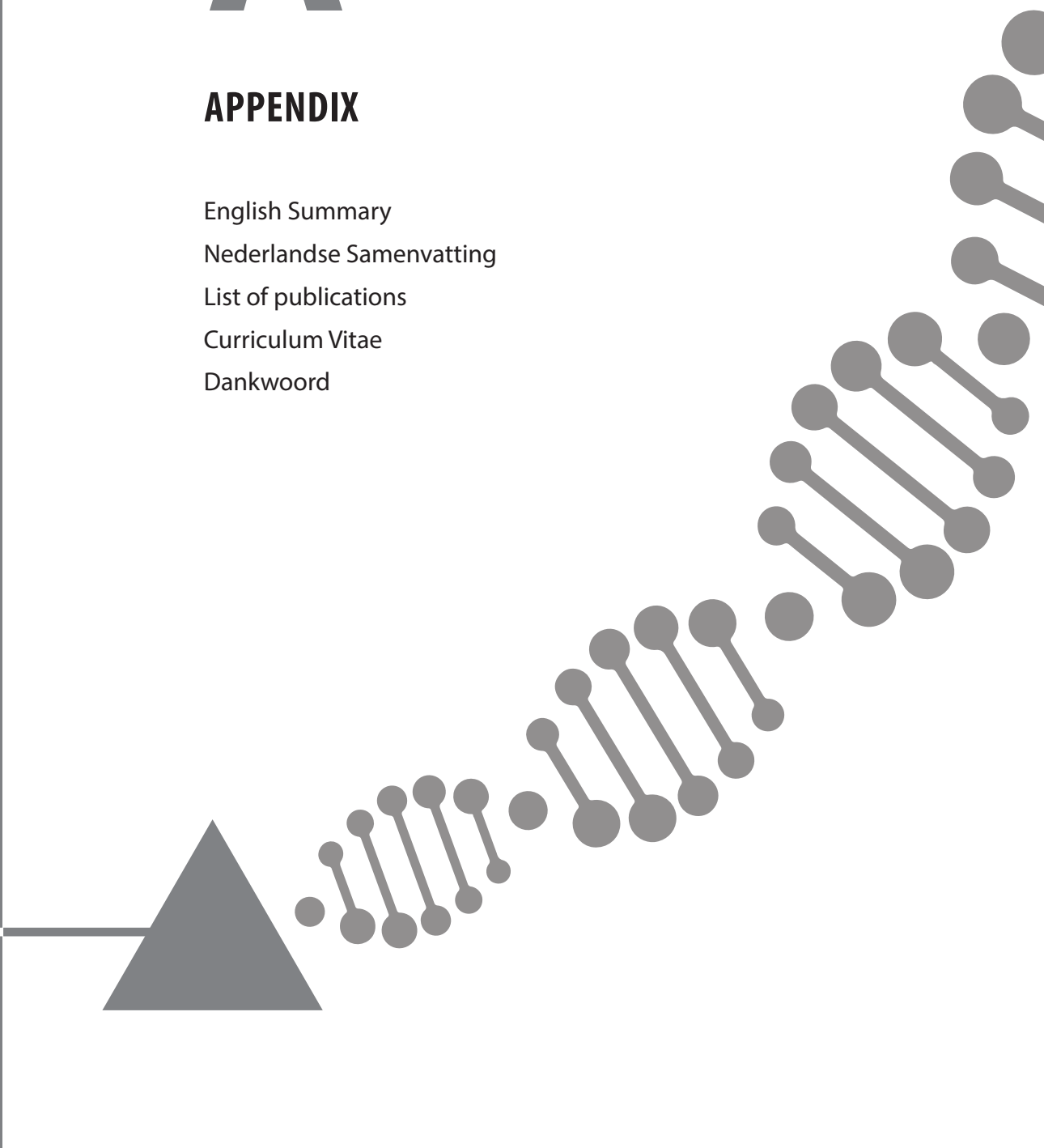
Expanding the mutation spectrum in FSHD and ICF syndrome

