



**Gdański Uniwersytet Medyczny
Wydział Lekarski**

Natalia Krawczyńska

Analiza genetycznego mozaicyzmu w zespolu Cornellii de Lange

Rozprawa na stopień doktora nauk medycznych

Promotor: dr hab. n. med. Bartosz Wasąg

Pracę wykonano w Katedrze i Zakładzie Biologii i Genetyki Medycznej
Gdańskiego Uniwersytetu Medycznego

Gdańsk 2019

Praca w całości została sfinansowana
z grantu nr UMO-2014/15/B/NZ5/03503
ze środków Narodowego Centrum Nauki.

SPIS TREŚCI

1. WYKAZ STOSOWANYCH SKRÓTÓW	4
2. WYKAZ PRAC WCHODZĄCYCH W SKŁAD ROZPRAWY	6
3. STRESZCZENIE.....	7
3.1 Wprowadzenie	7
3.2 Cele pracy	10
3.3 Metody	11
3.4 Omówienie publikacji wchodzących w skład rozprawy.....	12
3.5 Podsumowanie	14
4. SUMMARY	15
4.1 Introduction.....	15
4.2 The aims of doctoral dissertations	17
4.3 Methods	18
4.4 The overview of publications, which are part of doctoral dissertation.....	19
4.5 Conclusions.....	21
5. WYKAZ STOSOWANEGO PIŚMIENNICTWA	22
6. PUBLIKACJE WCHODZĄCE W SKŁAD ROZPRAWY DOKTORSKIEJ	24

1. WYKAZ STOSOWANYCH SKRÓTÓW

ATRX	gen kodujący czynnik remodelujący chromatynę ATRX (<i>ATRX chromatin remodeler</i>)
AURKB	gen kodujący kinazę B (<i>aurora kinase B</i>)
CDC45	gen kodujący białko sororinę (<i>cell division cycle associated 5</i>)
CdLS	zespół Cornelia de Lange (<i>Cornelia de Lange Syndrome</i>)
cDNA	komplementarny DNA (<i>complementary DNA</i>)
DDX11	gen kodujący helikazę DNA (<i>DEAD/H-box helicase 11</i>)
DNA	kwasy deoksyrybonukleinowe (<i>deoxyribonucleic acid</i>)
ESCO1	gen kodujący N-acetylotransferazę 1 (<i>establishment of sister chromatid cohesion N-acetyltransferase 1</i>)
ESCO2	gen kodujący N-acetylotransferazę 2 (<i>establishment of sister chromatid cohesion N-acetyltransferase 2</i>)
ESPL1	gen kodujący białko separtyne (<i>extra spindle pole bodies like 1, separate</i>)
HDAC8	gen kodujący deacetylazę histonową 8 (<i>histone deacetylase 8</i>)
HEAT	domena delanginy zbudowana z wielokrotnych powtórzeń aminokwasów (<i>Huntington, Elongation factor 3, the PR65/A subunit of PP2A and the lipid kinase Tor</i>)
KMT2A	gen kodujący metylotransferazę lizynową 2A (<i>lysine methyltransferase 2A</i>)
LOD	logarytm dziesiętny stosunku wiarygodności (<i>logarithm of odds</i>)
MAU2	gen kodujący białko MAU2 (<i>MAU2 sister chromatid cohesion factor</i>)
MLPA	zależna od ligacji łańcuchowa reakcja polimerazy (<i>Multiplex Ligation-dependent Probe Amplification</i>)
mRNA	matrycowy RNA (<i>messenger RNA</i>)
NGS	sekwencjonowanie nowej generacji (<i>next generation sequencing</i>)
NIPBL	gen kodujący delanginę (<i>NIPBL cohesin loading factor</i>)
p	krótkie ramie chromosomu (<i>z fr. petit – mały</i>)
PDS5A	gen kodujący białko PDS5A (<i>PDS5 cohesin associated factor A</i>)
PDS5B	gen kodujący białko PDS5B (<i>PDS5 cohesin associated factor B</i>)
PLK1	gen kodujący kinazę PLK1 (<i>polo like kinase 1</i>)
PTTG1	gen kodujący sekuryne (<i>PTTG1 regulator of sister chromatid separation, securit</i>)

