participation as determined by the treating physician or general practitioner.

Power calculation

Power calculation was performed according to two approaches. (1) Based on the Clinician Global Impression of Improvement (CGI-I) SD of 1.5 points, we calculated that our study would have an 80% power to detect a 1 point difference in CGI-I if 20 patients were included. (2) We also previously calculated the minimally detectable difference for all NeuroCart endpoints with 16 patients (see table 3 in Kruizinga *et al*[15]).

Study design

Randomised, double-blinded, placebo-controlled, two-way crossover study

Patients with pathogenic variants in *ARID1B* were recruited via the CSS expertise centre of the Leiden University Medical Center (LUMC), Leiden, the Netherlands. The schedule of assessments is listed in online supplemental table S1. The randomisation code was generated by a study-independent CHDR statistician. In this two-way crossover, placebo-controlled, randomised study, each study period or occasion was 22 days, and a 3-week washout separated each study period or occasion (figure 1A). The starting dose (i.e., 0.01 mg/kg daily) was decided with several physicians with experience in prescribing clonazepam to children, in an attempt to minimise the chance on sedative effects. This dose is half the minimum starting dosage of the administration of clonazepam for anxiety in children (i.e., 0.02 mg/kg daily). Study drug clonazepam or placebo was administered to the subjects as follows: day 1–3: starting dose was 0.005 mg/kg, two times per day (max. 0.5 mg); day 4–6: 0.01 mg/kg, two times per day (max. 0.5 mg); day 7–22: 0.015 mg/kg, two times per day (max. 0.5 mg). Patients received a single starting dose of 0.005 mg/kg (max. 0.5 mg) at their first visit and were subsequently monitored for 5 hours for safety and tolerability, PK and PD effects. If the dose on day 1 was well tolerated, patients took the same dose before bedtime and continued the regimen as described. If side effects related to the treatment were observed, dose was reduced to half of the current dose for 3 days. Thereafter, the original dose at which the side effects first occurred was restarted. If side effects re-emerged, patients remained on the lower dose. In case of clinically significant side effects at the lower dose, patients were excluded.

Furthermore, the clinician and parents determined the Clinician Global Impression of Severity (CGI-S) on day 1 and the CGI-I on day 22.[18] The teacher or mentor of the subject also assessed the CGI-S and CGI-I. In addition, the parents completed the Aberrant Behavior Checklist (ABC) on day 22.[19] Within days 1 and 22, patients continued their routine, wore a Steel HR smartwatch and performed home-based assessments (animal fluency, adaptive tracking and finger tapping) two times per week.

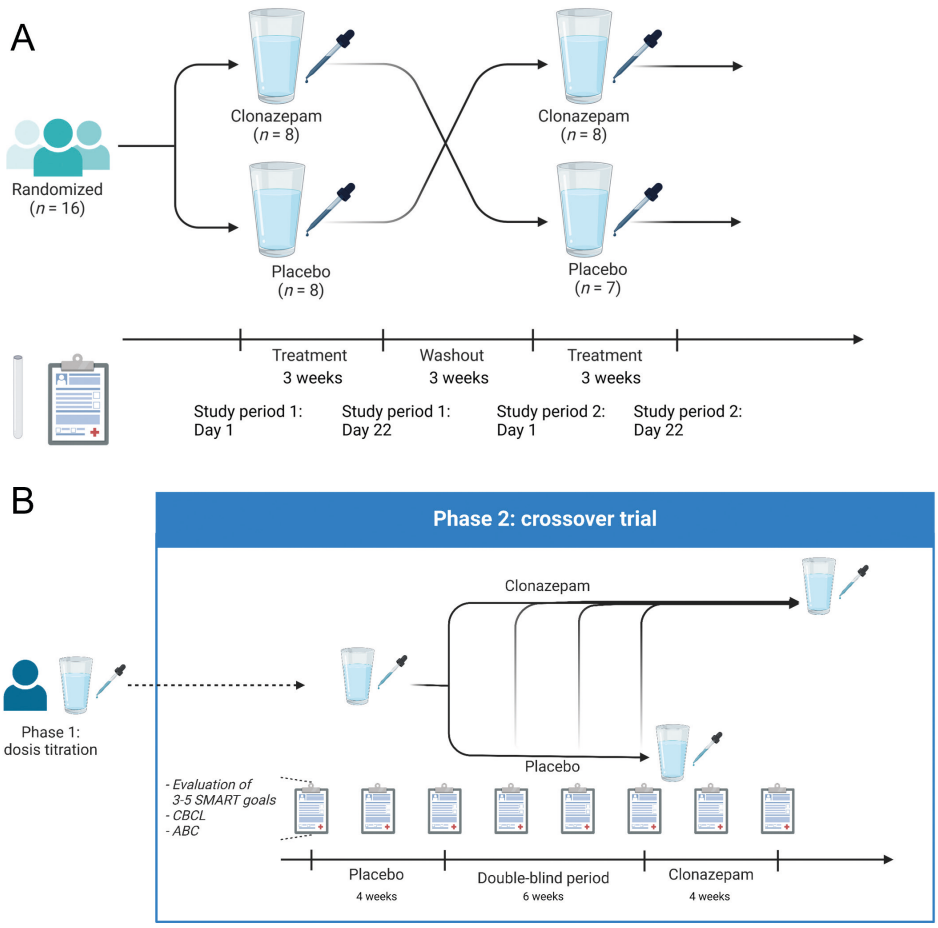


Figure 1. Study design. (**A**) Conventional crossover randomised controlled trial (RCT). (**B**) Additional crossover study creating an N-of-1 design. ABC, Aberrant Behavior Checklist; CBCL, Child Behavior Checklist; SMART, specific, measurable, achievable, relevant and time-bound. Figures 1A and 1B were created with BioRender.com (2024).

N-of-1 design

Parents who spontaneously indicated they wanted to continue clonazepam treatment were included in an N-of-1 follow-up protocol. The N-of-1 trials were performed before unblinding and no data of the crossover study were available. This trial was composed of two phases (figure 1B).

Phase I: dose titration (variable duration, max. 6 weeks)

The objective of the open-label titration phase was to identify the minimum effective dosage. The open-label titration phase started with a daily dose of 0.01 mg/kg/day in two doses. After 4–7 days, the effects were evaluated and compared with results of

the conventional RCT, and the dose was adjusted based on parental feedback, with a maximum dose of 0.03 mg/kg/day. Once parents were satisfied with the effect, this dose was maintained for at least 7 days. After this, the titration phase ended, and the dose was tapered off by 0.01 mg/kg every 3 days. During this phase, individual primary outcome treatment goals were set in consultation with parents using weighed Goal Attainment Scaling (GAS) with six defined levels per goal (see online supplemental table S5)[20 21] and a paper copy was given to parents.

Phase II: a single crossover trial consisting of three treatment periods: placebo, double blind and clonazepam

Phase II lasted 14 weeks and started with an open-label placebo period of 4 weeks. Subsequently, a double-blinded period was started until week 10. During this period, new study medication was started every 2 weeks. This study medication was prepared and provided by our pharmacy. The pharmacist determined randomly when the switch to clonazepam was made during the double-blinded period. This double-blinded period was followed by an open-label period of 4 weeks of clonazepam (figure 1B). There was only one crossover from placebo to clonazepam; therefore, no washout period was required. Regular contact with parents, teacher or mentor by the physician occurred every 2 weeks by phone to evaluate progress in achieving previously established goals. Each goal was scored individually during phase II; from these separate scores, a single aggregated T score was produced using a standardised formula.[22] Treatment goals were classified in the domains defined by the International Classification of Functioning, Disability and Health-Children and Youth version.[23] Additionally, parents were asked to complete the Child Behavior Checklist (CBCL) questionnaire every 2 weeks and the ABC questionnaire, as used in the conventional RCT study.

End of the N-of-1 study

At the end of the study, parents and the study physician were unblinded and decided whether to continue treatment by evaluating their experiences and GAS scores during the study period .

NeuroCart test procedures

All NeuroCart tests on days 1 and 22 of each study period were conducted according to the protocol by trained instructors. NeuroCart tests were performed as described in online supplemental textbox S1.

Statistics

Randomised controlled trial

All safety and statistical programming were conducted with SAS V.9.4 for Windows or newer (SAS Institute). PK variable programming was conducted with R V.3.6.1 for Windows or newer (R Foundation for Statistical Computing/R Development Core Team, Vienna, Austria, 2010). To establish treatment effects, the repeatedly measured crossover PD endpoints were analysed by mixed model analyses of variance with treatment, period, time and treatment by time as fixed effects, with subject, subject by treatment and subject by time as a random effect, and with the (average) baseline value as covariate.

N-of-1 design

All analyses were performed using SPSS V.29. For the primary outcome, linear mixed models (LMM) were used to compare GAS-T scores between placebo and clonazepam treatment, blinding was added as a fixed covariate. The individual was included as a random intercept. In a similar manner, LMM was used to determine whether CBCL and ABC questionnaire results were different between placebo and clonazepam treatment.

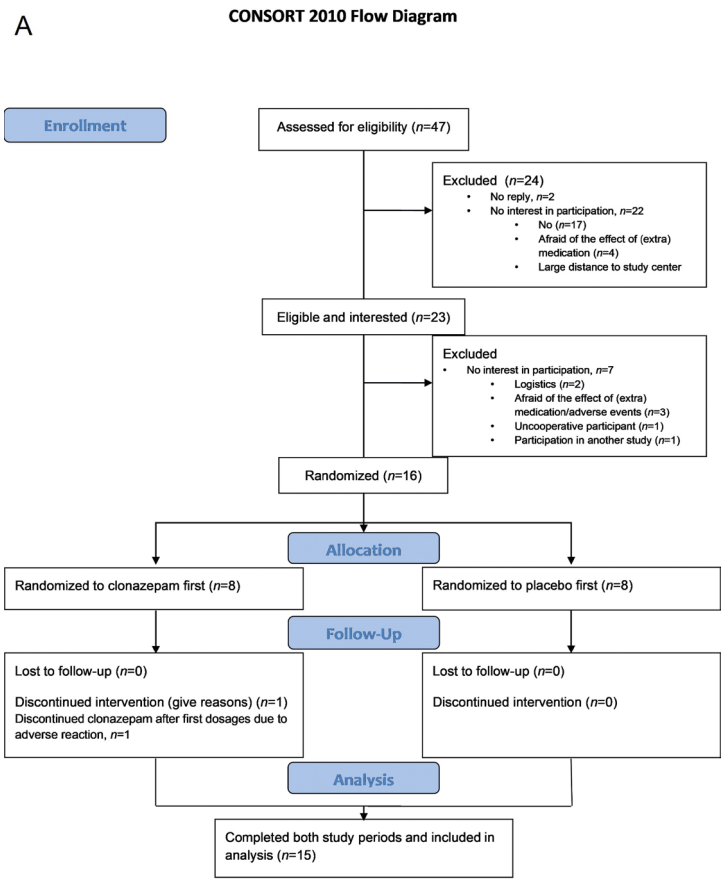
RESULTS

Randomised controlled trial

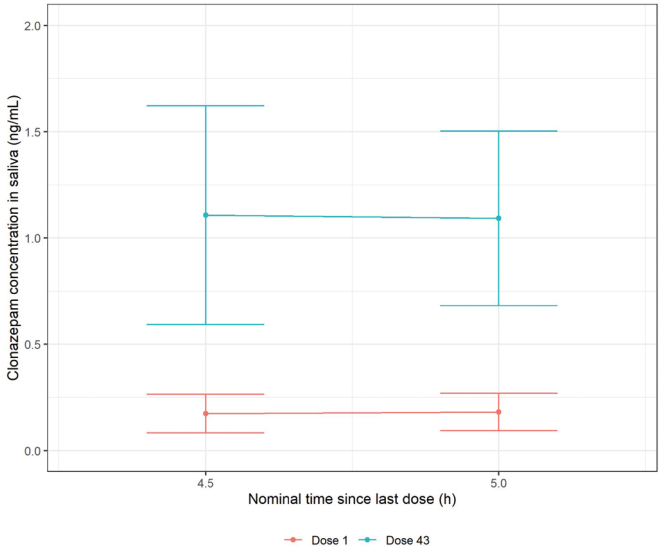
Patients

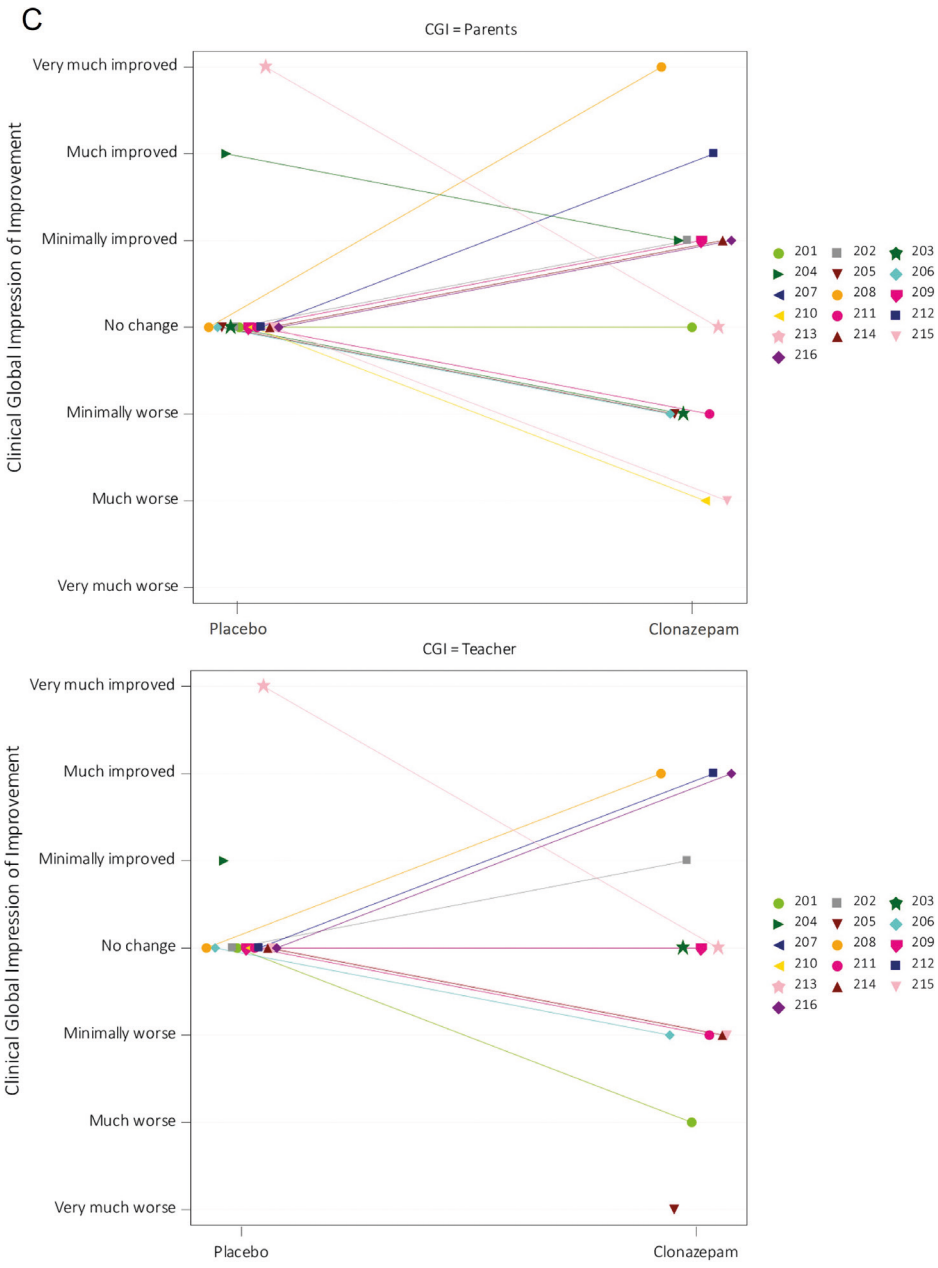
47 parents of patients (figure 2A) with a pathogenic variant in *ARID1B* were approached between May 2021 and January 2022, of which 16 consented to participate with their child (table 1 and online supplemental table S2). The pathogenic *ARID1B* variants identified in these 16 patients are all variants predicted to lead to haploinsufficiency. Of these 16 patients, 11 (69%) were female and the age varied between 6 and 33 years (median: 17). All but one patient had an intellectual disability, one patient had an IQ score within the normal range, but was diagnosed with developmental delays in combination with recurrent otitis media and laryngomalacia. These 16 patients were randomly assigned to receive either clonazepam or placebo, of which 15 completed both study periods. One subject stopped at day 4 of the clonazepam treatment period due to side effects consistent with a paradoxical reaction to clonazepam.