Boogaard, T.L. van den

A

APPENDIX



ENGLISH SUMMARY

Epigenetic regulation by chromatin modifiers needs to be tightly controlled to ensure correct patterns of gene expression. Misregulation of epigenetic processes can lead to disease, as is the case in FSHD and ICF syndrome. FSHD and ICF share a common hallmark of hypomethylation of repetitive DNA. FSHD is a muscular dystrophy mainly affecting the facial, shoulder girdle, and upper arm muscles. It is characterized by partial chromatin relaxation of the *DUX4* encoding D4Z4 macrosatellite repeat located on chromosome 4, and transcriptional derepression of the DUX4 transcription factor in skeletal muscle. D4Z4 chromatin relaxation is consistently marked by CpG hypomethylation of the D4Z4 repeat, including the *DUX4* promoter. In the most common form, FSHD1, this chromatin relaxation is the result of a repeat contraction of the D4Z4 repeat to 1-10 units (normal range 8-100 units). In the rare form of FSHD, FSHD2, D4Z4 chromatin relaxation occurs on all D4Z4 arrays and is most often caused by variants in the D4Z4 chromatin repressor SMCHD1.

ICF is an immunodeficiency syndrome mainly characterized by hypo- or agammaglobulinemia in the presence of B cells, developmental delay and facial anomalies. Compared to FSHD, ICF patients show more genome-wide hypomethylation at repetitive elements and this results in, amongst others, centromeric instability, which is the cytogenetic hallmark of ICF. Recessive variants in four genes have been identified to cause ICF syndrome: *DNMT3B* (ICF1), *ZBTB24* (ICF2), *CDCA7* (ICF3) and *HELLS* (ICF4). In this thesis the variant spectrum for chromatin modifiers involved in FSHD and ICF has been expanded.

In chapter 2 three FSHD2 families are described in which two potentially damaging variants in *SMCHD1* were identified in the proband. For each family we investigated whether these variants were located *in cis* or *in trans* as well as their contribution to D4Z4 hypomethylation. In the first family, a splice site variant and a nonsense variants *in cis* were identified in exon 21. The splice variant results in skipping of exon 21, which neutralizes the nonsense variant in exon 21. However, skipping of exon 21 is predicted to disrupt the open reading frame by a premature stop codon in exon 22. In the second family, the proband carries a missense variant in exon 28 and a splice site variant in exon 25. Although the missense variant in exon 28 was predicted to affect function, no D4Z4 hypomethylation was detected in the mother of the proband who also carries this variant. This suggests that only the splice site variant in exon 25, which seems to be in an *SMCHD1* mutation hotspot, affects SMCHD1 function and contributes to disease. This family also emphasizes the predictive value of D4Z4 methylation analysis in establishing the functional consequences of *SMCHD1* variants. In the third family, the proband carries two *SMCHD1* missense variants in exon 24 and exon 45 *in trans* and shows severe D4Z4 hypomethylation. D4Z4 hypomethylation was also detected in

the father and the mother of the proband, who carry the missense variant in exon 24 and exon 45, respectively. This revealed that both missense variants contribute to D4Z4 hypomethylation independently. The combination of these missense variants has an additive effect on D4Z4 hypomethylation, since this was more strongly reduced in the proband and his sister who carry both missense variants. Although the mother and the proband both carry a permissive D4Z4 allele of 27 repeat units, only the proband with two *SMCHD1* missense variants is affected. This suggests that these variants also have an additive effect on disease presentation. Interestingly, this family also shows that having two *SMCHD1* missense variants *in trans* is compatible with life in males and females.

In chapter 3 whole *SMCHD1* gene deletions were identified in two FSHD2 families, which were previously tested negative for exonic *SMCHD1* variants. Some neighbouring genes were also deleted in these families, however these genes have not been reported to be sensitive to hemizygosity. This study emphasizes that hemizygosity of *SMCHD1* can also cause FSHD2, and results in a haploinsufficiency mechanism as is also expected for the open reading frame disrupting variants identified in one third of the FSHD2 patients. *SMCHD1* hemizygosity is also present in most patients with 18p-deletion syndrome, however muscular dystrophy has not been reported in these patients. D4Z4 methylation analysis in 18p-deletion patients revealed a similar degree of hypomethylation in most of the 18p-deletion patients as is seen in FSHD2 patients. Based on the prevalence of permissive 4qA alleles with a D4Z4 repeat size in the FSHD2 range (11-16 units) in the control population, it is expected that approximately 1:8 18p-deletion patients with SMCHD1 hemizygosity might be at risk of developing FSHD.