<i>RAD21</i>	gen kodujący białko z rodziny kleizyn (<i>RAD21 cohesin complex component</i>)
<i>RECQL4</i>	gen kodujący helikazę RecQ (<i>RecQ like helicase 4</i>)
RNA	kwas rybonukleinowy (<i>ribonucleic acid</i>)
RT-PCR	reakcja łańcuchowa polimerazy z odwrotną transkryptazą (<i>reverse transcriptase PCR</i>)
<i>SGOL1</i>	gen kodujący białko SGOL1 (<i>shugoshin 1</i>)
<i>SMC1A</i>	gen kodujący białko utrzymujące strukturę chromosomów 1A (<i>structural maintenance of chromosomes 1A</i>)
<i>SMC3</i>	gen kodujący białko utrzymujące structure chromosomów 3 (<i>structural maintenance of chromosomes 3</i>)
<i>STAG1</i>	gen kodujący stromalinę 1 (<i>stromal antigen 1</i>)
<i>STAG2</i>	gen kodujący stromalinę 2 (<i>stromal antigen 2</i>)
<i>TAF6</i>	gen kodujący białko podjednostki TFIID współczynnika inicjacji transkrypcji 6 (<i>TATA-box binding protein associated factor 6</i>)
<i>WAPAL</i>	gen kodujący białko WAPL (<i>WAPL cohesin release factor</i>)

2. WYKAZ PRAC WCHODZĄCYCH W SKŁAD ROZPRAWY

1. **Natalia Krawczyńska, Alina Kuźniacka, Jolanta Wierzba, Ilaria Parenti, Frank J Kaiser, Bartosz Wasąg: Mosaic intronic NIPBL variant in a family with Cornelia de Lange Syndrom.** Front Genet. 2018; 9:1-5

Czasopismo umieszczone na Liście Filadelfijskiej, wskaźnik Impact Factor: **4,151**

2. **Natalia Krawczynska, Jolanta Wierzba, Jacek Jasiiecki, Bartosz Wasag: Molecular characterization of two novel intronic variants of NIPBL gene detected in unrelated Cornelia de Lange Syndrome patients.** BMC Med. Genet. 2019; 20:1-6

Czasopismo umieszczone na Liście Filadelfijskiej, wskaźnik Impact Factor: **1,913**

3. **Natalia Krawczynska, Jolanta Wierzba, Bartosz Wasag: Genetic mosaicism in a group of patients with Cornelia de Lange Syndrome.** Front Pediatr. 2019; 7:1-7

Czasopismo umieszczone na Liście Filadelfijskiej, wskaźnik Impact Factor: **2,335**

Łączny IF prezentowanych prac: 8,399

3. STRESZCZENIE

3.1 Wprowadzenie

Zespół Cornelia de Lange (*Cornelia de Lange Syndrome*, CdLS; OMIM #122470, 300590, 300882, 610759, 614701) został po raz pierwszy opisany przez niemieckiego lekarza Brachmana w 1916 roku [1]. Nazwa zespołu pochodzi od imienia i nazwiska duńskiej lekarki Cornelia de Lange, która w 1933 roku jako pierwsza opisała dwie niespokrewnione ze sobą dziewczynki z tym zespołem [2].

CdLS jest wielosystemowym zaburzeniem rozwojowym o heterogennym fenotypie i genotypie. Wśród pacjentów z CdLS, pomimo bardzo zróżnicowanego fenotypu, można zaobserwować specyficzne cechy odpowiadające zespołowi. Charakterystyczna dysmorfia twarzy jest jednym z najbardziej rozpoznawalnych objawów zespołu. W dysmorficznej twarzy istotnymi cechami są: długie i gęste brwi (często połączone po środku - synophrys), krótki nos, długa rynienka podnosowa, cienka górna warga, mała broda oraz nisko osadzone uszy [3,4]. Innymi charakterystycznymi wykładnikami CdLS są hirsutyzm, zaburzenia pokarmowe, opóźnienie wzrostu oraz deformacje kończyn górnych, począwszy od łagodnych cech, takich jak skrócenie paliczków, aż do ciężkich malformacji przedramienia. Niepełnosprawność intelektualna występuje z różnym nasileniem. Dodatkowo, wielu pacjentów wykazuje także cechy autystyczne oraz inne zaburzenia behawioralne, w tym tendencje do autoagresji.

