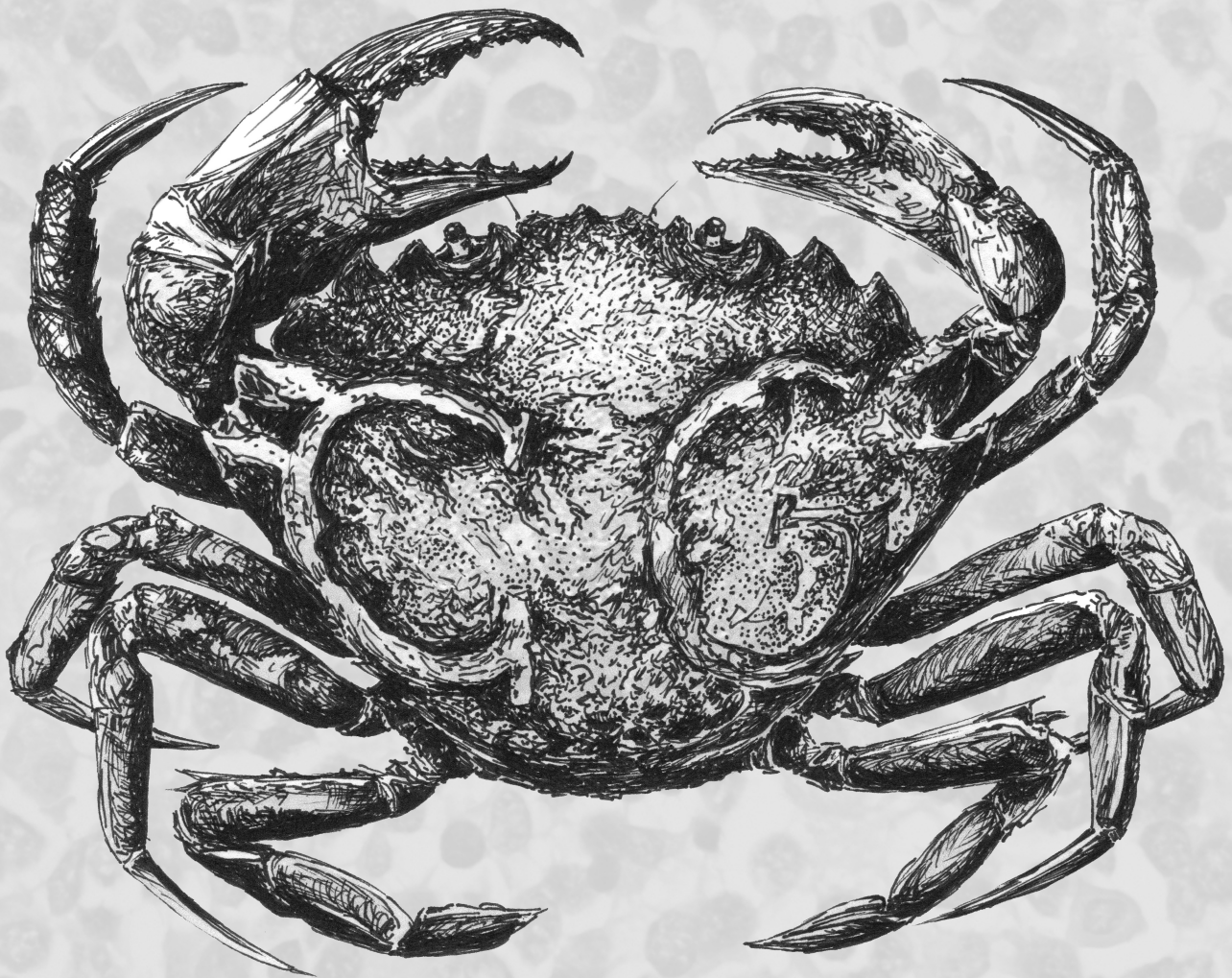


Exploring the neuroblastoma DNA methylome: from biology to biomarker





FACULTY OF MEDICINE
AND HEALTH SCIENCES

Ghent University, Faculty of Medicine and Health Sciences

EXPLORING THE NEUROBLASTOMA DNA METHYLOME: FROM BIOLOGY TO BIOMARKER

thesis submitted to fulfill the requirements for the degree of Doctor of Health Sciences
by Anneleen Decock, 2016

promotor

prof. dr. ir. Jo Vandesompele

co-promoters

prof. dr. Frank Speleman

dr. ir. Maté Ongenaert

Center for Medical Genetics
Cancer Research Institute Ghent
Ghent University Hospital, Medical Research Building 1
De Pintelaan 185, 9000 Ghent, Belgium
+32-9-3321951
a.decock@ugent.be