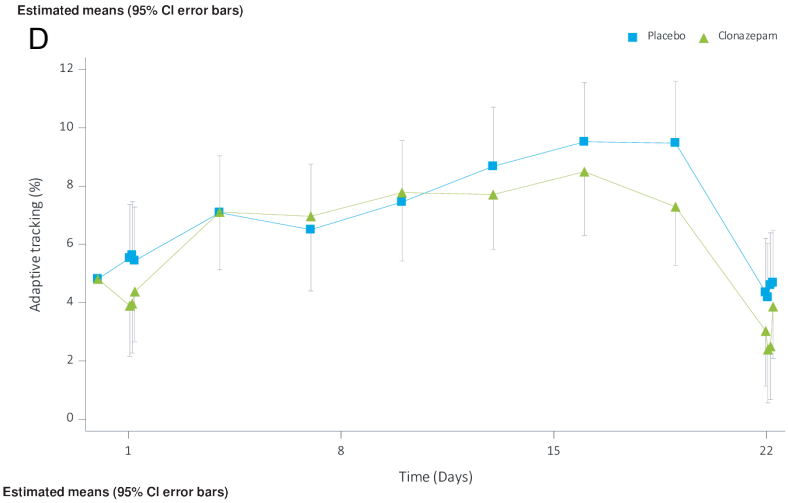
A

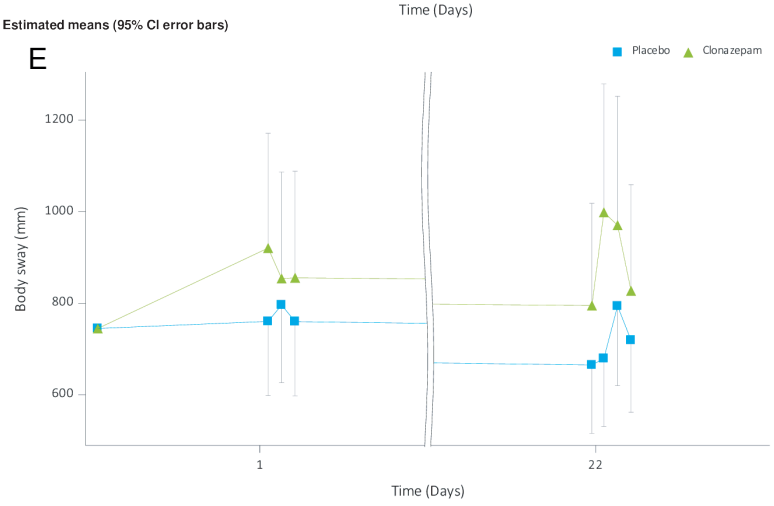

CONSORT 2010 Flow Diagram
Enrollment
Assessed for eligibility (n=47)
Excluded (n=24)
No reply, n=2
No interest in participation, n=22
No (n=17)
Afraid of the effect of (extra) medication (n=4)
Large distance to study center
Eligible and interested (n=23)
Excluded
No interest in participation, n=7
Logistics (n=2)
Afraid of the effect of (extra) medication/adverse events (n=3)
Uncooperative participant (n=1)
Participation in another study (n=1)
Randomized (n=16)
Allocation
Randomized to clonazepam first (n=8)
Randomized to placebo first (n=8)
Follow-Up
Lost to follow-up (n=0)
Discontinued intervention (give reasons) (n=1)
Discontinued clonazepam after first dosages due to adverse reaction, n=1
Lost to follow-up (n=0)
Discontinued intervention (n=0)
Analysis
Completed both study periods and included in analysis (n=15)

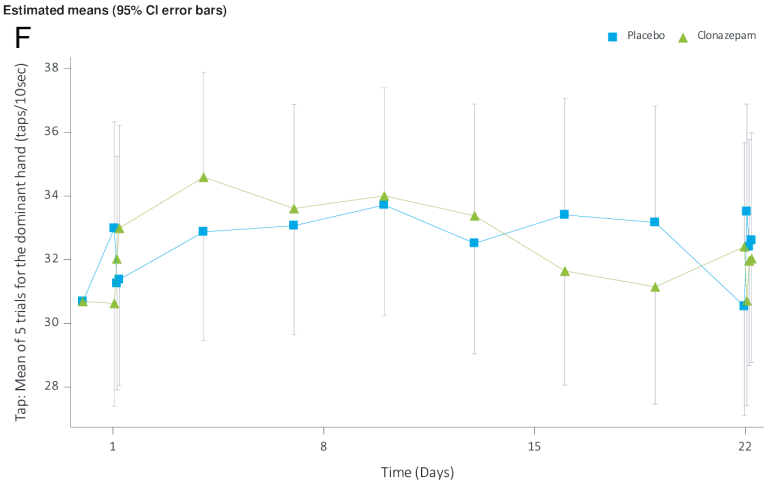
B


Clonazepam concentration in saliva (ng/mL)
2.0
1.5
1.0
0.5
0.0
4.5
5.0
Nominal time since last dose (h)
Dose 1
Dose 43


C
CGI = Parents
Very much improved
Much improved
Minimally improved
No change
Minimally worse
Much worse
Very much worse
Clinical Global Impression of Improvement
Placebo
Clonazepam
201 202 203
204 205 206
207 208 209
210 211 212
213 214 215
216
CGI = Teacher
Very much improved
Much improved
Minimally improved
No change
Minimally worse
Much worse
Very much worse
Clinical Global Impression of Improvement
Placebo
Clonazepam
201 202 203
204 205 206
207 208 209
210 211 212
213 214 215
216


Estimated means (95% CI error bars)
D
Placebo
Clonazepam
Adaptive tracking (%)
0
2
4
6
8
10
12
1
8
15
22
Time (Days)


Estimated means (95% CI error bars)
E
Placebo
Clonazepam
Body sway (mm)
600
800
1000
1200
1
22
Time (Days)


Estimated means (95% CI error bars)
F
Placebo
Clonazepam
Tap: Mean of 5 trials for the dominant hand (taps/10sec)
28
30
32
34
36
38
1
8
15
22
Time (Days)

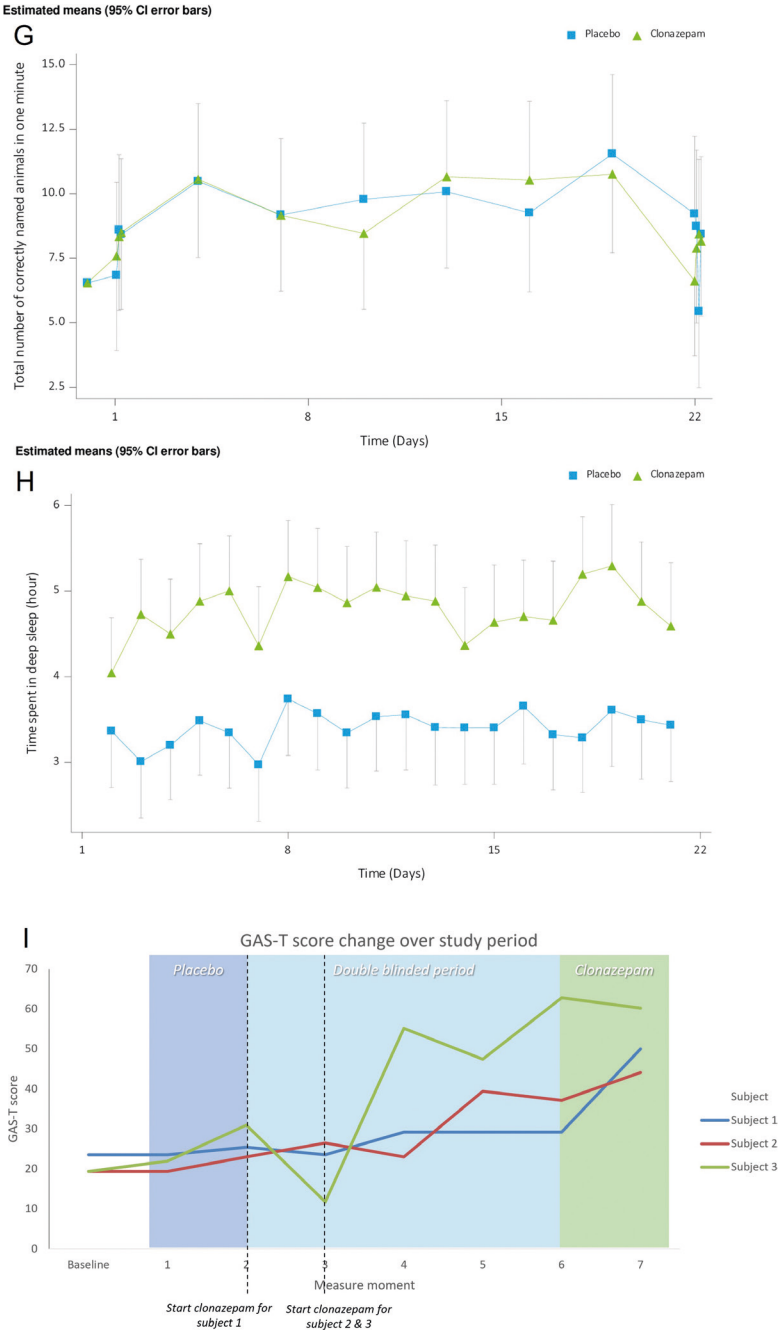


Figure 2. Randomised controlled trial (RCT): (A) Consolidated Standards of Reporting Trials (CONSORT) flow diagram; (**B**) individual concentrations of clonazepam concentration in saliva (ng/mL)—coloured per dose; (**C**) Clinician Global Impression of Improvement (CGI-I); (**D**) adaptive tracking; (**E**) body sway; (**F**) tapping; (**G**) animal fluency; (**H**) deep sleep; (**I**) N-of-1: Goal attainment scaling (GAS)-T change. CFB, Change From average Baseline.

Table 1. Demographics.

Sex	Weight (kg)	Height (cm)	BMI (kg/m²)	CGI-S
Female	65.0	154.0	27.4	Slightly ill
Female	60.5	164.0	22.5	Slightly ill
Male	57.1	164.9	21.0	Minimal ill
Female	55.4	147.7	25.4	Minimal ill
Female	24.0	125.8	15.2	Slightly ill
Female	51.0	150.2	22.6	Slightly ill
Female	28.0	131.6	15.9	Slightly ill
Male	39.0	157.9	15.7	Slightly ill
Female	50.0	152.7	21.2	Minimal ill
Male	36.0	121.7	24.1	Slightly ill
Female	38.6	139.3	19.9	Not ill
Female	68.5	160.8	26.5	Minimal ill
Female	19.1	115.0	14.4	Moderately ill
Female	55.0	167.0	19.0	Slightly ill
Male	74.0	179.5	22.2	Not ill
Male	54.0	170.0	18.7	Moderately ill
Female=11	Avg=52.50	Avg=153.4	Avg=21.12	

Avg average, BMI body mass index, CGI-S, Clinical Global Impression of Severity

PK assessment

Online supplemental table S3 shows the administered clonazepam dosages per subject. Figure 2B shows the individual clonazepam concentrations in saliva. After the first day of dosing, the drug's mean (SD) concentration was 0.17 ng/mL and 0.18 ng/mL (0.09) at 4.5 and 5 hours post-dose, respectively. On day 22, the mean (SD) concentration of the last dose was 1.11 ng/mL (0.52) and 1.10 ng/mL (0.41) at 4.5 and 5 hours post-dose, respectively. The accumulation ratio observed for the patients was approximately 6.1–6.5, indicating significant drug accumulation over time.

Titration adjustments of clonazepam

During the study, dosage adjustments necessitated by side effects were made for five out of 15 patients while taking clonazepam. These adjustments, consistently involving a halving of the dose, were executed on day 6, 7 or 8. In three patients, the dose was not increased to the prior dosing at the parents' request on day 11. No dosage modifications were required for patients while they were on the placebo.

PD assessment

Clinical Global Impression of Improvement

The mean and individual CGI-I values reported by parents and teachers are shown in figure 2C.

CGI-I interviews with parents

Based on CGI-I interviews with parents, out of 15 patients who received a placebo treatment, for 13 patients (87%) parents reported no change, for two patients (13%) parents reported improvement and none reported worsening (online supplemental table S4).