Wykonana w 12 rodzinach obciążonych CdLS analiza sprzężeń ujawniła cztery regiony chromosomowe mogące mieć potencjalne znaczenie w rozwoju zespołu [5]. Analiza sprzężeń regionu zlokalizowanego na chromosomie 5p13 z CdLS osiągnęła najwyższą wartość według skali oceniającej logarytm dziesiętny stosunku wiarygodności (*logarithm of odds*, LOD). Powyższa obserwacja wraz z wcześniej opisanymi rearanżacjami (translokacją t(5;13)(p13.1;q12.1) i delecją del5p14.2-p13.1) pozwoliły zidentyfikować region krytyczny o wielkości 1,1Mpz znajdujący się na krótkich ramionach (p) chromosomu 5 [6], a w konsekwencji umożliwiły identyfikację zmiany w genie *NIPBL* (OMIM*608667). Wkrótce po tym odkryciu - u pacjentów z CdLS wykryto także warianty patogenne w genach *SMC1A* (OMIM*300040) [7], *SMC3* (OMIM*606062) [8], *RAD21* (OMIM*606462) [9], oraz *HDAC8* (OMIM*300269)

[10]. Produkty pięciu wyżej wymienionych genów są niezbędne do prawidłowej budowy i funkcjonowania kompleksu kohezyny.

Kompleks kohezyny pełni niezwykle istotne funkcje w wielu procesach wewnątrzkomórkowych, w tym kontrolę rozchodzenia się chromatyd siostrzanych w trakcie podziałów komórkowych, regulację ekspresji genów oraz naprawę pęknięć nici DNA. W związku z powyższym choroby związane z zaburzeniami kompleksu kohezyny są klasyfikowane zarówno do grupy kohezynopatii jak również transkryptomopatii [11]. Ponadto dane literaturowe wskazują na udział szeregu innych genów, oprócz wcześniej wymienionych, w występowaniu fenotypu ze spektrum CdLS. W większości są to geny, których produkty biorą udział w budowie i funkcjonowaniu kompleksu kohezyny, w regulacji ekspresji i naprawie DNA. Ostatnie badania wskazują na udział *KMT2A* w patogenezę tego zespołu [11,12].

Na początku XXI wieku rozpoznanie CdLS opierało się na subiektywnej ocenie objawów klinicznych. Obecnie jednym z podstawowych kryteriów rozpoznania CdLS jest wykrycie wariantu patogennego w jednym z pięciu genów związanych z zespołem. Niemniej jednak u około 30% pacjentów z klinicznym rozpoznaniem CdLS wynik badania genetycznego jest negatywny, co potwierdza wcześniejszą obserwację dotyczącą heterogenności genetycznej tego zespołu. W związku z tym, coraz częściej w diagnostyce CdLS wykorzystywane są zaawansowane metody biologii molekularnej. Postawą w tym zakresie jest wykorzystanie sekwencjonowania nowej generacji (*next generation sequencing*, NGS) z zastosowaniem paneli genowych bądź sekwencjonowania całego genomu. Ponadto, dotychczasowe obserwacje naukowców wskazują na duże znaczenie analizy procesu składania mRNA w diagnostyce CdLS. W publikacjach wchodzących w skład niniejszej rozprawy doktorskiej wykazano, że empiryczne określenie roli poszczególnych wariantów genetycznych w składaniu mRNA nie zawsze jest tożsame z teoretycznym rozważaniem w oparciu o badania *in silico* [13,14]. W związku z powyższym analiza cDNA powinna stanowić istotny element badań diagnostycznych wariantów intronowych u pacjentów z CdLS.

W diagnostyce CdLS duże znaczenie ma również analiza mozaicyzmu komórkowego, będącego konsekwencją negatywnej selekcji komórek w stosunku do zmiany patogennej. Skutkiem tego jest brak możliwości wykrycia zmiany w DNA izolowanym z limfocytów krwi obwodowej co wiąże się z koniecznością pobrania alternatywnej tkanki do badań. W takiej sytuacji zazwyczaj materiałem badanym jest wymaz z

wewnętrznej części policzka pacjenta. Dotychczas odsetek mozaicyzmu, tylko w genie *NIPBL*, wyniósł 23% (11/44) [15], a należy zauważyć iż zjawisko to może w różnym stopniu dotyczyć także innych genów związanych z patogenezą tego zespołu [3,16,17].

W pracach prezentowanych w niniejszej rozprawie przedstawiono wyniki badań, które obrazują mozaicyzm genetyczny u pacjentów z CdLS [13,14,18].

3.2 Cele pracy

1. Określenie częstości i rodzaju germinalnych wariantów genetycznych występujących w genach kodujących układ kohezyny u pacjentów z klinicznie potwierdzonym CdLS.
2. Określenie częstości i rodzaju mozaikowych wariantów genetycznych badanych genów.