thesis submitted to fulfill the requirements for the degree of Doctor of Health Sciences

promotor

prof. dr. ir. Jo Vandesompele

Department of Pediatrics and Medical Genetics, Ghent University, Ghent, Belgium

co-promotors

prof. dr. Frank Speleman

Department of Pediatrics and Medical Genetics, Ghent University, Ghent, Belgium

dr. ir. Maté Ongenaert

Galapagos NV, Mechelen, Belgium

members of the examination committee

prof. dr. ir. Tim De Meyer

Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Ghent, Belgium

prof. dr. ir. Katleen De Preter

Department of Pediatrics and Medical Genetics, Ghent University, Ghent, Belgium

dr. Morgan Thénoz

Department of Pediatrics and Medical Genetics, Ghent University, Ghent, Belgium

prof. dr. Manon van Engeland

Cancer Pathology Laboratory, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands

prof. dr. Jolanda van Hengel

Department of Basic Medical Sciences, Ghent University, Ghent, Belgium

prof. dr. Max van Noesel

Princess Máxima Center for Pediatric Oncology, Utrecht University Medical Center, Utrecht, The Netherlands

The author and the promotor give the authorisation to consult and to copy parts of this work for personal use only. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using results from this thesis.

The research described in this thesis was conducted at the Center for Medical Genetics (Ghent University Hospital, Ghent, Belgium) and funded by grants from the Emmanuel Van der Schueren Foundation (scientific partner of the Flemish League Against Cancer (VLK)), the Ghent University Special Research Fund (GOA 01G01910), the Fournier-Majoie Foundation (FFM) and the Belgian National Lottery. Anneleen Decock was supported by a PhD fellowship of the Research Foundation Flanders (FWO) and an Emmanuel Van der Schueren research grant (VLK).

CONTENTS

ABBREVIATIONS	VII
FIGURES	XIII
BOXES	XV
TABLES	XVII
SUMMARY	- 1 -
SAMENVATTING	- 3 -
1 INTRODUCTION	- 5 -
1.1 DNA methylation: inherited information beyond the DNA sequence	- 7 -
1.1.1 Principles and functions of DNA methylation	- 8 -
1.1.1.1 DNA methylation and demethylation	- 8 -
1.1.1.2 DNA methylation regulates gene expression	- 9 -
1.1.2 DNA methylation in normal cellular processes	- 11 -
1.1.2.1 Subtelomeric and pericentromeric repeat silencing	- 11 -
1.1.2.2 Transposon silencing	- 11 -
1.1.2.3 X chromosome inactivation	- 11 -
1.1.2.4 Genomic imprinting	- 12 -
1.1.3 Cancer: when DNA methylation goes awry	- 12 -
1.1.3.1 Hypomethylation	- 12 -
1.1.3.2 Hypermethylation	- 13 -
1.1.3.3 Genetic lesions of the DNA methylation machinery	- 14 -
1.1.4 A plethora of technologies to illuminate DNA methylomes	- 14 -
1.1.4.1 Methylation-specific PCR	- 18 -
1.1.4.2 Methylation microarrays	- 19 -
1.1.4.3 Methyl-CpG-binding domain precipitation	- 20 -
1.1.4.4 Reduced representation and whole-genome bisulfite sequencing	- 20 -
1.1.5 DNA methylation as clinically useful biomarker	- 22 -
1.1.6 DNA methylation as therapeutic target	- 24 -
1.1.6.1 Nucleoside analogues	- 25 -
1.1.6.2 Small molecule inhibitors	- 26 -
1.1.6.3 Natural compounds	- 27 -
1.2 The quest for prognostic neuroblastoma markers	- 28 -
1.2.1 Widely used prognostic markers	- 30 -
1.2.1.1 Age of the patient at diagnosis	- 30 -
1.2.1.2 Tumor histology and differentiation grade	- 32 -
1.2.1.3 Tumor stage	- 32 -
1.2.1.4 <i>MYCN</i> amplification	- 37 -
1.2.1.5 Chromosome aberrations	- 38 -
1.2.1.5.1 DNA hyperdiploidy	- 38 -
1.2.1.5.2 Numerical and segmental chromosome aberrations	- 39 -
1.2.1.5.2.1 Chromosome 1p	- 39 -
1.2.1.5.2.2 Chromosome 11q	- 39 -
1.2.1.5.2.3 Chromosome 17q	- 40 -
1.2.2 mRNA and miRNA-based prognostic markers	- 40 -
1.2.3 DNA methylation-based prognostic markers	- 40 -
Review paper: neuroblastoma epigenetics: from candidate gene approaches to genome-wide screenings	- 43 -
1.2.4 Established prognostic markers are combined into a pretreatment risk classification system	- 57 -