Out of 15 patients who completed clonazepam treatment, two patients (13%) reported no change, seven patients (47%) reported an improvement and six patients (40%) reported worsening in the clonazepam treatment period compared with day 1 pre-dose. Examples of self or parent-reported improvement in the clonazepam group included increased expressive speech (*n*=4), self-reported calmness (*n*=3), experiencing fewer emotional outbursts and/or more susceptible to reason (*n*=2), better concentration (*n*=2), taking more initiative (*n*=1), better at indicating own boundaries (*n*=1) and better sleep (*n*=5). Self-reported or caregiver-reported worsening was linked to adverse events in all cases except for one case (see table 3 and the paragraph 'adverse events'). In this patient, the reported worsening involved an increased preoccupation with herself, more frequent outbursts of emotion directed at her surroundings and greater difficulty engaging with others.

CGI-I interviews with teacher or mentor

In most cases, the CGI-I scores based on interviews with teacher or mentor overlapped with the evaluation with parents. In four cases, CGI-I from both treatment periods was not available. Based on the CGI-I with the teacher or mentor, four patients improved, five patients worsened, and in two patients, no change was observed in the clonazepam treatment period.

Aberrant Behavior Checklist

Another self-reported outcome was the ABC which was separately completed by the parents and teacher or mentor (table 2). The parental ABC indicated a significant increase in hyperactivity score during the clonazepam compared with the placebo treatment period after adjusting for baseline measures.

Table 2. NeuroCart results, Steel HR sleep and ABC questionnaire.

			Least squares mean	
		Treatment P value	**Placebo (*n*=15)**	**Clonazepam (*n*=15)**
NeuroCart	Smooth pursuit	0.64	12.65	12.36
	Body sway	0.16	737.36	885.98
	Tap: mean of 5 trials for the dominant hand (taps/10 s)	0.92	32.58	32.40
	Adaptive tracking (%)	0.10	6.40	5.34
	Animal fluency	0.96	8.9	8.9
Withings Steel HR	Time spent in light sleep (hour)	<0.01	5.10	4.14
	Time spent in deep sleep (hour)	<0.01	3.41	4.79
	Time spent waking up (hour)	0.49	0.05	0.04
	Time spent in total sleep	0.06	8.49	8.92
	Number of times waking up during sleep (count)	<0.01	4.1	2.3
ABC parents	1: Irritability	0.95	6.2	6.3
	2: Lethargy	0.84	11.4	12.1
	3: Stereotypy	0.44	1.6	2.5
	4: Hyperactivity	0.05	7.9	14.3
	5: Inappropriate speech	0.73	1.3	1.4

ABC, Aberrant Behavior Checklist

NeuroCart and Withings Steel HR

No differences between the clonazepam and placebo period were observed on the NeuroCart test (table 2, figure 2D–H). Saccadic eye movement measurements were successful for only two patients on both study periods' first and last days. Therefore, no analyses were performed on these data. The Steel HR watch measurements showed that compared with the placebo period, patients during the clonazepam period spent more time in deep sleep and less time in light sleep. However, their total sleep time was comparable between study periods (table 2).

Safety and tolerability/adverse events

In total, for nine patients (56%), more than one adverse event was reported (table 3) while taking clonazepam, compared with six patients (38%) taking a placebo. During the clonazepam treatment, fatigue was the most frequently reported adverse event (*n*=8), followed by somnolence (*n*=2). For five of the six patients with a worsening CGI-I score, adverse events linked to clonazepam were reported (fatigue, *n*=5; nausea, *n*=1). A paradoxical drug reaction was observed in one patient on day 4 during clonazepam treatment. After the first increase in dosing on day 4, this patient displayed agitated

behaviour at home and school, restlessness and sleepiness. In consultation with the parents, it was decided to exclude this patient from the study early.

Table 3. Summary of number of subjects in the RCT with TEAEs by treatment, SOC, PT and severity.

System Organ Class/Preferred Term	Clonazepam (*n*=16)			Placebo (*n*=15)		
	Mild	Moderate	Severe	Mild	Moderate	Severe
Any events	**10**	**2**	**–**	**7**	**–**	**–**
Gastrointestinal disorders						
Upper abdominal pain	1	1	–	–	–	–
Nausea	–	–	–	1	–	–
General disorders and administration site conditions						
Fatigue	8	–	–	2	–	–
Feeling drunk	1	–	–	–	–	–
Paradoxical drug reaction	–	1	–	–	–	–
Infections and infestations						
Coronavirus infection	1	–	–	–	–	–
Nasopharyngitis	1	–	–	1	–	–
Injury, poisoning and procedural complications						
Head injury	1	–	–	–	–	–
Musculoskeletal and connective tissue disorders						
Bursitis	1	–	–	–	–	–
Nervous system disorders						
Balance disorder	2	–	–	–	–	–
Disturbance in attention	1	–	–	–	–	–
Dysarthria	1	–	–	–	–	–
Headache	–	–	–	–	–	–
Slow response to stimuli	1	–	–	–	–	–
Somnolence	1	–	–	–	–	–
Psychiatric disorders	1			2		
Bradyphrenia		–	–	1	–	–
Disinhibition	1	–	–	–	–	–
Impulse control disorder	1	–	–	–	–	–
Inappropriate affect		–	–	1	–	–
Respiratory, thoracic and mediastinal disorders	2	–	–	–	–	–
Oropharyngeal pain	1	–	–	–	–	–
Rhinorrhoea	1	–	–	–	–	–

PT, Preferred Term, RCT randomised controlled trial, SOC, System Organ Class; TEAE, treatment-emergent adverse event

There was no observed dose-dependent treatment effect for any of the parameters, and there were no differences in treatment effects based on phenotype severity, sex or age.

N-of-1

Three parents of the 15 patients expressed interest in continuing clonazepam treatment after completing the trial (see online supplemental table S6).

Phase I

Several patients or parents reported a sedative effect of clonazepam. Therefore, to identify the minimal effective dosage of clonazepam, we added a titration phase I in the N-of-1 design. The optimal dose at the end of the open-label titration phase I differed from the dose used during the RCT for all patients. Online supplemental table S7 provides an overview of the domains for which specific goals were defined.

Phase II

All three enrolled patients completed this phase and their data were analysed. Online supplemental table S8 and figure 2I show the change in the GAS-T score during phase 2. While all patients exhibited improvement on the GAS-T score, only subject 3 showed a distinct improvement in the GAS-T score, specifically on switching from placebo to clonazepam during the double-blinded period. Further exploration of the indicated effect using the GAS scoring and parental feedback indicated an effect on sleep in two patients and a small effect on behaviour in one patient. No relevant changes at the individual level or group level were measured using the behaviour questionnaires ABC and CBCL (see online supplemental table S9), and no adverse events, side effects or paradoxical reactions were reported.

Treatment evaluation and informed decision-making

At the end of phase II, the results were discussed with the parents of patient 1 and her physician in intellectual disability medicine. The parents expressed the need to take some time to consider the results, and care was transferred to the patient's own physician, who was also present during the evaluation. For patient 2, who demonstrated an observed effect only on sleep, the parents decided to discontinue clonazepam treatment. Similarly, for subject 3, who showed an effect on sleep and a potentially small effect on behaviour, the parents, together with their child's physician, concluded that clonazepam did not significantly improve the child's functioning and chose to discontinue its use.

DISCUSSION

This randomised, double-blinded, placebo-controlled crossover study, followed by N-of-1 design, is the first to investigate the effects of clonazepam in patients with an *ARID1B*-related intellectual disability. In our studies, aside from the effect on sleep observed by parents, no significant improvements in functioning were observed for the clonazepam group compared with the placebo group. A large heterogeneity in treatment response was present, with nearly half of the parents and patients themselves reporting notable improvement, and nearly half reporting a worsening in function. The reported improvements were further investigated in a subset of our population using an N-of-1 study. Th results of this study confirmed the effect on sleep and showed no significant improvements on the other domains, illustrating the difficulties when assessing drug effects without objective clinical endpoints and indirectly through parents/caregivers in this highly variable population.

Ultimately, 16 patients started the study, and 15 patients completed the full study. Therefore, the power to detect a 1 point difference on the CGI-I was not 80% (at *n*=20) but 67% (at *n*=15). Given the complete lack of signal and the subsequent results from the N-of-1 trials we consider it unlikely that including five additional patients would have led to a different conclusion.

The titration phase of the N-of-1 trials resulted in an observed effect with slightly lower dosages for all three patients compared with the dose they received during the crossover study. We consider it unlikely that the lack of effect on other domains, except sleep, in the N-of-1 trials can be explained by the lower dosage in these studies. Since the parents did observe an effect of clonazepam in phase I, and a clear effect of clonazepam on sleep, similar to the crossover study, was measured in phase II. We therefore consider it unlikely that a higher dose would have resulted in better efficacy.

Previous research in an Arid1b$^{+/-}$ mouse model showed promising results for clonazepam compared with saline treatment.[11] In the Arid1b$^{+/-}$ mouse model, clonazepam improves novel object recognition, social novelty and anxiety.[11] Although it is always critical to approach extrapolation from mouse models to human patients with caution, the observed improvements in the mouse model suggested potential overlap with positive effects reported in patients, such as improved concentration, emotion regulation and calmness. However, in this trial, these encouraging results from the mouse model did not translate into positive outcomes for humans.

CGI and ABC

Variable treatment effects of clonazepam were observed in the *ARID1B* patient population based on the parental CGI-I results, with seven patients improving, six worsening and two showing no change. This variability may be attributed to population heterogeneity,[8] but could also be explained by random variations or parents who are highly motivated to observe drug effects. Reported side effects, such as fatigue and somnolence, align with clonazepam's known central nervous system effects and likely contributed to the reported worsening in the functioning of a subset of patients. Lowering the dose may have improved the benefit-risk ratio for this group.

An unexpected finding was that parents reported a higher hyperactivity domain score on the ABC questionnaire during the clonazepam period compared with placebo (borderline statistically significant). This domain, known as 'Hyperactivity/Noncompliance', measures levels of hyperactivity and disobedience. Given clonazepam's sedative properties, it would be expected that patients' activity levels during clonazepam treatment would be lower or similar to those during placebo treatment. However, the observed agitation and restlessness in one patient may have been present in a milder form in more patients. We speculate that clonazepam more frequently displays paradoxical effects in patients with *ARID1B*-related intellectual disability.

NeuroCart

Contrary to what we had anticipated, we could not show an improved performance for clonazepam compared with placebo treatment on NeuroCart tests.[16 17] The reasons behind the lack of a significant diminished effect remain uncertain. The most likely explanation is that clonazepam has no effect on these tests. Alternatively, the side effect of clonazepam could offset the expected improvement in psychomotor tests. The possibility that this patient population was unable to perform the test accurately could be another consideration, though this seems unlikely given that prior research has shown that these tests could be administered to this patient group.[15]

Goal attainment scaling

We used GAS in our N-of-1 design to evaluate individual treatment effects and offer personalised assessments. GAS is sensitive to change, and its patient-centered focus increases relevancy and the likelihood of detecting relevant differences in treatment outcomes. However, it is important to consider the subjectivity of goal setting and rating, the lack of standardisation, difficulty of choosing goals that are specific, measurable, achievable, relevant and time-bound, and limited generalisability. As demonstrated in our N-of-1 design, although GAS provides meaningful and personalised outcomes, the potential bias towards scoring progress[24] should be interpreted cautiously,

considering the subjective nature of the defined goals and the need to discuss the clinical significance of the measured differences with parents or caregivers. Additionally, in the evaluation process, it is important to assess the overall effect of the medication and consider the possibility of treatment effects not captured by the predefined goals.[24]

Strengths and limitations

This is the first clinical study for *ARID1B* and/or CSS. Strengths of this study include the use of the innovative trial design combining a crossover design that allows each patient to serve as their control, with an optional additional placebo-controlled RCT, resulting in an N-of-1 design, a sizeable cohort, a low dropout rate, repeated measurement design and the involvement of parents in the study design.

Our study also has several limitations. Only patients able to handle multiple testing on the study days were included. This may have introduced a selection bias towards less affected patients. An additional issue complicating determination of a clinically meaningful change is the reported placebo response of 20% in autism clinical trials.[25] This report indicates that studies involving children and adolescents, employing caregiver ratings and ensuring minimal biases and allocation concealment risks were associated with the highest risk of placebo response. In addition, caregiver or self-reported measures may introduce expectation bias. This may have led to an accentuation of this placebo effect by overshadowing subtle changes that may happen in the treatment period.[26] We tried to limit the expectation bias by using a double-blinded design of both studies, and the low number of patients that improved during the placebo period of the RCT makes it probable that these effects were minimal in this study. On the other hand, the results of the N-of-1 study of subject 1 (figure 2I) clearly demonstrate the presence of expectation bias during the last unblinded period. Another potential limitation of this study is that the endpoints were very general. We chose CGI since it was unknown what kind of effect clonazepam might have in *ARID1B* patients, which complicates the choice of appropriate, specific study endpoints. On the other hand, these broad CGI-I evaluations made it possible to catch any potential effect that may have been missed with a more specific endpoint and enabled us to estimate the treatment effect of clonazepam compared with placebo. Selection bias may be present in the selection of patients for our N-of-1 trial. Only patients whose parents approached the study team about the continuation of clonazepam treatment were included, creating a bias towards patients with a potential positive effect of clonazepam during the crossover study. However, we would expect this bias to be in favour of clonazepam effect. Furthermore, we cannot exclude that the dose titration phase of the N-of-1 might have resulted in patients or the parents or caretaker recognising the effects of clonazepam in the double-blinded phase II. However, we felt this disadvantage

was outweighed by the positive effects of allowing a thorough determination of the optimal dose, and given the results it is unlikely that this effect changed the outcome.

Conventional randomised clinical trial and N-of-1 studies

The conventional/parallel-group RCTs and the N-of-1 design are distinct but complementary approaches in clinical research, each with unique strengths. RCTs provide a rigorous framework, minimising biases and allowing for unbiased assessment of treatment efficacy on a larger scale. They offer more objective results through randomisation and blinding. However, RCTs may not capture individual variations and personalised responses to interventions.

On the other hand, the N-of-1 design offers a personalised and patient-centred approach, monitoring individual responses to treatments over time. They capture nuances and fluctuations in symptoms, providing valuable insights at the individual level. However, the smaller sample size in N-of-1 studies and subjective endpoints may increase the risk of type 1 errors.[24 27]

To confirm clonazepam's effects in individual patients and limit expectation bias, our study combined a conventional RCT with a crossover design, followed by an additional crossover in three patients, resulting in an N-of-1 trial. This approach maximised the strengths of both methods while considering time and cost constraints.