3.3 Metody

Zadania badawcze zrealizowano poprzez analizę:

1. sekwencji kodującej genów związanych z kompleksem kohezyny: *NIPBL*, *SMC1A*, *SMC3*, *RAD21*, *HDAC8*, *STAG1*, *SGOL1*, *PDS5A*, *PTTG1*, *TAF6*, *ESCO2*, *WAPAL*, *CDCA5*, *KMT2A*, *DDX11*, *ESPL1*, *PDS5B*, *PLK1*, *AURKB*, *ESCO1*, *MAU2*, *ATR*, *STAG2* oraz *RECQL4* z wykorzystaniem techniki NGS oraz sekwencjonowania metodą Sanger,
2. genu *NIPBL* z wykorzystaniem techniki MLPA (*multiplex ligation-dependent probe amplification*) w celu identyfikacji dużych rearanżacji tego genu,
3. RNA wybranych pacjentów w celu określenia potencjalnego wpływu wariantów genetycznych na proces składania mRNA.

3.4 Omówienie publikacji wchodzących w skład rozprawy

W publikacji pt: „*Mosaic intronic NIPBL variant in a family with Cornelia de Lange Syndrome*” (Front Genet., 2018) przedstawiono wyniki NGS przeprowadzone u syna z klasycznym fenotypem CdLS oraz matki z łagodną postacią tego zespołu. Upřednio wykonana analiza molekularna u tych osób z wykorzystaniem techniki sekwencjonowania metodą Sangerą pozwoliła na identyfikację heterozygotycznego wariantu patogennego w genie *NIPBL* (c.869-2A>G) [19]. Analiza z użyciem NGS wykazała obecność tego wariantu w DNA wyizolowanym z krwi syna i matki na poziomie odpowiednio 46% (522/1132) i 23% (251/1115). Uzyskane wyniki pozwoliły na stwierdzenie, że u matki występuje wariant mozaikowy. Analiza cDNA potwierdziła, że zidentyfikowany wariant patogenny w genie *NIPBL* zaburza proces składania mRNA - p.(Gly290_Lys498del). Prawdopodobnym mechanizmem powstania wariantu mozaikowego wykrytego u matki jest spontaniczna mutacja, która miała miejsce po zapłodnieniu a przed gastrulacją zygoty. Przedstawione w omawianej publikacji wyniki dowodzą, że zastosowanie technik o wysokiej czułości pozwala na identyfikację wariantów mozaikowych. Wykorzystanie tych technik pozwoli na bardziej precyzyjną diagnostykę molekularną pacjentów z CdLS oraz szczegółową korelację uzyskanych wyników z danymi klinicznymi.

Wyniki analizy dwóch zmian intronowych przeprowadzonych u dwojga niespokrewnionych pacjentów z klasycznym CdLS zostały przedstawione w pracy pt. „*Molecular characterization of two novel intronic variants of NIPBL gene detected in unrelated Cornelia de Lange Syndrome patients.*” (BMC Med. Genet., 2019). Pierwszy etap analizy materiału genetycznego pacjentów wykonano z wykorzystaniem techniki NGS. Dalsze analizy przeprowadzono u pacjentów, u których zidentyfikowano dwa warianty intronowe w genie *NIPBL*: c.6954+3A>G oraz c.5862+1delG. W tym celu zamplifikowane fragmenty cDNA wklonowano do wektora plazmidowego i zsekwencjonowano. Analiza uzyskanych elektroforogramów wykazała, że zmiana c.6954+3A>G powoduje delecję na poziomie RNA (r.6764_6954del), co prowadzi do powstania zmodyfikowanego produktu białkowego - p.(Ser2255LeyfsTer20). W przypadku drugiego pacjenta zaobserwowano, że zmiana c.5862+1delG jest przyczyną insercji na poziomie RNA (r.5862_5863ins5862+2_5862+41) czego efektem jest wprowadzenie kodonu STOP w pozycji 1955 - p.(Leu1955Ter). W obydwu

przypadkach skrócenie sekwencji aminokwasów doprowadziło do utraty wysoce konserwowanych powtórzeń HEAT (*Huntington, Elongation factor 3, the PR65/A subunit of PP2A and the lipid kinase Tor*) w produkcie białkowym. Przedstawione w pracy wyniki wskazują, że każda zmiana intronowa u pacjenta z CdLS powinna być analizowana na poziomie RNA co pozwoli na prawidłową klasyfikację zidentyfikowanych wariantów i usprawni analizę fenotyp-genotyp u pacjentów z CdLS.