2	RESEARCH OUTLINE	- 61 -
3	RESULTS	- 65 -
	Paper 1: Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers	- 69 -
	Paper 2: Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma	- 95 -
	Paper 3: DNA methylation profiling of primary neuroblastoma tumors using methyl-CpG-binding domain sequencing	- 117 -
	Paper 4: Stage 4S neuroblastoma tumors show a characteristic DNA methylation portrait	- 133 -
4	DISCUSSION AND FUTURE PERSPECTIVES	- 153 -
4.1	Defeating pitfalls and limitations in biomarker research: the successful establishment of a prognostic DNA methylation signature for neuroblastoma	- 155 -
4.2	Finding prognostic biomarkers for high-risk neuroblastoma: a tough nut to crack	- 158 -
4.3	Sharing is caring	- 158 -
4.4	The stage 4S DNA methylome sheds light on mechanisms contributing to spontaneous regression	- 159 -
4.5	Latent mysteries in neuroblastoma epigenetics may decipher the clinical enigma	- 160 -
4.6	Conclusions	- 161 -
	REFERENCES	- 163 -
	CURRICULUM VITAE	- 175 -

ABBREVIATIONS

A	adenine
ACC	accuracy
aCGH	array comparative genomic hybridization
AIMS	amplification of intermethylated sites
AJCC	American Joint Committee on Cancer
ALT	alternative lengthening of telomeres
ASO	antisense oligonucleotide
BAC	balanced accuracy
BIG N2N	Bioinformatics Institute Ghent - From Nucleotides to Networks
BMP	bone morphogenetic protein
BOF	Ghent University Special Research Fund
bp	base pair
BSPP	bisulfite sequencing padlock probes
BS seq	bisulfite sequencing
BT	bisulfite-treated
C	cytosine
CARD	caspase recruitment domain
CCLG	Children's Cancer and Leukaemia Group
CCSG	Children's Cancer Study Group
CEEHRC	Canadian Epigenetics, Environment and Health Research Consortium
CGH	comparative genomic hybridization
CHARM	comprehensive high-throughput arrays for relative methylation
ChIP	chromatin immunoprecipitation
CI	confidence interval
CIMP	CpG island methylator phenotype
COBRA	combined bisulfite restriction analysis
COG	Children's Oncology Group
Cq	quantification cycle
CREST	Core Research for Evolutional Science and Technology
CRIG	Cancer Research Institute Ghent
CT	computed tomography
DAC	5-aza-2'-deoxycytidine
DcR	decoy receptor
DD	death domain
ddPCR	digital droplet polymerase chain reaction
DED	death effector domain
DISC	death-inducing signaling complex
DM	double minute
DMH	differential methylation hybridization
DMR	differentially methylated region
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
DOD	deceased

abbreviations

DR	death receptor
EFS	event-free survival
EGCG	(–)-Epigallocatechin-3-gallate
ENCODE	Encyclopedia of DNA Elements
ERDF	European Regional Development Fund
FDA	Food and Drug Administration
FDR	false discovery rate
FFM	Fournier-Majoie Foundation
FIS	Spanish Health Research Fund
FISH	fluorescence <i>in situ</i> hybridization
FN	false negative
FNR	false negative rate
FP	false positive
FPR	false positive rate
FWO	Research Foundation Flanders
G	guanine
GEO	Gene Expression Omnibus
GN	ganglioneuroma
GNB	ganglioneuroblastoma
GPOH	German Pediatric Oncology and Hematology Group
GSEA	gene set enrichment analysis
GST	glutathione-S-transferase
HAT	histone acetyl transferase
HDAC	histone deacetylase
HDM	histone demethylase
HELP	<i>HpaII</i> tiny fragment enrichment by ligation-mediated PCR
HMT	histone methyltransferase
HM450 array	Illumina Infinium HumanMethylation450 BeadChip Kit
HR	high-risk
HSR	homogeneously staining region
HT-TREBS	high-throughput targeted repeat element bisulfite sequencing
HVA	homovanillic acid
ICR	imprinting control region
IDRF	image-defined risk factor
I-FISH	interphase fluorescence <i>in situ</i> hybridization
IGV	Integrative Genomics Viewer
IHEC	International Human Epigenome Consortium
iINRGdb	interactive International Neuroblastoma Risk Group database
INPC	International Neuroblastoma Pathology Committee
INSS	International Neuroblastoma Staging System
INRG	International Neuroblastoma Risk Group
INRGSS	International Neuroblastoma Risk Group Staging System
ISCIII	Institute of Health Carlos III
IWT	Institute for the Promotion of Innovation by Science and Technology in Flanders
JANB	Japanese Advanced Neuroblastoma Study Group