In chapter 4 we described a (deep) intronic variant in two FSHD families. In the first family an intronic variant located 15 base pairs proximal to exon 14 was identified (c.1843-15A>G), which creates a splice acceptor site. The variant results in the inclusion of the distal 14 nucleotides of intron 13 into the transcript and this is predicted to disrupt the open reading frame by a premature stop codon in exon 14. The variant was identified in the proband and his sister, and segregates with D4Z4 hypomethylation. In the second family a deep intronic variant in *SMCHD1* was identified in intron 34 (c.4347-236A>G). This variant creates a splice acceptor site, while a cryptic splice donor site is predicted at position c.4347-183 of *SMCHD1*. Together this results in exonisation of 53 nucleotides from intron 34 into the transcript, which is predicted to disrupt the open reading frame by a premature stop codon in exon 35. In this family this deep intronic variant segregates with D4Z4 hypomethylation and modifies disease severity in family members with a permissive 7 units D4Z4 repeat array.

D4Z4 hypomethylation cannot always be explained by *SMCHD1* variants in all FSHD2 families. In chapter 5 we described the identification of heterozygous missense variants in *DNMT3B* in two of these families. These *DNMT3B* variants segregate with D4Z4

hypomethylation and increased penetrance of FSHD. We identified *DUX4* expression in MyoD transduced fibroblasts from an affected individual from one of these families with a permissive D4Z4 repeat array of 13 units, but not in an unaffected individual from the other family with a permissive D4Z4 repeat array of 44 units. This suggests that heterozygous *DNMT3B* variants can only derepress *DUX4* in combination with smaller D4Z4 repeat arrays.

Since homozygous or compound heterozygous *DNMT3B* variants cause autosomal recessive ICF1 syndrome, we made a clinical and epigenetic comparison between some ICF1 patients and FSHD2 families with a *DNMT3B* variant. D4Z4 hypomethylation is shared between FSHD2 patients and ICF1 patients, and is also found in some parents of ICF1 patients who carry a heterozygous *DNMT3B* variant, like our FSHD2 patients. In addition, DUX4 expression was detected in myoblasts from one ICF1 patient, as well as in MyoD transduced fibroblasts from two other ICF1 patients. Some epigenetic characteristics of ICF were also detected in one *DNMT3B* variant carrier from these FSHD2 families. Although some epigenetic characteristics are shared, muscular dystrophy has not been reported in ICF1 patients or their parents and no immunodeficiency was detected in FSHD2 families, in concordance with parents of ICF1 patients. We propose that the effect of *DNMT3B* variants on *DUX4* expression and disease presentation, like for *SMCHD1*, depends on several aspects associated with the FSHD1 locus including D4Z4 repeat array size, and the presence of a polymorphic *DUX4* polyadenylation signal. This study suggests that multiple factors are involved in the epigenetic state at D4Z4 and the regulation of *DUX4* in muscle cells.

Finally, in chapter 6 we expand the spectrum of ICF1 and ICF2 variants. In seven ICF1 patients from four different families we identified in total six missense variants in the catalytic domain of DNMT3B. These variants were not previously described in ICF1 patients. In addition, we describe five ICF2 cases from five different families. Four of these ICF2 patients carry homozygous nonsense variants in *ZBTB24*, including one which was not previously identified in ICF2 patients. The fifth ICF2 patient carries a homozygous deletion on chromosome 6 which includes *ZBTB24*. This observation shows that complete absence of ZBTB24 protein is compatible with life and supports the hypothesis that most ICF2 patients suffer from a functional loss of ZBTB24. With these 12 ICF patients included, a total of 77 genetically confirmed patients has been reported. 56% of the ICF patients carry *DNMT3B* variants, 31% carry *ZBTB24* variants, 7% carry *CDCA7* variants and 7% carry *HELLS* variants. Only for ICF2, a gender bias was detected, with 79% male patients. This suggests that *ZBTB24* variants might be more deleterious for females.