W pracy „*Genetic mosaicism in a group of patients with Cornelia de Lange Syndrome.*” (Front Pediatr., 2019) opisano wyniki badań molekularnych wykonanych w grupie 69 pacjentów z CdLS. Materiał badany stanowił DNA wyizolowany zarówno z komórek jądrzastych krwi obwodowej (68 pacjentów), jak i pochodzący z komórek nabłonkowych uzyskanych drogą wymazu z wewnętrznej części policzka (13 pacjentów). Badania molekularne przeprowadzono z wykorzystaniem panelu NimbleGen SeqCap EZ Hyper Cap (Roch Diagnostics) oraz urządzenia MiSeq (Illumina). Analizę bioinformatyczną przeprowadzono za pomocą oprogramowań IGV (Broad Institute), SeqNext (JSI Medical Systems) oraz Alamut (Interactive Biosoftware). Następnie wykorzystując sekwencjonowanie metodą Sangera potwierdzano obecność wariantów patogennych i prawdopodobnie patogennych w wybranych genach. W przypadku wariantów intronowych, potencjalny wpływ na alternatywne składanie mRNA analizowano za pomocą reakcji RT-PCR oraz sekwencjonowania metodą Sangera. Ponadto, w celu wykrycia dużych rearanzacji w genie *NIPBL* u pacjentów przeprowadzono analizę MLPA z zastosowaniem komercyjnie dostępnego zestawu (MRC Holland). Germinalne warianty patogene stwierdzono u 18 pacjentów – 14 wariantów w genie *NIPBL* oraz dwa w genach *SMC1A* oraz *HDAC8*. Mozaikowe warianty patogene wykryto u czterech pacjentów – trzy warianty w genie *NIPBL* oraz jeden w genie *KMT2A*. Uzyskane wyniki dowodzą, iż zastosowanie technik o wysokiej czułości pozwala na wykrycie wariantów patogennych u większej liczby pacjentów z CdLS. Wykorzystanie DNA pochodzącego z różnych listków zarodkowych pozwala na identyfikację moziacyzmu u części pacjentów z CdLS. Natomiast analiza na poziomie RNA pozwala na prawidłową klasyfikację wariantów intronowych.

3.5 Podsumowanie

Heterogenność genetyczna i fenotypowa CdLS stanowi wyzwanie dla prawidłowej diagnostyki pacjentów z tym zespołem. Patogenne warianty zarówno germinalne jak i mozaikowe mogą występować w genach związanych nie tylko z kompleksem kohezyny.

- Wykorzystanie wysokoprzepustowych i czułych metod molekularnych takich jak NGS pozwala na identyfikację podłoża molekularnego u większej liczby pacjentów z CdLS.
- Wykorzystanie do badań molekularnych DNA wyizolowanego z różnych listków zarodkowych pozwala na identyfikację mozaikowych wariantów genetycznych co pozwala na potwierdzenie klinicznego rozpoznania zespołu CdLS.
- Dodatkowa analiza na poziomie RNA pozwala na określenie patogenności i prawidłową klasyfikację wykrytych wariantów intronowych.

4. SUMMARY

4.1 Introduction

Cornelia de Lange Syndrome (CdLS; OMIM # 122470, 300590, 300882, 610759, 614701) was first described by the German physician Brachman in 1916 [1]. Nevertheless, the name of the syndrome comes from the name of the Danish doctor Cornelia de Lange, who described two unrelated girls with this syndrome in 1933 [2].

CdLS is a multisystemic developmental disorder with a heterogeneous phenotype and genotype. Despite very distinct phenotype among CdLS patients specific features corresponding to the syndrome are observed. The characteristic facial dysmorphism is one of the most recognizable symptoms of the CdLS. In the dysmorphic face, the important features are long and thick eyebrows (often connected in the middle - synophrys), a short nose, a long philtrum, a thin upper lip, a small beard and low set ears [3,4]. Other characteristic CdLS hallmarks are hirsutism, digestive disorders, growth retardation, and upper limb deformities, ranging from benign features such as shortened phalanges to severe forearm malformations. Intellectual disability occurs with varying intensity. In addition, many patients also show autistic symptoms and other behavioral disorders, including self-injury.

The linkage analysis performed in 12 families with CdLS revealed four chromosomal regions that could be of potential importance in the development of CdLS [5]. The linkage analysis of the region located on the 5p13 chromosome with CdLS reached the highest value in the range of logarithm of odds (LOD). The above observation together with the previously described rearrangements (translocation t(5;13)(p13.1;q12.1) and deletion del5p14.2-p13.1) identified the critical region of 1.1 Mbp on the short arms (p) of the chromosome 5 [6]. Based on the previously mentioned data, future studies could have become the basis for identifying variants in *NIPBL* (OMIM*608667). Shortly afterwards *NIPBL* variants discovery - pathogenic variants in *SMC1A* (OMIM*300040) [7], *SMC3* (OMIM*606062) [8], *RAD21* (OMIM*606462) [9], and *HDAC8* (OMIM*300269) [10] were detected in patients with CdLS. The products of all five genes are involved in the construction and functioning of the cohesin complex.