abbreviations

JINCS	Japanese Infantile Neuroblastoma Co-operative Study Group
KMT	lysine methyltransferase
LC480	Roche LightCycler 480
LD-BSP	limiting dilution bisulfite pyrosequencing
LDH	lactate dehydrogenase
LHC-BS	liquid hybridization capture-based bisulfite sequencing
LINES	Low and Intermediate Neuroblastoma European Study
LOH	loss of heterozygosity
LR	low-risk
MAB seq	methylation-assisted bisulfite sequencing
MBD	methyl-CpG-binding domain
MB-PCR	methyl-binding polymerase chain reaction
MCAM	methyated CpG island amplification microarray
MCA-RDA	methyated CpG island amplification with representational difference analysis
MeDIP	methyated DNA immunoprecipitation
MethylC seq	cytosine methylome sequencing
Methyl-MAPS	methylation mapping analysis by paired-end sequencing
MiGS	methyl-CpG-binding domain-isolated genome sequencing
MIRA	methyated CpG island recovery assay
miRNA	microRNA
MKI	mitosis-karyorrhexis index
MLPA	multiplex ligation-dependent probe amplification
MMASS	microarray-based methylation assessment of single samples
MRE seq	methylation-sensitive restriction enzyme sequencing
MRI	magnetic resonance imaging
mRNA	messenger RNA
MS-AP-PCR	methylation-sensitive arbitrarily primed polymerase chain reaction
MSBE	methylation-specific single base extension
MSCC	methyl-sensitive cut counting
MS-DGGE	methylation-specific denaturing gradient gel electrophoresis
MSDK	methylation-specific digital karyotyping
MS-HRM	methylation-sensitive high-resolution melting
MS-MCA	methylation-specific melting curve analysis
MS-MLPA	methylation-specific multiplex ligation-dependent probe amplification
MSNP	methylation-sensitive single nucleotide polymorphism chip analysis
MSO	methylation-specific oligonucleotide
MSP	methylation-specific polymerase chain reaction
MSRF	methylation-sensitive restriction fingerprinting
Ms-SNuPE	methylation-sensitive single nucleotide primer extension
MS-SSCA	methylation-sensitive single strand conformation analysis
MYCNO	<i>MYCN</i> non-amplified
MYCN1	<i>MYCN</i> amplified
NB	neuroblastoma
NCA	numerical chromosomal aberration
NDR	nucleosome-depleted region

abbreviations

NES	normalized enrichment score
NGS	next-generation sequencing
NIH	National Institutes of Health
NPV	negative predictive value
NSE	neuron-specific enolase
NTC	no template control
OR	odds ratio
OS	overall survival
PC	principal component
PCR	polymerase chain reaction
PE	paired-end
PGC	primordial germ cells
PNET	primitive neuroectodermal tumor
PNS	peripheral nervous system
POG	Pediatric Oncology Group
PPV	positive predictive value
PRMT	protein arginine methyltransferase
PTP	permeability transition pore
RLGS	restriction landmark genomic scanning
RNA	ribonucleic acid
ROC	receiver operating characteristic
RPKM	reads per kilobase CpG island per million
RRBS	reduced representation bisulfite sequencing
RRMAB seq	reduced-representation methylation-assisted bisulfite sequencing
RTICC	Thematic Network of Cooperative Research on Cancer
RTOP	Recherche Translationnelle en Oncologie Pédiatrique
SAM	S-adenosyl-L-methionine
SCA	segmental chromosomal aberration
scBS seq	single cell bisulfite sequencing
scWGBS	single cell whole-genome bisulfite sequencing
scRRBS	single cell reduced representation bisulfite sequencing
SEER	Surveillance, Epidemiology, and End Results
SIOPEN	International Society of Pediatric Oncology Europe Neuroblastoma Group
SJCRH	St Jude Children's Research Hospital
SRO	smallest region of overlap
ST	subtelomere
SURV	surviving
Sz	size
T	thymine
TAB-array	TET-assisted bisulfite conversion with array analysis
TDG	thymine-DNA glycosylase
TF	transcription factor
Tm	melting temperature
TN	true negative
TNR	true negative rate (specificity)

abbreviations

TNFR	tumor necrosis factor receptor
TNM	tumor-node-metastases
TP	true positive
TPR	true positive rate (sensitivity)
TSA	trichostatin-A
TSS	transcription start site
UICC	International Union against Cancer
VLK	Flemish League Against Cancer
VMA	vanillylmandelic acid
WGBS	whole-genome bisulfite sequencing
XIC	X-inactivation center
μ WGBS	whole-genome bisulfite sequencing in very small cell populations
¹²³ I-mIBG	iodine-123 meta-iodobenzylguanidine
^{99m} Tc-MDP	technetium-99m-methylene diphosphonate