In our case, the N-of-1 trials further explored observed effects from the RCT. Alternatively, starting with N-of-1 initially could have allowed for a detailed understanding of intervention effects and informed the design of a more targeted RCT with specific endpoints.

CONCLUSIONS

Altogether, the results of our RCT and following N-of-1 trials suggest that clonazepam has no effect on *ARID1B* patients. Some effects were observed in a subset (n=6) of patients in the RCT. The N-of-1s in three of these patients only replicated the effect on sleep. Our study may serve as an example of navigating the complexities of research in rare diseases and contribute to improving the care and outcomes of patients with *ARID1B*-related disorders by addressing the challenges of validating biomarkers and identifying parental expectation bias. We demonstrated how the N-of-1 trial is effectively employed to explore and validate observed effects, enabling us to differentiate between genuine treatment effects and expectation bias.

SUPPLEMENTARY MATERIAL

Supplementary Information (463.7KB, pdf)

ACKNOWLEDGEMENTS

The authors thank the patients, their families and caregivers, the research assistants who conducted the tests and the data management officers and support staff at CHDR who contributed to this study.

Funding: ZonMW (Project No 10140261910002).

Patient consent for publication: Consent obtained from parent(s)/guardian(s)

Ethics approval: The protocol (NL71395.056.19) was reviewed and approved by the Beoordeling Ethiek Biomedisch Onderzoek (BEBO) Foundation Review Board (Assen, the Netherlands). Informed consent was received and archived from all participants or their representative(s) as required by the Institutional Review Board (IRB). Individual data were deidentified using subject numbers. Participants gave informed consent to participate in the study before taking part.

The N-of-1 compassionate use of clonazepam in *ARID1B* patients was exempted from review by the IRB of the LUMC, Leiden, the Netherlands, on 18 February 18th, 2022.

DATA AVAILABILITY STATEMENT

The data generated during this study are available via https://www.isrctn.com/ISRCTN11225608.

REFERENCES

1. Hoyer J, Ekici AB, Endele S, et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am J Hum Genet. 2012;90:565–72. doi: 10.1016/j.ajhg.2012.02.007.
2. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. Nature New Biol. 2017;542:433–8. doi: 10.1038/nature21062.
3. Gillentine MA, Wang T, Eichler EE. Estimating the Prevalence of De Novo Monogenic Neurodevelopmental Disorders from Large Cohort Studies. Biomedicines. 2022;10:2865. doi: 10.3390/biomedicines10112865.
4. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet. 2012;44:376–8. doi: 10.1038/ng.2219.
5. Santen GWE, Aten E, Vulto-van Silfhout AT, et al. Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum Mutat. 2013;34:1519–28. doi: 10.1002/humu.22394.
6. Wieczorek D, Bögershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum Mol Genet. 2013;22:5121–35. doi: 10.1093/hmg/ddt366.
7. Chen C-A, Lattier J, Zhu W, et al. Retrospective analysis of a clinical exome sequencing cohort reveals the mutational spectrum and identifies candidate disease-associated loci for BAFopathies. Genet Med. 2022;24:364–73. doi: 10.1016/j.gim.2021.09.017.
8. van der Sluijs PJ, Jansen S, Vergano SA, et al. The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin-Siris syndrome. Genet Med. 2019;21:1295–307. doi: 10.1038/s41436-018-0330-z.
9. Santen GWE, Aten E, Sun Y, et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. Nat Genet. 2012;44:379–80. doi: 10.1038/ng.2217.
10. Celen C, Chuang J-C, Luo X, et al. Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. Elife. 2017;6:e25730. doi: 10.7554/eLife.25730.
11. Jung E-M, Moffat JJ, Liu J, et al. Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. Nat Neurosci. 2017;20:1694–707. doi: 10.1038/s41593-017-0013-0.
12. Shibutani M, Horii T, Shoji H, et al. Arid1b Haploinsufficiency Causes Abnormal Brain Gene Expression and Autism-Related Behaviors in Mice. Int J Mol Sci. 2017;18:1872. doi: 10.3390/ijms18091872.
13. Moffat JJ, Jung E-M, Ka M, et al. The role of ARID1B, a BAF chromatin remodeling complex subunit, in neural development and behavior. Prog Neuropsychopharmacol Biol Psychiatry. 2019;89:30–8. doi: 10.1016/j.pnpbp.2018.08.021.
14. Kruizinga MD, Zuiker RGJA, Bergmann KR, et al. Population pharmacokinetics of clonazepam in saliva and plasma: Steps towards noninvasive pharmacokinetic studies in vulnerable populations. Br J Clin Pharmacol. 2022;88:2236–45. doi: 10.1111/bcp.15152.
15. Kruizinga MD, Zuiker RGJA, Sali E, et al. Finding Suitable Clinical Endpoints for a Potential Treatment of a Rare Genetic Disease: the Case of ARID1B. Neurotherapeutics. 2020;17:1300–10. doi: 10.1007/s13311-020-00868-9.
16. Zuiker RGJA, Chen X, Østerberg O, et al. NS11821, a partial subtype-selective GABAA agonist, elicits selective effects on the central nervous system in randomized controlled trial with healthy subjects. J Psychopharmacol . 2016;30:253–62. doi: 10.1177/0269881115620435.

17. Groeneveld GJ, Hay JL, Van Gerven JM. Measuring blood-brain barrier penetration using the NeuroCart, a CNS test battery. Drug Discov Today Technol. 2016;20:27–34. doi: 10.1016/j.ddtec.2016.07.004.
18. Guy W. Clinical global impression. Assess man for Psychopharmacol. 1976:217–22.
19. Aman MG, Singh NN, Stewart AW, et al. The aberrant behavior checklist: a behavior rating scale for the assessment of treatment effects. Am J Ment Defic. 1985;89:485–91.
20. Kiresuk TJ, Sherman RE. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Ment Health J. 1968;4:443–53. doi: 10.1007/BF01530764.
21. Dekkers K, Viet E, Eilander H, et al. Handleiding goal attainment scaling (gas) https://revant.nl/dynamic/media/1/documents/Gas/handleiding_GAS.pdf Available.
22. Kiresuk TJ, Smith A, Cardillo JE. Goal attainment scaling: applications, theory, and measurement. L. Erlbaum Associates. 1994
23. World Health Organization International classification of functioning, disability and health: children & youth version: ICF-CY. 2007. [-Jun-2022]. http://www.who.int/classifications/icf/ Available. Accessed.
24. Steenbeek D, Ketelaar M, Galama K, et al. Goal attainment scaling in paediatric rehabilitation: a critical review of the literature. Dev Med Child Neurol. 2007;49:550–6. doi: 10.1111/j.1469-8749.2007.00550.x.
25. Siafis S, Çıray O, Schneider-Thoma J, et al. Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis. Mol Autism. 2020;11:66. doi: 10.1186/s13229-020-00372-z.
26. McConachie H, Parr JR, Glod M, et al. Systematic review of tools to measure outcomes for young children with autism spectrum disorder. Health Technol Assess. 2015;19:1–506. doi: 10.3310/hta19410.
27. Urach S, Gaasterland C, Posch M, et al. Statistical analysis of Goal Attainment Scaling endpoints in randomised trials. Stat Methods Med Res. 2019;28:1893–910. doi: 10.1177/0962280218777896.

Chapter 9

Discussion

INTRODUCTION

This thesis was started only a few years after *ARID1B* had been recognized as an intellectual disability (ID) gene[1,2] and was associated with Coffin-Siris syndrome(CSS)[3,4]. These discoveries resulted in the publication of the first large genotype-phenotype correlation studies of CSS[3-6]. These studies[3-6] identified pathogenic variants in BAF-complex related genes (such as *ARID1B*) in 55 to 71% of in total 180 patients with a clinical diagnosis of CSS. The application of next-generation sequencing (NGS) was a pivotal development enabling the identification of the CSS-causing genes. Since then, the increasing use of NGS in diagnostics has uncovered numerous variants in BAF-complex-related genes, some of which are clearly pathogenic. This has made it possible to diagnose these patients with a BAFopathy, but an increasing number of variants are considered variants of uncertain significance (VUS). Each patient has their own specific phenotype and the clinician needs to assess whether the pathogenic variant can explain the phenotype or whether further studies are needed. This has highlighted the need for new genotype-phenotype studies to expand our understanding of the BAFopathy phenotypes. Simultaneously, the identification of more patients with a VUS in BAFopathy genes has created a demand for tools to assist in understanding these uncertain variants.

Another clinical challenge in the field of BAFopathies is the lack of treatment options. In order to find viable therapies, one needs to understand the etiology of the disease, find suitable disease models and develop a framework to study the therapy in patients. Clinical trials for rare diseases face challenges such as small sample sizes and phenotypic variability. To address this, it is crucial to develop disease-specific, and perhaps even patient-specific endpoints, and take the feasibility of a trial design into account.

The first aim or this thesis was to provide further insight into the phenotypes associated with pathogenic variants in BAFopathy genes and to discover new (epi) genetic phenotypes associated with BAFopathies. Our next aim was to create a model to study mechanisms of disease underlying CSS and to look for a possibility of a therapy for patients with a pathogenic variant in *ARID1B*.

INSIGHTS INTO CLINICAL AND EPIGENETIC PHENOTYPES

An important first step in a genotype-phenotype study is a thorough assessment of the identified variants. Below we describe the different tools to aid interpretation of the variants, followed by a discussion on the phenotype-genotype correlations in various genotype subgroups.

Interpretation of variants

In 2015, the American College of Medical Genetics (ACMG) formulated a framework to create a standardized approach to aid variant interpretation[7]. In this thesis we have applied the criteria formulated within this framework together with Human Phenotype Ontology (HPO)-term based analysis (Chapters 5-6), facial photograph analysis (Chapters 4-6), and DNA methylation studies (Chapters 4-6) to interpret variants.

HPO-terms

The description of the phenotype of a patient consists of objectively defined and measurable features, such as specific congenital anomalies and ID. Other features such as dysmorphic or facial features, are usually regarded as more subjective because their assessment depends on the clinical experience and training of the observer and can vary between populations. HPO-terms can be used to systematically assess the first type of features and are usually used to characterize patient cohorts, as demonstrated by the DECIPHER database[8] and the Human Disease Genes website series (https://humandiseasegenes.nl)[9]. HPO-terms can also be used to identify and compare patient subgroups[10]. An advantage of HPO-based comparison is that multiple features can be assessed simultaneously. However, meticulous phenotyping remains crucial in HPO-based studies, as it is not possible to report an 'unknown' status of a feature in most analyses.

HPO-terms can also be used to aid VUS interpretation. If patients with pathogenic variants have a phenotype different from that of controls matched for age, sex and ethnicity, HPO-term analysis can be used to determine whether a patient with a VUS has a phenotype that closely resembles those with pathogenic variants. We applied this method in Chapter 6 and 7 using the PhenoScore algorithm[10]. In Chapter 6 we investigated whole gene duplications of *ARID1A* and *ARID1B*. Using HPO-term based analyses we showed that indeed the phenotype of patients with a whole gene duplication of *ARID1A* and *ARID1B* have a distinct phenotype, different from that of controls and from typical CSS patients. We also showed that these two groups of patients are different from each other. Next, we used HPO-based analyses to investigate if patients with a CNV of uncertain significance clustered with patients with a whole gene duplication of

either *ARID1A* or *ARID1B*. Of the five cases with a CNV of uncertain significance three clustered with duplication cases. In our final interpretation, we confirmed this for two cases but the third, based on the predicted effect of the variant and DNA methylation results, was later classified as having CSS. This shows that although PhenoScore can be a useful tool to aid the interpretation of variants, the results should be interpreted with care. Especially if a patient's phenotype differs from the known pathogenic group, it does not necessarily confirm that the variant is benign, as the phenotype may be broader than is currently recognized. Therefore, the results of a phenotype-based analysis should always be considered alongside other evidence for or against the variant's pathogenicity.

Facial photograph analysis

One way to objectively analyze facial features is by examining facial photographs. Facial photograph analyses are being increasingly applied in diagnostic and research settings. In the diagnostic setting, two systems, Face2Gene (FDNA Inc., Boston, MA, USA; https://www.face2gene.com) and GestaltMatcher[11], are used for such analyses. In both a photo of a patient is uploaded and is compared with photos of patients with the disorders in the database. The PhenoScore algorithm[10] can be used to perform facial photograph analysis in a research setting. PhenoScore can use the data of facial photographs, HPO-terms or both to identify whether a patient group clusters separately from matched controls with neurodevelopmental delay. Conveniently, if a patient group clusters separately from controls based on HPO-terms and/or facial features, features differing between patients and controls can be retrieved; these can be HPO-terms or distinct facial features. If a group of patients with a pathogenic variant in a particular gene, clusters separately from controls, data from a single patient (e.g. patient with a VUS in the same gene) can be entered in the analysis and used to see whether this patient's data and/or facial photograph clusters with patients or with controls. In this manner HPO-based analyses and facial photograph analyses may also be used in a diagnostic setting.

We applied facial photograph analyses in Chapter 4-6. In Chapter 4 we confirmed that *ARID1B* patients have a distinct facial phenotype, distinguishable from controls. We also showed the importance of taking the ages at which photos were taken into account since photos of older patients were harder to distinguish from controls. In Chapter 5, we verified these results by analyzing recent photos of adult patients as well as their childhood photos. Photos from all age groups clustered separately from controls. However, in the highest age group, clustering was less significant. This group also had the fewest photos available.

In Chapter 5 we showed that patients with pathogenic variants in exon 1 of *ARID1B* are generally less severely affected. We also showed that facial photographs from eight patients with a pathogenic variant in exon 1 did not cluster separately from controls. To investigate whether this lack of clustering was linked to their milder phenotype or could be explained by the limited number of available photos, we compared them to patients with variants in other exons. No significant differences were found, suggesting that the absence of a distinct clustering pattern is primarily due to the small number of photos available.

Therefore, if facial analysis is implemented to assist with variant interpretation, we recommend using childhood photographs, especially since most available photos are from below the age of 18 years. This approach can provide a more reliable basis for accurate diagnosis and variant interpretation.

We went on to explore the use of facial analysis in VUS interpretation. In Chapter 6 we showed that clustering of VUS based on facial photographs produced results similar to clustering based on HPO-terms.