The cohesin complex plays an important role in many intracellular processes. It is involved in controlling the separation of sister chromatids during cell division,

as well as in the regulation of gene expression and the repair of double-strand breaks of DNA. Regarding these phenomena, diseases associated with disturbances of the cohesin complex are included in the group of cohesinopathies and also classified as transcriptomopathies [11]. Literature data indicate the participation of other genes to those previously mentioned, in the CdLS occurrence. Most of them are involved in the proper functioning of the cohesin complex and in the regulation of gene expression and DNA repair processes. One of the latest papers, report the involvement of the *KMT2A* in the pathogenesis of the CdLS [11, 12].

At the beginning of the twenty-first century, the diagnosis of CdLS was based on a subjective assessment of clinical symptoms. Currently, one of the basic criteria for diagnosing CdLS is the detection of a pathogenic variant in one of the five genes associated with the syndrome. However, in about 30% of patients with a clinically diagnosed syndrome the result of genetic test is negative, which confirms the genetic heterogeneity of CdLS. Therefore, more advanced methods of molecular biology are used in CdLS diagnosis. The approach in this area is the use of next-generation sequencing (NGS) using gene panels or whole genome sequencing. Previous scientific observations confirmed that the analysis of the impact of intronic variants on mRNA splicing could have a pivotal role in CdLS diagnostics in the future. The publications included in this doctoral dissertation show that the empirical determination of the role of particular genetic variants in the mRNA splicing is not always identical to the theoretical consideration based on *in silico* research [13, 14]. Regarding the aforementioned reports, cDNA analysis should become an important element of diagnostic tests of intronic variants in CdLS patients.

In the CdLS diagnostics, it is extremely important to analyze cellular mosaicism, which is a consequence of the negative selection of cells with the pathogenic variant. This results in the inability to detect variants in DNA isolated from peripheral blood lymphocytes, which involves the need to take an alternative tissue for testing. In this situation, the tested material usually is a buccal swab. Until now, the percentage of mosaicism only in the *NIPBL* in patients with CdLS was estimated at 23% (11/44) [15]. Should be noted that the phenomenon of mosaicism is not only related to the occurrence of variants in the *NIPBL*, but also to the other four genes related to CdLS [3,16,17].

The papers presented in this dissertation present the results of studies that illustrate genetic mosaicism in patients with CdLS [13,14,18].

4.2 The aims of doctoral dissertations

1. To determine the frequency and type of germinal genetic variants in genes encoding the cohesin complex in a group of patients with clinically confirmed CdLS.
2. To determine the frequency and type of mosaic genetic variants in analyzed genes.

4.3 Methods

The tasks were accomplished by analysis of:

1. coding sequences of cohesion complex genes: *NIPBL*, *SMC1A*, *SMC3*, *RAD21*, *HDAC8*, *STAG1*, *SGOL1*, *PDS5A*, *PTTG1*, *TAF6*, *ESCO2*, *WAPAL*, *CDCA5*, *KMT2A*, *DDX11*, *ESPL1*, *PDS5B*, *PLK1*, *AURKB*, *ESCO1*, *MAU2*, *ATRX*, *STAG2* and *RECQL4* using NGS and bidirectional Sanger sequencing,
2. large rearrangements within the coding sequence of *NIPBL* using MLPA (*multiplex ligation-dependent probe amplification*) technique,
3. RNA isolated from selected patients in order to determine the potential impact of genetic variants on the mRNA splicing.

4.4 The overview of publications, which are part of doctoral dissertation

The publication entitled "*Mosaic intronic NIPBL variant in a family with Cornelia de Lange Syndrome*" (Front Genet., 2018) presents the results of NGS performed in the son with a classic phenotype and the mother with a mild form of CdLS. The previously completed molecular analysis of these patients using the Sanger sequencing technique allowed the identification of heterozygous pathogenic variant in *NIPBL* (c.869-2A>G) [19]. Analysis with the use of NGS showed the presence of this variant in DNA isolated from the blood of the son and the mother at the level of 46% (522/1132) and 23% (251/1115), respectively. The obtained results allowed to conclude that the mother has a mosaic variant. cDNA analysis confirmed that the identified pathogenic variant in the *NIPBL* interferes mRNA splicing – p.(Gly290_Lys498del). The probable mechanism of the mosaic variant detected in the mother is a spontaneous mutation that occurred after fertilization and before zygote gastrulation. The results presented in this publication suggest that the use of highly sensitive techniques allows the identification of mosaic variants. The use of these techniques will allow more precise molecular diagnostics of patients with CdLS and a detailed correlation of the obtained results with clinical data.

The results of the analysis of two unrelated patients with classic CdLS with two intronic variants are presented in the paper "*Molecular characterization of two novel intronic variants of NIPBL gene detected in unrelated Cornelia de Lange Syndrome patients.*" (BMC Med. Genet., 2019). The first stage of the analysis was performed using NGS. Further analyses were performed in two patients with intronic variants in *NIPBL*: c.6954+3A>G and c.5862+1delG. For this purpose, the amplified cDNA fragments were cloned into a plasmid vector and sequenced. Analysis of the obtained electropherograms showed that the variant c.6954+3A>G causes deletion at the RNA level (r.6764_6954del), which leads to the creation of a modified protein product - p.(Ser2255LeyfsTer20). In the second patient, the variant c.5862+1delG causes insertion at the RNA level (r.5862_5863ins5862+2_5862+41), which leads to the creation of STOP codon in 1955 position - p.(Leu1955Ter). In both cases, the shortening of the amino acid sequence led to the loss of highly conserved HEAT repeats (*Huntington, Elongation factor 3, the PR65 / A subunit of PP2A and the lipid kinase Tor*) in the protein. The results presented in the paper indicate that each intronic variant

in a patient with CdLS should be analyzed at the RNA level, which will allow the correct classification of identified variants and improve the phenotype-genotype analysis in patients.

The work entitled "*Genetic mosaicism in a group of patients with Cornelia de Lange Syndrome.*" (Front Pediatr., 2019) describes the results of molecular analyses performed in a group of 69 patients with CdLS. The tested material consisted of DNA isolated from both peripheral blood cells (68 patients) and epithelial cells obtained from buccal swabs (13 patients). Molecular studies were performed using the NimbleGen SeqCap EZ Hyper Cap (Roch Diagnostics) panel and the MiSeq (Illumina). The bioinformatic analysis was performed using IGV (Broad Institute), SeqNext (JSI Medical Systems) and Alamut (Interactive Biosoftware) software. Using Sanger sequencing, the presence of pathogenic and probably pathogenic variants was confirmed. In the case of intronic variants, the potential impact on alternative mRNA splicing was analyzed by RT-PCR followed by Sanger sequencing. In addition, to detect large rearrangements in *NIPBL*, MLPA analysis was performed using a commercially available kit (MRC Holland). Germline pathogenic variants were found in 18 patients - 14 variants in *NIPBL* and two in *SMC1A* and *HDAC8*. Mosaic pathogenic variants were detected in four patients - three variants in *NIPBL* and one in *KMT2A*. The obtained results prove that the use of highly sensitive techniques allows detecting of pathogenic variants in a larger number of patients with CdLS. The use of DNA derived from various germ layers allows the identification of the disease in part of patients with CdLS. Moreover, analysis at the RNA level allows for proper classification of intronic variants.

4.5 Conclusions

Genetic and phenotypic heterogeneity of CdLS is challenging for the correct diagnosis of patients with this syndrome. Pathogenic germinal and mosaic variants can be detected in genes associated not only with the cohesin complex.

- The use of high-throughput and sensitive molecular methods such as NGS allows identifying the molecular basis of the syndrome in a larger number of CdLS patients.
- The use of DNA isolated from different germ layers allows to identify mosaic genetic variants which confirm the clinical diagnosis of CdLS.
- Additional RNA analysis allows to determine the pathogenicity and correctly classify detected intronic variants.