DNA methylation

In 2018 Aref-Eshghi *et al*[12] published an article identifying a specific DNA methylation pattern in the blood of CSS patients. They also showed that this pattern could be used to interpret VUS in *ARID1B*. In Chapters 4-6 we have used DNA methylation analyses to aid VUS interpretation. We also discovered specific (sub)episignatures, enabling us to distinguish between patient subgroups. In Chapter 4 and 5, we used DNA methylation to interpret inherited *ARID1B* variants. In Chapter 6, we discovered an episignature for *ARID1A* whole-gene duplications and identified a new microduplication syndrome caused by *ARID1B* whole-gene duplications.

While DNA methylation analysis holds promise for diagnostic applications, there are several challenges and limitations to consider. These include technical variability (including significant batch effects), sample heterogeneity, and the need for robust validation in large cohorts to establish its usefulness in a clinical setting. DNA methylation analysis can be used as a diagnostic tool, particularly when a sample of a VUS clusters with a specific disease. However, an intermediate or a negative DNA methylation clustering result of a VUS requires cautious interpretation. While an intermediate clustering may suggest an intermediate phenotype[13], it could also signify a benign variant; a negative clustering result does not rule out pathogenicity since an alternative episignature could be present, as we and others have shown[14].

Variant interpretation in the future

Several other approaches can be used to aid variant interpretation. For example, RNA-sequencing can measure the impact of the variant on the transcript in the tissue from which the RNA is obtained. RNA-seq can reveal how a variant impacts the splicing of the gene or the downstream consequences of the variant. Another approach to assess the effect of a variant on cellular phenotypes is to use a combination of RNA-seq, ChiP-seq, and ATAC-seq. Applying these techniques to patient and control cell lines reveals how chromatin structure, gene expression, and regulatory mechanisms are affected in cases with pathogenic BAF complex variants. If these studies are used to interpret VUS, samples from an unaffected control as well as from the patient with the pathogenic variant in the same gene as the VUS would be needed.

The investigation of the effects of pathogenic variants in BAF complex subunits could lead to the identification of genes or disorders with a similar functional effect as BAFopathies. For instance, *ARID2* has been linked to RASopathies[15].

While investigating variant pathogenicity and function of the BAF-complex, new episignatures are likely to emerge as we have shown in Chapter 6, in which we report new episignatures. While a substantial number of episignatures exist, the timing and impact of DNA methylation changes in most disorders are still unclear. Disease models offer a promising method to investigate these and many other aspects.

Genotype-Phenotype correlations

Using all of the above-mentioned methods in combination with the ACMG-criteria, most known variants in BAFopathy genes can be classified and patients with (likely) pathogenic variants are eligible for inclusion in the clinical overviews of genotype-phenotype correlation studies. In our studies we have expanded the phenotypic spectrum of CSS (Chapters 2-5), phenotypes associated with *ARID1A* whole gene duplications (Chapters 6) and we described, for the first time, the *ARID1B* duplication phenotype (Chapter 6). For the sake of accuracy, we divided the phenotype data in our overview into specific subgroups. Below we discuss the different subgroups and provide a discussion on biases that may arise in genotype-phenotype correlation studies.

Genotype Subgroups

Differences in the phenotype of patients can depend on a) differences between genes, b) differences between variants, and c) on the location of the variant.

Gene-specific subgroups

In Chapter 3 we reported the prenatal phenotype of CSS. In our overview we divided the phenotype into gene-specific subgroups because earlier research[3-6] had shown phenotypic differences in disease severity between, for example, *ARID1B* and *SMARCB1* patients. We observed phenotypic differences between patient groups and differences in the distribution of prenatal versus postnatal cases over the different genes. Examples of gene-specific features are the increased risk of developing hepatoblastoma in *ARID1A* patients[16] and the hypogonadotropic hypogonadism in several *SOX11* patients[17]. These findings underline the importance of creating gene-specific subgroups when analyzing phenotypic data for disorders associated with multiple genes such as CSS.

Variant-specific subgroups

In Chapter 6 we show that whole gene duplications of *ARID1A* and *ARID1B* lead to distinct phenotypes with their own episignature. These phenotypes are different from those of patients with CSS with a pathogenic variant or a deletion in *ARID1A* or *ARID1B*. The data can, therefore, not be pooled in one overview; one needs to make variant-specific subgroups. A similar example is *SMARCC2*, another CSS-associated gene. Missense variants in *SMARCC2* lead to a CSS phenotype, and the typical BAFopathy episignature[18,19]. On the other hand, truncating variants in *SMARCC2* lead to a milder phenotype with a reduced penetrance[20].

Location-specific subgroups

In the phenotype table in Chapter 5, we divided patients into subgroups based not only on the affected gene, but also on the location of the variant[21]. In Chapter 4 we had already noted that all inherited non-mosaic pathogenic variants clustered in exon 1 and hypothesized that exon 1 variants may result in a milder phenotype compared to other variants. This hypothesis was confirmed in Chapter 5.

Thus, in CSS cohorts, differences can be present between patients with variants in different genes, between patients with a different type of variants in the same gene and between patients with variants at a different location on the transcript of the same gene. When assessing a genotype-phenotype correlation it is therefore important to keep in mind the genotype and to determine whether the appropriate subgroup analysis has been done.

Findings of published studies that disregard the above-mentioned aspects should be interpreted with caution. Unfortunately, several such studies have been published, which can be grouped as 1) CSS cohorts that do not report genetic variants identified in their patients[22,23], because interpretation of variants is not always straightforward,

could have included patients with false-positive variants, and 2) CSS cohorts that do not distinguish between patients with pathogenic variants in different genes[23,24].

Bias in genotype-phenotype studies

An important aspect to consider in genotype-phenotype studies is bias. There are several types of bias present in our studies and in published BAFopathy cohorts. Recognizing these biases is crucial when using the data in clinical practice. Common biases in genotype-phenotype studies are ascertainment bias and phenotype reporting bias.

Ascertainment bias

There are different types of ascertainment bias. Below we discuss four examples: 1) the overrepresentation of syndromic patients, 2) the tendency to include more severe patients, 3) postnatal ascertainment bias and 4) age-related bias.

Overrepresentation of syndromic patients

We explored the bias caused by overrepresentation of syndromic patients in Chapter 2, where we investigated whether patients with a clinical suspicion of CSS before genetic testing are phenotypically different from those without an *a priori* suspicion. We discovered that, aside from CSS-specific features, such as short fifth finger, dysmorphic features and hypertrichosis, these patient groups were comparable and should, therefore, be treated similarly. Thus, we showed that although there was ascertainment bias regarding CSS-related features, it did not lead to a biased estimate of other features.

Inclusion of more severe patients

Ascertainment bias can also arise due to the tendency to include more severely affected patients. In Chapter 3, we identified a distinct part of the CSS spectrum by focusing only on patients with prenatal anomalies. This approach revealed new aspects of the clinical spectrum. However, not all CSS patients exhibit prenatal anomalies, and our findings are, therefore, not representative of the entire group.

Postnatal ascertainment bias

Postnatal ascertain bias arises because only individuals who survive the prenatal period are included, potentially underrepresenting severe cases that result in miscarriage, stillbirth, or termination due to congenital anomalies detected during pregnancy. To tackle this issue, we reported in Chapter 3 44 CSS cases with a prenatal presentation with the aim of defining the prenatal CSS phenotype and to assess whether features emerging during pregnancy were overlooked in postnatal cohorts. In this chapter, we

expanded the CSS spectrum to include hydrocephalus, hypoplastic left heart syndrome, persistent left vena cava, diaphragmatic hernia, renal agenesis, intrauterine growth restriction, and anal anomalies.

Age-related bias

Age-related ascertainment bias can lead to an underreporting of clinical features. In pediatric cohorts, for instance, age-related ascertainment bias may prevent the reporting of features that develop later in life. For example, the cohort we described in Chapter 2 consisted mostly of pediatric patients; we, therefore, hypothesized that age-related bias affects this cohort.

To address this, in Chapter 5 we described a cohort of adult *ARID1B* patients, including both index cases and affected family members to illustrate the development and self-sustainability of adult *ARID1B* patients. We used survival plots to report features that develop later in life. This approach explicitly takes into account the fact that a phenotype may not yet be fully developed, and therefore gives a better estimate of the prevalence of these features than a cross-sectional percentage. Using this approach, we found that seizures develop in circa 55% of patients in this cohort; this is more than the circa 40% we previously reported in Chapter 2. We also assessed whether features developing later in life were missed in the previous (primarily pediatric) cohort in Chapter 2, and identified several new features in this adult population including loss of skills and recurrent patella luxation. The adult cohort also allowed us to show that patients with a variant in exon 1 appeared to have a higher level of self-sustainability. Despite efforts to overcome age-related bias, ascertainment bias remains in our adult cohort, as mildly affected individuals are less likely to undergo genetic testing than those with more severe symptoms. When assessing a genotype-phenotype study it is therefore essential to consider how the cohorts were established and what the reasons for genetic testing in specific individuals were.

Phenotype reporting bias

Phenotype reporting bias arises when a clinical feature is either missing from the questionnaire or is not assessed in all patients, making it impossible to calculate the true prevalence. In order not to miss unrecognized features we added open fields for responses in our online questionnaires. In Chapter 2, we identified phenotype reporting bias in the DECIPHER database. For example, at that time ID was reported in 20% and seizures in 5.6% of *ARID1B* patients in the DECIPHER database, while in our cohort, these rates were 99.2% and 27.5%, respectively. To avoid reporting such an underestimation we added the option 'unknown' as an answer to our 'yes/no' questions. This may lead to an overestimation if the absence of the features is not recognized in clinical charts

and results in the scoring 'unknown' rather than 'no' or 'absent' for the given feature. Therefore, larger genotype-phenotype studies with more patients and more precise phenotype assessment are needed to refine the prevalence estimates in the phenotype tables of this thesis.

METHODS FOR STUDYING MECHANISMS OF DISEASE AND POSSIBLE TREATMENT

The identification of a disease gene may be the starting point for the search of a disorder-specific treatment, but in-depth investigation of the involved pathways is an essential next step. Disease models offer a promising approach to explore these pathways and the mechanisms in greater depth. While this type of research is usually far from the patients' bedside its results may aid variant interpretation (e.g., by replicating successful experiments with cell lines containing VUS instead of a pathogenic variant). This could enhance our understanding of the disorder (Chapter 7) and potentially lead to targeted therapeutic interventions (Chapter 8). Below is a closer look at disease models, disease-specific therapies and possible treatment options.

Disease models

In Chapters 8 and 9 we elaborated on the findings in the *Arid1b*-haploinsufficient mouse model described by Jung *et al*[25]. In this model the authors identified increased apoptosis and decreased proliferation of inhibitory GABAergic interneurons. Inhibitory GABAergic interneurons regulate brain activity by reducing neuronal excitability, maintaining balance in neural circuits, and preventing excessive excitation. Based on their finding of fewer GABAergic interneurons in the brains of mice, we created an *in vitro* model to investigate the effects of genetic variants in CSS on neuronal differentiation and on the epigenetic landscapes in induced pluripotent stem cells (iPSC) and neuronal progenitor cells (NPCs). To create this model, we used two *ARID1B*, three *SMARCB1* patient cell lines and four control cell lines (Chapter 7).

The results of our cell model suggest that the effect of pathogenic CSS variants is tissue dependent. We found, as expected, a difference in methylation between iPSCs and NPCs. This difference was present in both patient and control samples. Subsequently, we wondered whether the distinct episignature in blood of CSS patients compared to controls is present in CSS patient derived iPSCs and/or NPCs. We did not find the BAFopathy episignature from blood in NPCs of the patients. This makes it likely that the effect of pathogenic CSS variants on methylation is either tissue dependent or occurs later in the differentiation process.

We also observed morphological differences between patient and control cell lines as early as day 16 of differentiation. This raises a critical question about how the brain develops in patients, if phenotypic changes emerge so early. To address this, future research should focus on differentiating patient-derived iPSCs into mature GABAergic interneurons to assess how these early differences influence neuronal function. Although we have repeated our differentiation process twice, additional experiments with a larger set of patient-derived cell lines are essential to replicate these findings. Such studies will help determine whether the observed differences are due to the pathogenic variant itself or to genetic heterogeneity between cell lines.

Another approach in the future, to answer the question about how the brain develops in patients, if phenotypic changes emerge so early, is by generating an organoid of the brain (as described in Chapter 7). Subsequently such an organoid could be used to study the effect of a potential therapy.

Disease-specific therapies

Disease models can be used to investigate the underlying pathology of a disease. They are also of great value in drug development. Two types of disease-specific therapies are a) therapies targeting the cause or the genetic variant (i.e. gene therapy), b) therapies targeting affected pathways or processes of the genetic variant.

Therapies targeting the genetic variant

Currently, no genetic therapies are available for CSS or other BAFopathies, though research is actively being conducted by several groups. Each type of genetic therapy presents unique challenges. Some challenges are broader and apply to many conditions, such as selecting the optimal method of administering the treatment (e.g., tablet, droplets, or injection into the bloodstream or spinal fluid) and ensuring that the therapy reaches the appropriate tissue and target (e.g., cell nucleus, DNA, or RNA). Others are gene-specific, requiring tailored strategies to address the particular characteristics of the gene or condition being treated.

Even if a genetic therapy is developed that would, for example, successfully restore *ARID1B* expression in the cells of an *ARID1B* haploinsufficient patient without causing overexpression, there are still aspects that need further investigation. One key aspect is whether correcting the effects of a pathogenic variant in a brain that is already fully developed, can result in meaningful improvements of the phenotype.

Therapies targeting the affected pathway or processes

Aside from therapies targeting the pathogenic variants, there are also therapies targeting the affected pathways or processes of the pathogenic variant. An example of this type of disease-specific therapy is illustrated by Jung *et al*[25] who used this approach in their ARID1B haploinsufficient mouse model. Upon observing that these mice exhibited fewer GABAergic interneurons due to decreased proliferation and increased apoptosis, they administered clonazepam (a $GABA_A$ receptor positive allosteric modulator) and observed that this rescued part of the phenotype in mice.