5. WYKAZ STOSOWANEGO PIŚMIENICTWA

1. Brachmann W. Ein fall von symmetrischer monodactylie durch ulnadefekt, mit symmetrischer ughaut- bildung in den ellenbeugen, sowie anderen abnormalitaten (zwergha igkeit, halsrippen, behaarung). Jahrbuch Kinderheilkd 1916;84.
2. de Lange C. Sur un type nouveau de dégénération (typus Amstelodamensis) [On a new type of degen- eration (type Amsterdam)]. Arch Méd Enfants 1933;36.
3. Ansari M, Poke G, Ferry Q, Williamson K, Aldridge R, Meynert AM, et.al. Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. J Med Genet. 2014;51:659-683.
4. Mannini L, Cucco F, Quarantotti V, Krantz ID, Musio A. Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome. Hum Mutat. 2013;34(12):1589-96.
5. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nat Genet. 2004;36:631-635.
6. Hulinsky R, Byrne JLB, Lowichik A, Viskochil DH: Fetus With Interstitial del(5)(p13.1p14.2) Diagnosed Postnatally With Cornelia de Lange Syndrome Am J Hum Genet. 2005;173A:336-338.
7. Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. Nat Genet 2006;38:528–530.
8. Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, et al. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of Cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet. 2007;80:485–494.
9. Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, et al. RAD21 mutations cause a human cohesinopathy. Am J Hum Genet. 2012(2);90:1014–1027.
10. Kaiser FJ, Ansari M, Braunholz D, Concepción Gil-Rodríguez M, Decroos C, Wilde J, et al. Loss-of-function *HDAC8* mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance. *Human Molecular Genetics*. 2014;23(11):2888-2900.
11. Yuan B, Pehlivan D, Karaca E, Patel N, Charng WL, Gambin T, et al. Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. J clin invest. 2015;125(2):636-51.
12. Parenti I, Teresa-Rodrigo ME, Pozojevic J, Ruiz Gil S, Bader I, Braunholz D, et al. Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes. Hum Genet. 2017;136:307–320.

13. Krawczyńska N, Wierzba J, Jasiński J, Wasąg B: Molecular characterization of two novel intronic variants of NIPBL gene detected in unrelated Cornelia de Lange Syndrome patients. *BMC Med. Genet.* 2019; 20:1.
14. Krawczyńska N, Wierzba J, Wasąg B: Genetic mosaicism in a group of patients with Cornelia de Lange Syndrome. *Front Pediatr.* 2019; 7:1-7.
15. Castronovo P, Delahaye-Duriez A, Gervasini C, Azzollini J, Minier F, Russo S, et al. Somatic mosaicism in Cornelia de Lange syndrome: a further contributor to the wide clinical expressivity? *Clin Genet.* 2010;78:560–564.
16. Boyle MI, Jespersgaard C, Brøndum-Nielsen K, Bisgaard AM, Tümer Z. Cornelia de Lange syndrome. *Clin Genet.* 2015;88(1):1-12.
17. Parenti I, Gervasini C, Pozojevic J, Wendt KS, Watrin E, Azzollini J, et al. Expanding the clinical spectrum of the ‘HDAC8-phenotype’ – implications for molecular diagnostics, counseling and risk prediction. *Clin Genet.* 2016;89:564–573.
18. Krawczyńska N, Kuźniacka A, Wierzba J, Parenti I, Kaiser FJ, Wasąg B: Mosaic intronic *NIPBL* variant in a family with Cornelia de Lange Syndrom. *Front Genet.* 2018;9:255.
19. Teresa-Rodrigo ME, Eckhold J, Puisac B, Dalski A, Gil-Rodríguez MC, Braunholz D, et al. Functional characterization of NIPBL physiological splice variants and eight splicing mutations in patients with Cornelia de Lange syndrome. *Int J Mol Sci.* 2014;15:10350-64.

6. PUBLIKACJE WCHODZĄCE W SKŁAD ROZPRAWY DOKTORSKIEJ

1. Natalia Krawczyńska, Alina Kuźniacka, Jolanta Wierzba, Ilaria Parenti, Frank J Kaiser, Bartosz Wasąg: Mosaic intronic *NIPBL* variant in a family with Cornelia de Lange Syndrom. *Front Genet.* 2018; 9:1-5. doi: [10.3389/fgene.2018.00255](https://doi.org/10.3389/fgene.2018.00255)
2. Natalia Krawczynska, Jolanta Wierzba, Jacek Jasiński, Bartosz Wasąg: Molecular characterization of two novel intronic variants of NIPBL gene detected in unrelated Cornelia de Lange Syndrome patients. *BMC Med. Genet.* 2019; 20:1-6. doi: [10.1186/s12881-018-0738-y](https://doi.org/10.1186/s12881-018-0738-y)
3. Natalia Krawczynska, Jolanta Wierzba, Bartosz Wasąg: Genetic mosaicism in a group of patients with Cornelia de Lange Syndrome. *Front Pediatr.* 2019; 7:1-7. doi: [10.3389/fped.2019.00203](https://doi.org/10.3389/fped.2019.00203)