Investigating treatment options in patients

Considering these effects of clonazepam in ARID1B haploinsufficient mice (Jung *et al.*[25]), the next logical step was to evaluate its impact in *ARID1B* patients. One advantage of repurposing clonazepam as a targeted therapy is that it is an FDA-approved, well-studied, and affordable drug. However, conducting RCTs in vulnerable populations, such as children and individuals with ID, presents unique challenges. These include population heterogeneity, ethical considerations, and difficulties in obtaining informed consent. Children and individuals with an ID often have complex needs and may require personalized interventions and goals, making it challenging to implement standardized treatments and study endpoints. Moreover, obtaining consent from these populations can be challenging due to their limited capacity to understand and communicate. In such scenarios, N-of-1 trials offer a promising solution by allowing for individualized treatment evaluations. These N-of-1 trials involve repeated measurements of an outcome in the same individual over multiple treatment periods. This enables clinicians to assess treatment efficacy at the individual level. N-of-1 trials can provide valuable insights into treatment effectiveness in children and individuals with ID. N-of-1 trials can serve as a precursor to larger RCTs by assessing the feasibility and potential efficacy of interventions. Alternatively, N-of-1 trials can complement RCTs by providing supplementary data to further elucidate treatment effects observed in larger cohorts (e.g. Chapter 8).

We investigated the effect of clonazepam in *ARID1B* patients (Chapter 8) in a randomized control trial (RCT), which was a double-blind, placebo-controlled and two-way crossover study followed by an N-of-1 design. In our study an effect of clonazepam was observed in a subset of patients in the RCT, however, the N-of-1 trials only replicated the effect on sleep. We thus concluded that clonazepam has no additional therapeutic value in *ARID1B* patients. In our clinical trial we also show the feasibility and strength of combining a conventional RCT with N-of-1 studies, which we believe to be a robust method for conducting therapeutic research in populations with ID.

CONCLUSIONS

We have addressed key challenges in BAFopathy research by broadening the clinical spectrum through genotype-phenotype correlation studies, discovering two new episignatures, and interpreting VUS in multiple BAFopathy genes. Additionally, we have outlined a model for studying the pathophysiology of pathogenic variants in CSS and we have reported the first clinical trial in the most prevalent BAFopathy, *ARID1B*-related ID.

Future work remains necessary to further investigate the phenotypes of different BAFopathies, including studying the underlying pathophysiology, improving interpretation of variants, broadening clinical phenotypes, discovering new BAFopathy genes, and exploring potential treatment strategies.

The work described in this thesis also highlights various essential aspects of clinical genetics, by focusing on the interpretation of genetic variants and, where applicable, providing recommendations for screening. Moreover, it illustrates an example of an aspect that is becoming increasingly relevant: disease-specific therapies. It shows the expanding scope of clinical genetics, from diagnosis and counselling to actively contributing to therapeutic strategies aimed at addressing the underlying genetic mechanisms of diseases.

REFERENCES

1. Hoyer J, Ekici AB, Endele S, et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am J Hum Genet*. Mar 09 2012;90(3):565-72. doi:10.1016/j.ajhg.2012.02.007
2. Halgren C, Kjaergaard S, Bak M, et al. Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin Genet*. Sep 2012;82(3):248-55. doi:10.1111/j.1399-0004.2011.01755.x
3. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet*. Mar 18 2012;44(4):376-8. doi:10.1038/ng.2219
4. Santen GW, Aten E, Vulto-van Silfhout AT, et al. Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum Mutat*. Nov 2013;34(11):1519-28. doi:10.1002/humu.22394
5. Wieczorek D, Bogershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet*. Dec 20 2013;22(25):5121-35. doi:10.1093/hmg/ddt366
6. Tsurusaki Y, Okamoto N, Ohashi H, et al. Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clin Genet*. Jun 2014;85(6):548-54. doi:10.1111/cge.12225
7. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. May 2015;17(5):405-24. doi:10.1038/gim.2015.30
8. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet*. Apr 2009;84(4):524-33. doi:10.1016/j.ajhg.2009.03.010
9. Dingemans AJM, Stremmelaar DE, Vissers L, et al. Human disease genes website series: An international, open and dynamic library for up-to-date clinical information. *Am J Med Genet A*. Apr 2021;185(4):1039-1046. doi:10.1002/ajmg.a.62057
10. Dingemans AJM, Hinne M, Truijen KMG, et al. PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. *Nat Genet*. Aug 7 2023;doi:10.1038/s41588-023-01469-w
11. Hsieh TC, Bar-Haim A, Moosa S, et al. GestaltMatcher facilitates rare disease matching using facial phenotype descriptors. *Nat Genet*. Mar 2022;54(3):349-357. doi:10.1038/s41588-021-01010-x
12. Aref-Eshghi E, Bend EG, Hood RL, et al. BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin-Siris and Nicolaides-Baraitser syndromes. *Nat Commun*. Nov 20 2018;9(1):4885. doi:10.1038/s41467-018-07193-y
13. Awamleh Z, Goodman S, Choufani S, Weksberg R. DNA methylation signatures for chromatinopathies: current challenges and future applications. *Hum Genet*. Apr 6 2023;doi:10.1007/s00439-023-02544-2
14. Levy MA, McConkey H, Kerkhof J, et al. Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. *HGG Adv*. Jan 13 2022;3(1):100075. doi:10.1016/j.xhgg.2021.100075
15. Kang E, Kang M, Ju Y, et al. Association between ARID2 and RAS-MAPK pathway in intellectual disability and short stature. *J Med Genet*. Nov 2021;58(11):767-777. doi:10.1136/jmedgenet-2020-107111

16. van der Sluijs PJ, Vergano SA, Roeder ER, Jongmans MCJ, Santen GWE. Recommending revised hepatoblastoma surveillance in children with a pathogenic ARID1A variant. Reply to "Cancer in ARID1A-Coffin-Siris syndrome: Review and report of a child with hepatoblastoma" by Carcamo et al. 2022. *Eur J Med Genet*. Feb 2023;66(2):104694. doi:10.1016/j.ejmg.2022.104694
17. Al-Jawahiri R, Foroutan A, Kerkhof J, et al. SOX11 variants cause a neurodevelopmental disorder with infrequent ocular malformations and hypogonadotropic hypogonadism and with distinct DNA methylation profile. *Genet Med*. Jun 2022;24(6):1261-1273. doi:10.1016/j.gim.2022.02.013
18. Machol K, Rousseau J, Ehresmann S, et al. Expanding the Spectrum of BAF-Related Disorders: De Novo Variants in SMARCC2 Cause a Syndrome with Intellectual Disability and Developmental Delay. *Am J Hum Genet*. Jan 3 2019;104(1):164-178. doi:10.1016/j.ajhg.2018.11.007
19. Sadikovic B, Levy MA, Kerkhof J, et al. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet Med*. Jun 2021;23(6):1065-1074. doi:10.1038/s41436-020-01096-4
20. Bosch E, Popp B, Guse E, et al. Elucidating the clinical and molecular spectrum of SMARCC2-associated NDD in a cohort of 65 affected individuals. *Genet Med*. Nov 2023;25(11):100950. doi:10.1016/j.gim.2023.100950
21. Beaumont RN, Hawkes G, Gunning AC, Wright CF. Clustering of predicted loss-of-function variants in genes linked with monogenic disease can explain incomplete penetrance. *Genome Med*. Apr 26 2024;16(1):64. doi:10.1186/s13073-024-01333-4
22. Mannino EA, Miyawaki H, Santen G, Schrier Vergano SA. First data from a parent-reported registry of 81 individuals with Coffin-Siris syndrome: Natural history and management recommendations. *Am J Med Genet A*. Oct 1 2018;doi:10.1002/ajmg.a.40471
23. McCague EA, Lamichhane R, Holt N, Schrier Vergano SA. Growth charts for individuals with Coffin-Siris syndrome. *Am J Med Genet A*. Oct 2020;182(10):2253-2262. doi:10.1002/ajmg.a.61823
24. Schmetz A, Ludecke HJ, Surowy H, et al. Delineation of the adult phenotype of Coffin-Siris syndrome in 35 individuals. *Hum Genet*. Jan 2024;143(1):71-84. doi:10.1007/s00439-023-02622-5
25. Jung EM, Moffat JJ, Liu J, Dravid SM, Gurumurthy CB, Kim WY. Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. *Nat Neurosci*. Dec 2017;20(12):1694-1707. doi:10.1038/s41593-017-0013-0

A

Addendum

SUMMARY

BAFopathies are a group of neurodevelopmental disorders caused by pathogenic variants in genes encoding subunits of the BAF (BRG1/BRM-associated factor) chromatin remodeling complex. Coffin-Siris syndrome (CSS) is the most well-known and prevalent BAFopathy. CSS patients typically have an intellectual disability (ID), distinctive facial features, and hypoplasia of the fifth digits and/or nail(s). CSS patients also have a specific DNA methylation pattern in their blood compared to controls. These patterns are especially useful to discriminate between pathogenic and harmless variants when a variant of uncertain significance (VUS) is identified in a CSS-associated gene.

Due to the widespread application of next-generation sequencing (NGS) in diagnostics, an increasing number of variants are being identified in BAFopathy-related genes. Many of these are clearly pathogenic, enabling clinicians to diagnose affected individuals. However, an important proportion of variants are classified as VUS, posing a significant challenge. In clinical practice VUS can be resolved by a thorough comparison between patient and literature phenotype, and/or by additional tools. This highlights the need for solid genotype-phenotype studies to broaden our understanding of the BAFopathy-associated phenotype as well as for tools, such as episignatures and facial analysis, to aid in variant interpretation.

Another pressing clinical challenge in the field of BAFopathies is the lack of disease specific treatments. Developing viable therapies requires a deeper understanding of disease etiology, the establishment of relevant cellular or animal models, and a clear framework to evaluate treatment efficacy in patients. Clinical trials for rare diseases like BAFopathies are particularly complex, facing obstacles such as small patient populations, phenotypic variability, and high expectations from families. Therefore, it is essential to design trials with disease-specific—and potentially even patient-specific—outcomes, while carefully considering the feasibility of the study design.

The aims of the first part of this thesis were to provide further insight into the phenotypic spectrum associated with pathogenic variants in BAFopathy genes and to discover new (epi)genetic features linked to these disorders. The aims of the second part of this thesis were to develop an *in vitro* model to study the disease mechanisms underlying CSS and to explore potential therapeutic options, particularly for patients with *ARID1B*-related disorders.

In **chapter 2**, we compared two cohorts of patients with a pathogenic variant in *ARID1B*: I) those clinically diagnosed with or a suspicion of Coffin-Siris syndrome (*ARID1B*-CSS)

before genetic testing and; II) those identified through exome sequencing (*ARID1B*-ID). Our study of 143 patients showed that while CSS-associated dysmorphic features were more common in the *ARID1B*-CSS group, overall, there were only minor differences between the two groups. These findings show that *ARID1B*-related disorders form a broad phenotypic spectrum rather than two distinct entities, emphasizing that all *ARID1B* patients should receive similar clinical management. We also showed that common methods for reporting phenotypes may miss important gene-related features. This highlights the need for thorough data collection methods.

In **chapter 3** we focused on the prenatal presentation of CSS. We collected data on 44 patients with prenatal anomalies and a pathogenic variant in a CSS gene. Frequently observed prenatal features included hydrocephalus, agenesis of the corpus callosum, and congenital heart defects. Notably, *ARID1A* variants were much more common in this prenatal cohort compared to postnatal cohorts. Our findings provide valuable new insights into the prenatal spectrum of CSS, improving variant interpretation in fetal diagnostics.

In **chapter 4**, we tackled the challenge of interpreting inherited uncertain loss-of-function (LoF) variants in ARID1B. We studied 12 families with variants that were hard to classify because they were inherited or their inheritance was unknown. Using DNA methylation episignatures and clinical assessments, we were able to reclassify most of these variants. Our results showed that some pathogenic variants exhibit highly variable expression, even within families, and we hypothesized about an alternative start site in *ARID1B*. We suggested an update for the ACMG-criteria (genetic variant classification guidelines) to incorporate epigenetic profiling and facial analysis.

In **chapter 5** we examined the natural history of *ARID1B*-related disorder in adults. We collected data on 87 adult patients, revealing a broad spectrum ranging from severe ID to normal IQ scores. We also identified new age-related features such as loss of skill and recurrent patellar luxation. The broad spectrum was also reflected in the self-sustainability of *ARID1B* patients. Facial photo analysis showed consistent clustering separate from controls across ages, reinforcing the distinctiveness of the facial features of *ARID1B* patients. We also investigated the genotype-phenotype correlation we first reported in chapter 4. In chapter 4 we identified a number of cases with an inherited pathogenic variant in *ARID1B.* Most of these cases had a variant in the first exon of *ARID1B*. In this chapter, we confirmed that patients with a variant in the first exon of *ARID1B* generally exhibit a milder phenotype.

In **chapter 6**, we shifted focus to whole gene duplications in *ARID1A* and *ARID1B*, an area with previously limited data. We analyzed 29 patients with duplications. Both whole gene duplications of *ARID1A* and *ARID1B* result in ID phenotypes, which differ from CSS. Patients with an *ARID1A* duplication tend to have more severe symptoms compared to *ARID1B* patients. DNA methylation analyses revealed specific episignatures, supporting the idea that these duplications represent a distinct type of BAFopathy. We successfully reclassified several duplications previously of uncertain significance as pathogenic, illustrating the diagnostic utility of epigenetic profiling.

In **chapter 7** we described the development of an *in vitro* disease model using patient-derived induced pluripotent stem cells (iPSCs) with *ARID1B* or *SMARCB1* variants. Differentiation into neuronal cells and subsequent analyses confirmed the model's suitability for studying CSS variant effects on neuronal differentiation, offering a platform to investigate neuronal differentiation in CSS and potential therapies.

Finally, in **chapter 8**, we reported on a clinical trial evaluating clonazepam in *ARID1B* patients. Despite promising preclinical data in mouse models, our randomized controlled crossover trial and subsequent N-of-1 studies did not show convincing therapeutic benefit in patients. Our study highlighted the feasibility of combining conventional RCT and N-of-1 studies for therapeutic studies in populations with intellectual disabilities, distinguishing real treatment effects from expectation bias.

This thesis enhances the understanding of the genetic and phenotypic landscape of BAFopathies, particularly Coffin-Siris syndrome and *ARID1B*-related disorders, and lays the foundation for precision medicine approaches in this field. By integrating detailed genotype-phenotype correlations, novel episignature discovery, disease modeling, and the first therapeutic trial in BAFopathy patients, we deliver new scientific, diagnostic and clinical tools to support personalized care.

SAMENVATTING

BAFopathieën vormen een groep neurologische ontwikkelingsstoornissen veroorzaakt door pathogene varianten in genen die coderen voor subunits van het BAF (BRG1/BRM-associated factor) chromatine-remodelleringscomplex. Coffin-Siris syndroom (CSS) is de bekendste en meest voorkomende BAFopathie. Patiënten met CSS hebben meestal een verstandelijke beperking (VB), karakteristieke gelaatskenmerken en korte vijfde vingers en/of kleine of afwezige nagels. Daarnaast hebben CSS-patiënten een specifiek DNA-methylatiepatroon (episignature) in hun bloed vergeleken met controles. Deze patronen zijn vooral nuttig om wanneer een variant van onbekende betekenis (VUS) wordt gevonden in een CSS-geassocieerd gen onderscheid te maken tussen ziekteveroorzakende (pathogene) en onschuldige varianten.

Door het brede gebruik van *next-generation sequencing* in de diagnostiek worden er steeds meer varianten gevonden in BAFopathie-gerelateerde genen. Veel hiervan zijn duidelijk pathogeen, waardoor clinici diagnoses kunnen stellen. Echter, een aanzienlijk deel wordt als VUS geclassificeerd, wat een belangrijke uitdaging vormt. In de kliniek kan zo een VUS verder worden geclassificeerd door vergelijking van het patiëntfenotype met de literatuur en/of door inzetten van aanvullende tools. Dit onderstreept de noodzaak van gedegen genotype-fenotype studies en hulpmiddelen zoals episignatures en gezichtsanalyses om genetische varianten beter te interpreteren.

Een andere belangrijke klinische uitdaging is het ontbreken van ziektespecifieke behandelingen voor BAFopathieën. De ontwikkeling van therapieën vereist een beter begrip van de ziekteoorzaak, geschikte cel- of diermodellen en een duidelijke studie opzet om de effectiviteit van behandelingen te beoordelen. Klinische studies bij zeldzame ziekten zoals BAFopathieën zijn uitdagend door de kleine patiënten aantallen, variabele symptomen en hoge verwachtingen van families. Het is daarom essentieel om studies zorgvuldig te ontwerpen, met ziekte specifieke en mogelijk zelfs patiënt specifieke uitkomstmaten.

Het doel van het eerste deel van dit proefschrift (hoofdstuk 2-6) is om meer inzicht te geven in het fenotypisch spectrum van pathogene varianten in BAFopathie-genen en nieuwe (epi)genetische kenmerken van deze aandoeningen te ontdekken. In het tweede deel (hoofdstuk 7-8) richtten we ons op het ontwikkelen van een *in vitro* model om de ziekteprocessen bij CSS te bestuderen en om een mogelijke behandelingsoptie, voor patiënten met een *ARID1B*-gerelateerde aandoening, te verkennen.

In **hoofdstuk 2** vergeleken we twee cohorten van patiënten met een pathogene variant in *ARID1B*: I) patiënten die klinisch als CSS waren gediagnosticeerd of verdacht werden van CSS vóór genetische testen (*ARID1B*-CSS), en II) patiënten geïdentificeerd via exoom *sequencing* (*ARID1B*-ID). Onze studie met 143 patiënten toonde aan dat hoewel CSS-kenmerken vaker voorkwamen in de *ARID1B*-CSS groep, de verschillen tussen beide groepen klein waren. Dit bevestigt dat het fenotype van *ARID1B* patiënten uit een breed overlappend spectrum bestaat in plaats van twee verschillende entiteiten en dat alle *ARID1B*-patiënten vergelijkbare zorg verdienen. We toonden ook aan dat gangbare methoden voor het rapporteren van fenotypes belangrijke gen-gerelateerde kenmerken kunnen missen. Dit benadrukt de noodzaak van grondige gegevensverzamelingsmethoden.

In **hoofdstuk 3** onderzochten we de prenatale presentatie van CSS. We verzamelden gegevens van 44 patiënten met prenatale afwijkingen en een pathogene variant in een CSS-gen. Veelvoorkomende prenatale kenmerken waren hydrocefalie, agenesie van het corpus callosum en hartafwijkingen. Opvallend was dat *ARID1A*-varianten veel vaker voorkwamen in dit prenatale cohort. Onze bevindingen bieden waardevolle nieuwe inzichten in het prenatale spectrum van CSS, welke bijdragen aan de interpretatie van varianten in de prenatale setting.

In **hoofdstuk 4** gingen we de uitdaging aan om overgeërfde *loss-of-function* (LoF) varianten in *ARID1B* te interpreteren. We bestudeerden 12 families met varianten die moeilijk te classificeren waren omdat ze overgeërfd waren of omdat de overerving onbekend was. Met behulp van DNA-methylatie analyse en klinische beoordelingen, waren we in staat om de meeste van deze varianten te herclassificeren. Onze resultaten lieten zien dat sommige pathogene varianten sterk variabele expressie vertonen, zelfs binnen families. We hypotheseerden over een alternatief startcodon in *ARID1B* en we stelden een update voor van de ACMG-criteria (richtlijnen voor classificatie van genetische varianten) om DNA methylatie analyse en gezichtsanalyse hierin op te nemen.

In **hoofdstuk 5** bestudeerden we het natuurlijke beloop van *ARID1B*-gerelateede aandoening bij volwassenen. We verzamelden gegevens van 87 volwassen patiënten en vonden een breed spectrum, van ernstige VB tot normale IQ-scores. We ontdekten ook nieuwe leeftijdsgebonden kenmerken, zoals verlies van bepaalde vaardigheden en terugkerende patella luxaties. De variabiliteit binnen de aandoening werd ook weerspiegeld in de mate van zelfstandigheid van de patiënten. Gezichtsanalyse toonde aan dat *ARID1B*-patiënten duidelijke, leeftijdsonafhankelijke gelaatskenmerken hebben. Ook onderzochten we de genotype-fenotype correlatie die we in hoofdstuk 4 hadden

beschreven. Daar zagen we dat overgeërfde pathogene varianten vaak in exon 1 van *ARID1B* lagen en bevestigden we dat patiënten met zo'n variant meestal een milder fenotype hebben.

In **hoofdstuk 6** richtten we ons op duplicaties van het hele gen in *ARID1A* en *ARID1B*, waarover eerder weinig bekend was. We analyseerden 29 patiënten met zulke duplicaties. Zowel *ARID1A*- als *ARID1B*-duplicaties veroorzaken verstandelijke beperkingen. De patiënten verschillen van CSS patiënten. Patiënten met een *ARID1A*-duplicatie hadden meestal ernstigere ontwikkelings- en groeiachterstand dan *ARID1B*-patiënten. DNA-methylatie analyse onthulde specifieke episignatures voor de patiënten met een duplicatie van het hele *ARID1A*-gen, wat aangeeft dat deze duplicaties een aparte soort BAFopathie vormen. We konden meerdere voorheen duplicaties met onzekere betekenis nu als pathogeen herclassificeren.

In **hoofdstuk 7** beschreven we de ontwikkeling van een *in vitro* ziekte model met geïnduceerde pluripotente stamcellen (iPSCs) van patiënten met *ARID1B*- of *SMARCB1*-varianten. Differentiatie naar neurale voorlopercellen bevestigde de geschiktheid van dit model om effecten van CSS-varianten op neuronale ontwikkeling te bestuderen. Dit biedt een platform voor verder onderzoek naar pathofysiologie van CSS.

Tot slot beschreven we in **hoofdstuk 8** een klinische trial met clonazepam bij *ARID1B*-patiënten. Onze klinische trial bestond uit een gerandomiseerde gecontroleerde cross-over studie (RCT) aangevuld met 3 N=1 studies. Een N=1 studie is een RCT uitgevoerd bij 1 patiënt met gepersonaliseerde uitkomstmaten. Ondanks veelbelovende preklinische data bij muizen, toonde onze studie geen overtuigend effect van clonazepam bij patiënten. Onze studie benadrukte de haalbaarheid van het combineren van conventionele RCT en N=1 studies voor therapeutische studies in populaties met verstandelijke beperkingen, waarbij echte behandelingseffecten worden onderscheiden van verwachtingsbias.

Dit proefschrift vergroot het begrip van het genetische en fenotypische landschap van BAFopathieën, met name het Coffin-Siris syndroom en *ARID1B*-gerelateerde aandoeningen, en legt de basis voor *precision medicine* benaderingen op dit gebied. Door het integreren van gedetailleerde genotype-fenotype correlaties, het ontdekken van nieuwe episignaturen, ziektemodellering en de eerste therapeutische trial in BAFopathie patiënten, leveren we nieuwe wetenschappelijke, diagnostische en klinische hulpmiddelen om gepersonaliseerde zorg te ondersteunen.

LIST OF PUBLICATIONS

Houdayer C, Rooney K, van der Laan L, Bris C, Alders M, Bahr A, Barcia G, Battault C, Begemann A, Bonneau D, Bonnevalle A, Boughalem A, Bourges A, Bournez M, Bruel AL, Buhas D, Carallis F, Cogné B, Cormier-Daire V, Delanne J, Demaret T, Denommé-Pichon AS, Désir J, Dubourg C, Fradin M, Geneviève D, Goel H, Goldenberg A, Gripp KW, Guichet A, Guimier A, Jacquinet A, Keren B, Legoff L, Levy MA, McConkey H, Mendelsohn BA, Mignot C, Milon V, Nizon M, Oneda B, Pasquier L, Patat O, Philippe C, Procaccio V, Procopio R, Prouteau C, Rambaud T, Rauch A, Relator R, Rondeau S, Santen GWE, Schleit J, Sorlin A, Steindl K, Tedder M, Tessarech M, Mau-Them FT, Trost D, **van der Sluijs PJ,** Vincent M, Whalen S, Thauvin-Robinet C, Isidor B, Sadikovic B, Vitobello A, Colin E.
ARID2-related disorder: further delineation of the clinical phenotype of 27 novel individuals and description of an epigenetic signature.
Eur J Hum Genet. 2025 Mar 5.

van der Sluijs PJ, Safai Pour K, Berends CL, Kruizinga MD, Müller AR, van Eeghen AM, Rodríguez-Girondo M, Juachon MJ, Steenbeek D, Cohen AF, Zuiker RGJA, Santen GWE.
Clonazepam repurposing in ARID1B patients through conventional RCT and N-of-1 trials: an experimental strategy for orphan disease development.
J Med Genet. 2025 Feb 26;62(3):210-218.

van der Sluijs PJ, Gösgens M, Dingemans AJM, Striano P, Riva A, Mignot C, Faudet A, Vasileiou G, Walther M, Schrier Vergano SA, Alders M, Alkuraya FS, Alorainy I, Alsaif HS, Anderlid B, Bache I, van Beek I, Blanluet M, van Bon BW, Brunet T, Brunner H, Carriero ML, Charles P, Chatron N, Coccia E, Dubourg C, Earl RK, Eichler EE, Faivre L, Foulds N, Graziano C, Guerrot AM, Hashem MO, Heide S, Heron D, Hickey SE, Hopman SMJ, Kattentidt-Mouravieva A, Kerkhof J, Klein Wassink-Ruiter JS, Kurtz-Nelson EC, Kušíková K, Kvarnung M, Lecoquierre F, Leszinski GS, Loberti L, Magoulas PL, Mari F, Maystadt I, Merla G, Milunsky JM, Moortgat S, Nicolas G, Leary MO', Odent S, Ozmore JR, Parbhoo K, Pfundt R, Piccione M, Pinto AM, Popp B, Putoux A, Rehm HL, Reis A, Renieri A, Rosenfeld JA, Rossi M, Salzano E, Saugier-Veber P, Seri M, Severi G, Sonmez FM, Strobl-Wildemann G, Stuurman KE, Uctepe E, Van Esch H, Vitetta G, de Vries BBA, Wahl D, Wang T, Zacher P, Heitink KR, Ropers FG, Steenbeek D, Rybak T, Santen GWE.
ARID1B-related disorder in 87 adults: Natural history and self-sustainability.
Genet Med Open. 2024 Jul 23;2:101873.

van der Sluijs PJ, Moutton S, Dingemans AJM, Weis D, Levy MA, Boycott KM, Arberas C, Baldassarri M, Beneteau C, Brusco A, Coutton C, Dabir T, Dentici ML, Devriendt K, Faivre L, van Haelst MM, Jizi K, Kempers MJ, Kerkhof J, Kharbanda M, Lachlan K, Marle N,

McConkey H, Mencarelli MA, Mowat D, Niceta M, Nicolas C, Novelli A, Orlando V, Pichon O, Rankin J, Relator R, Ropers FG, Rosenfeld JA, Sachdev R, Sandaradura SA, Shukarova-Angelovska E, Steenbeek D, Tartaglia M, Tedder MA, Trajkova S, Winer N, Woods J, de Vries BBA, Sadikovic B, Alders M, Santen GWE.
Microduplications of ARID1A and ARID1B cause a novel clinical and epigenetic distinct BAFopathy.
Genet Med. 2025 Jan;27(1):101283.

Dingemans AJM, Jansen S, van Reeuwijk J, de Leeuw N, Pfundt R, Schuurs-Hoeijmakers J, van Bon BW, Marcelis C, Ockeloen CW, Willemsen M, **van der Sluijs PJ**, Santen GWE, Kooy RF, Vulto-van Silfhout AT, Kleefstra T, Koolen DA, Vissers LELM, de Vries BBA.
Prevalence of comorbidities in individuals with neurodevelopmental disorders from the aggregated phenomics data of 51,227 pediatric individuals.
Nat Med. 2024 Jul;30(7):1994-2003.

Bosch E, Popp B, Güse E, Skinner C, **van der Sluijs PJ,** Maystadt I, Pinto AM, Renieri A, Bruno LP, Granata S, Marcelis C, Baysal Ö, Hartwich D, Holthöfer L, Isidor B, Cogne B, Wieczorek D, Capra V, Scala M, De Marco P, Ognibene M, Jamra RA, Platzer K, Carter LB, Kuismin O, van Haeringen A, Maroofian R, Valenzuela I, Cuscó I, Martinez-Agosto JA, Rabani AM, Mefford HC, Pereira EM, Close C, Anyane-Yeboa K, Wagner M, Hannibal MC, Zacher P, Thiffault I, Beunders G, Umair M, Bhola PT, McGinnis E, Millichap J, van de Kamp JM, Prijoles EJ, Dobson A, Shillington A, Graham BH, Garcia EJ, Galindo MK, Ropers FG, Nibbeling EAR, Hubbard G, Karimov C, Goj G, Bend R, Rath J, Morrow MM, Millan F, Salpietro V, Torella A, Nigro V, Kurki M, Stevenson RE, Santen GWE, Zweier M, Campeau PM, Severino M, Reis A, Accogli A, Vasileiou G.
Elucidating the clinical and molecular spectrum of SMARCC2-associated NDD in a cohort of 65 affected individuals.
Genet Med. 2023 Nov;25(11):100950.

Dingemans AJM, Hinne M, Truijen KMG, Goltstein L, van Reeuwijk J, de Leeuw N, Schuurs-Hoeijmakers J, Pfundt R, Diets IJ, den Hoed J, de Boer E, Coenen-van der Spek J, Jansen S, van Bon BW, Jonis N, Ockeloen CW, Vulto-van Silfhout AT, Kleefstra T, Koolen DA, Campeau PM, Palmer EE, Van Esch H, Lyon GJ, Alkuraya FS, Rauch A, Marom R, Baralle D, **van der Sluijs PJ**, Santen GWE, Kooy RF, van Gerven MAJ, Vissers LELM, de Vries BBA.
PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework.
Nat Genet. 2023 Sep;55(9):1598-1607.

Valencia AM, Sankar A, **van der Sluijs PJ**, Satterstrom FK, Fu J, Talkowski ME, Vergano SAS, Santen GWE, Kadoch C.
Landscape of mSWI/SNF chromatin remodeling complex perturbations in neuro-developmental disorders.
Nat Genet. 2023 Aug;55(8):1400-1412.

van der Sluijs PJ, Joosten M, Alby C, Attié-Bitach T, Gilmore K, Dubourg C, Fradin M, Wang T, Kurtz-Nelson EC, Ahlers KP, Arts P, Barnett CP, Ashfaq M, Baban A, van den Born M, Borrie S, Busa T, Byrne A, Carriero M, Cesario C, Chong K, Cueto-González AM, Dempsey JC, Diderich KEM, Doherty D, Farholt S, Gerkes EH, Gorokhova S, Govaerts LCP, Gregersen PA, Hickey SE, Lefebvre M, Mari F, Martinovic J, Northrup H, O'Leary M, Parbhoo K, Patrier S, Popp B, Santos-Simarro F, Stoltenburg C, Thauvin-Robinet C, Thompson E, Vulto-van Silfhout AT, Zahir FR, Scott HS, Earl RK, Eichler EE, Vora NL, Wilnai Y, Giordano JL, Wapner RJ, Rosenfeld JA, Haak MC, Santen GWE.
Discovering a new part of the phenotypic spectrum of Coffin-Siris syndrome in a fetal cohort.
Genet Med. 2023 Feb;25(2):100004.

van der Sluijs PJ, Vergano SA, Roeder ER, Jongmans MCJ, Santen GWE.
Recommending revised hepatoblastoma surveillance in children with a pathogenic ARID1A variant. Reply to "Cancer in ARID1A-Coffin-Siris syndrome: Review and report of a child with hepatoblastoma" by Cárcamo et al. 2022.
Eur J Med Genet. 2023 Feb;66(2):104694.

Levy MA, Relator R, McConkey H, Pranckeviciene E, Kerkhof J, Barat-Houari M, Bargiacchi S, Biamino E, Palomares Bralo M, Cappuccio G, Ciolfi A, Clarke A, DuPont BR, Elting MW, Faivre L, Fee T, Ferilli M, Fletcher RS, Cherick F, Foroutan A, Friez MJ, Gervasini C, Haghshenas S, Hilton BA, Jenkins Z, Kaur S, Lewis S, Louie RJ, Maitz S, Milani D, Morgan AT, Oegema R, Østergaard E, Pallares NR, Piccione M, Plomp AS, Poulton C, Reilly J, Rius R, Robertson S, Rooney K, Rousseau J, Santen GWE, Santos-Simarro F, Schijns J, Squeo GM, John MS, Thauvin-Robinet C, Traficante G, **van der Sluijs PJ**, Vergano SA, Vos N, Walden KK, Azmanov D, Balci TB, Banka S, Gecz J, Henneman P, Lee JA, Mannens MMAM, Roscioli T, Siu V, Amor DJ, Baynam G, Bend EG, Boycott K, Brunetti-Pierri N, Campeau PM, Campion D, Christodoulou J, Dyment D, Esber N, Fahrner JA, Fleming MD, Genevieve D, Heron D, Husson T, Kernohan KD, McNeill A, Menke LA, Merla G, Prontera P, Rockman-Greenberg C, Schwartz C, Skinner SA, Stevenson RE, Vincent M, Vitobello A, Tartaglia M, Alders M, Tedder ML, Sadikovic B.
Functional correlation of genome-wide DNA methylation profiles in genetic neuro-developmental disorders.
Hum Mutat. 2022 Nov;43(11):1609-1628.

Levy MA, McConkey H, Kerkhof J, Barat-Houari M, Bargiacchi S, Biamino E, Bralo MP, Cappuccio G, Ciolfi A, Clarke A, DuPont BR, Elting MW, Faivre L, Fee T, Fletcher RS, Cherik F, Foroutan A, Friez MJ, Gervasini C, Haghshenas S, Hilton BA, Jenkins Z, Kaur S, Lewis S, Louie RJ, Maitz S, Milani D, Morgan AT, Oegema R, Østergaard E, Pallares NR, Piccione M, Pizzi S, Plomp AS, Poulton C, Reilly J, Relator R, Rius R, Robertson S, Rooney K, Rousseau J, Santen GWE, Santos-Simarro F, Schijns J, Squeo GM, St John M, Thauvin-Robinet C, Traficante G, **van der Sluijs PJ**, Vergano SA, Vos N, Walden KK, Azmanov D, Balci T, Banka S, Gecz J, Henneman P, Lee JA, Mannens MMAM, Roscioli T, Siu V, Amor DJ, Baynam G, Bend EG, Boycott K, Brunetti-Pierri N, Campeau PM, Christodoulou J, Dyment D, Esber N, Fahrner JA, Fleming MD, Genevieve D, Kerrnohan KD, McNeill A, Menke LA, Merla G, Prontera P, Rockman-Greenberg C, Schwartz C, Skinner SA, Stevenson RE, Vitobello A, Tartaglia M, Alders M, Tedder ML, Sadikovic B.
Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders.
HGG Adv. 2021 Dec 3;3(1):100075.

van der Sluijs PJ, Alders M, Dingemans AJM, Parbhoo K, van Bon BW, Dempsey JC, Doherty D, den Dunnen JT, Gerkes EH, Millller IM, Moortgat S, Regier DS, Ruivenkamp CAL, Schmalz B, Smol T, Stuurman KE, Vincent-Delorme C, de Vries BBA, Sadikovic B, Hickey SE, Rosenfeld JA, Maystadt I, Santen GWE.
A Case Series of Familial ARID1B Variants Illustrating Variable Expression and Suggestions to Update the ACMG Criteria.
Genes (Basel). 2021 Aug 20;12(8):1275.

van der Sluijs PJ, Sleeboom C, Peetsold MG.
Een neonaat met weefsel uit de navel.
Nederlands Tijdschrift voor Geneeskunde. 2020;164:D4769

van der Sluijs PJ, Santen GWE.
Letter regarding the article: "Striking phenotypic overlap between Nicolaides-Baraitser and Coffin-Siris syndromes in monozygotic twins with ARID1B intragenic deletion".
Eur J Med Genet. 2020 Apr;63(4):103813.

Vergano SA, **van der Sluijs PJ**, Santen GWE.
ARID1B-Related Disorder.
In: Adam MP, Ardinger HH, Pagon RA, et al., eds. GeneReviews®. Seattle (WA) 1993.

van der Sluijs PJ, Ruivenkamp CAL, Santen GWE.
Response to Gorokhova et al.
Genet Med. 2019 Sep;21(9):2159-2164.

van der Sluijs PJ, Jansen S, Vergano SA, Adachi-Fukuda M, Alanay Y, AlKindy A, Baban A, Bayat A, Beck-Wödl S, Berry K, Bijlsma EK, Bok LA, Brouwer AFJ, van der Burgt I, Campeau PM, Canham N, Chrzanowska K, Chu YWY, Chung BHY, Dahan K, De Rademaeker M, Destree A, Dudding-Byth T, Earl R, Elcioglu N, Elias ER, Fagerberg C, Gardham A, Gener B, Gerkes EH, Grasshoff U, van Haeringen A, Heitink KR, Herkert JC, den Hollander NS, Horn D, Hunt D, Kant SG, Kato M, Kayserili H, Kersseboom R, Kilic E, Krajewska-Walasek M, Lammers K, Laulund LW, Lederer D, Lees M, López-González V, Maas S, Mancini GMS, Marcelis C, Martinez F, Maystadt I, McGuire M, McKee S, Mehta S, Metcalfe K, Milunsky J, Mizuno S, Moeschler JB, Netzer C, Ockeloen CW, Oehl-Jaschkowitz B, Okamoto N, Olminkhof SNM, Orellana C, Pasquier L, Pottinger C, Riehmer V, Robertson SP, Roifman M, Rooryck C, Ropers FG, Rosello M, Ruivenkamp CAL, Sagiroglu MS, Sallevelt SCEH, Sanchis Calvo A, Simsek-Kiper PO, Soares G, Solaeche L, Sonmez FM, Splitt M, Steenbeek D, Stegmann APA, Stumpel CTRM, Tanabe S, Uctepe E, Utine GE, Veenstra-Knol HE, Venkateswaran S, Vilain C, Vincent-Delorme C, Vulto-van Silfhout AT, Wheeler P, Wilson GN, Wilson LC, Wollnik B, Kosho T, Wieczorek D, Eichler EE, Pfundt R, de Vries BBA, Clayton-Smith J, Santen GWE.
The ARID1B spectrum in 143 patients: From nonsyndromic intellectual disability to Coffin-Siris syndrome.
Genet Med. 2019 Jun;21(6):1295-1307.

van der Sluijs PJ, Aten E, Barge-Schaapveld DQCM, Bijlsma EK, Bökenkamp-Gramann R, Donker Kaat L, van Doorn R, van de Putte DF, van Haeringen A, Ten Harkel ADJ, Hilhorst-Hofstee Y, Hoffer MJV, den Hollander NS, van Ierland Y, Koopmans M, Kriek M, Moghadasi S, Nibbeling EAR, Peeters-Scholte CMPCD, Potjer TP, van Rij M, Ruivenkamp CAL, Rutten JW, Steggerda SJ, Suerink M, Tan RNGB, van der Tuin K, Visser R, van der Werf-'t Lam AS, Williams M, Witlox R, Santen GWE.
Putting genome-wide sequencing in neonates into perspective.
Genet Med. 2019 May;21(5):1074-1082.

CURRICULUM VITAE

Eline van der Sluijs was born on September 10, 1995, in Gorinchem, the Netherlands. In 2013 she completed her pre-university education (*atheneum*) at the Gomarus Scholengemeenschap in Gorinchem and began her medical studies at the Leiden University Medical Center (LUMC) that same year. During her bachelor's program, she joined the *ARID1B* patient database project under the supervision of Dr. Gijs Santen.

After graduating with honors in 2016, she pursued a one-year Master's degree in Vitality and Ageing at Leiden University. As part of this program she completed a scientific internship at the Department of Public Health and Primary Care at the LUMC. In 2017 she continued with her Master's in Medicine. She began this program with a scientific internship at the Department of Clinical Genetics at the LUMC, resulting in two publications and co-authorship of the *ARID1B* GeneReview. These contributions formed the basis for securing an MD-PhD grant in 2019.

In September 2019 she paused her medical studies to begin her PhD research. From September 2020 to June 2021 she interrupted her PhD to complete her medical degree. In September 2023 she started working as a resident in Clinical Genetics at the LUMC.

DANKWOORD

Aan het einde van dit proefschrift wil ik graag iedereen bedanken die heeft bijgedragen. Naast de patiënten en hun ouders wil ik enkele mensen in het bijzonder danken.

Allereerst Gijs, zonder jou was dit proefschrift er niet geweest. Dank je wel voor je begeleiding en de mogelijkheid om dit onderzoek te doen. Jouw kennis, ideeën en kritische blik waren fundamenteel voor de totstandkoming van dit proefschrift. Lucia, thank you for your guidance and patience with me as a clinician without any prior lab work experience. Christi, hoewel je inhoudelijk niet direct betrokken was bij dit onderzoek, weerhield dat je er niet van om betrokken mee te leven en de voortgang nauwgezet te volgen. Bedankt voor je betrokkenheid en het meedenken als promotor.

Kelly, jij hebt mij wegwijs gemaakt in de wereld van kweken, DNA-isolatie, PCR, western blots en kleuringen. Dank je wel voor je geduld en begeleiding. I would also like to thank the other members of the DevEpi group and the Human Genetics department who were always willing to help me with my questions, especially Maja, Cor, Serkan, and JP.

Natuurlijk wil ik ook mijn collega's uit de kantoortuin bedanken voor de gezelligheid en steun, waaronder Anne-Sophie, Diantha, Maayke, Minne, Noah en Remco. Jullie gezelligheid, discussies en congresbezoeken hebben mijn tijd als PhD-student compleet gemaakt.

Ook wil ik de masterstudenten bedanken die ik met plezier heb begeleid: Ileana, Lisa, Marjolein, Annelinde, Nan en Eva. Dank je wel voor jullie enthousiasme en inzet.

Er hebben veel mensen bijgedragen aan het onderzoek in dit proefschrift. Daarom wil ik iedereen bedanken: mijn coauteurs, collega's van de (CSS-)poli, het lab en het secretariaat. Dear Kamlesh, thank you for thinking along with me while wrapping up the last part of this thesis. Lex en Mariëlle, speciaal dank aan jullie voor jullie snelle reacties op mijn mails, de analyses en de fijne samenwerking.

Lieve Resi, in dit dankwoord verdien jij zeker een bijzondere plaats. Ik had anderhalf jaar geleden niet kunnen bedenken dat jouw betrokkenheid bij dit proefschrift ertoe zou leiden dat het vol zou komen te staan met jouw mooie creaties. Telkens als ik met een idee kwam, maakte jij er een kunstwerk van, mooier dan ik me had kunnen voorstellen.

Lieve paranimfen, Jona en Jan, wat fijn dat jullie vandaag aan mijn zijde staan, dank jullie wel!

Tot slot wil ik iedereen om mij heen bedanken: mijn lieve (schoon)familie en vrienden, hartelijk dank voor jullie betrokkenheid, interesse, steun en de gezellige momenten samen. Lieve papa en mama, dank voor jullie liefde en steun, ook al was het soms misschien niet helemaal duidelijk wat ik precies wilde doen en waarom. Zonder jullie had ik dit niet kunnen bereiken.

Lieve Benjamin, dank je wel dat je er bent, voor je geduld als ik weer eens over mijn onderzoek aan het praten was, en je steun in de afgelopen jaren. Ik kijk uit naar onze toekomst samen!