FIGURES

Figure 1. Various epigenetic mechanisms contribute to modulation of the chromatin structure.	- 7 -
Figure 2. DNA methylation levels change dynamically during human development.	- 8 -
Figure 3. Enzymatic DNA demethylation occurs via multiple intermediates.	- 10 -
Figure 4. DNA methylation alterations contribute to carcinogenesis via several mechanisms.	- 13 -
Figure 5. Bisulfite converts cytosine to uracil.	- 18 -
Figure 6. Methylation-specific PCR (MSP) allows site-specific DNA methylation analysis.	- 19 -
Figure 7. Methyl-CpG-binding domain (MBD) sequencing allows genome-wide DNA methylation analysis.	- 21 -
Figure 8. DNA methylation biomarkers are used in many clinical applications.	- 24 -
Figure 9. The neural crest is a transient population of cells during embryonal development.	- 29 -
Figure 10. Kaplan-Meier curves are used to analyze survival data.	- 31 -
Figure 11. Neuroblastoma tumors are categorized into undifferentiated, poorly differentiated and differentiating subtypes.	- 34 -
Figure 12. Fluorescence in situ hybridization of neuroblastoma cells depicts <i>MYCN</i> amplification as double minutes or intrachromosomal homogeneously staining regions.	- 38 -
Figure 13. Using cutoffs on the event-free survival data in the survival tree regression analysis of the International Neuroblastoma Risk Group (INRG) cohort, four pretreatment risk group categories were determined: very low-risk, low-risk, intermediate-risk and high-risk.	- 59 -
Figure 14. The research project resulted in five publications.	- 156 -

BOXES

- | | | |
|---------|--|--------|
| Box 1. | Several large-scale epigenomics projects have been initiated. | - 10 - |
| Box 2. | Differential methylation analyses based on MBD sequencing data make use of several bioinformatics tools. | - 21 - |
| Box 3. | Several statistical metrics can be used to evaluate biomarker performance. | - 23 - |
| Box 4. | Several classes of histone modulators have been developed and clinically implemented. | - 25 - |
| Box 5. | Survival probabilities can be estimated using the Kaplan-Meier method. | - 31 - |
| Box 6. | The presence or absence of Homer Wright rosettes is used to determine the neuroblastoma differentiation grade. | - 33 - |
| Box 7. | The International Neuroblastoma Risk Group Staging System (INRGSS) stratifies neuroblastoma tumors into several tumor stages, based on image-defined risk factors. | - 35 - |
| Box 8. | The International Neuroblastoma Risk Group Staging System (INRGSS) makes use of a list of image-defined risk factors (IDRFs) to determine the neuroblastoma tumor stage. | - 36 - |
| Box 9. | Radiological imaging is used to evaluate the presence or absence of image-defined risk factors. | - 37 - |
| Box 10. | Several technologies can be used to assess DNA gains and losses. | - 38 - |

TABLES

Table 1.	Multiple DNA methylation detection methodologies, making use of restriction enzymes, affinity enrichment and/or bisulfite conversion, have been developed.	- 15 -
Table 2.	Biomarkers can be evaluated by calculating several statistical metrics.	- 23 -
Table 3.	Different populations of neural crest cells arising along the neural axis give rise to distinct cell types.	- 30 -
Table 4.	The International Neuroblastoma Pathology Classification (Shimada System) stratifies neuroblastomas into favorable and unfavorable tumors, using the age of the patient and tumor differentiation grade.	- 33 -
Table 5.	The International Neuroblastoma Staging System (INSS) stratifies neuroblastoma tumors into several tumor stages, based on the degree of surgical tumor resection and lymph node involvement.	- 35 -
Table 6.	The International Neuroblastoma Risk Group Staging System (INRGSS) stratifies neuroblastoma tumors into several tumor stages, based on image-defined risk factors.	- 35 -
Table 7.	Using the International Neuroblastoma Risk Group (INRG) consensus pretreatment classification, neuroblastoma patients are stratified into risk groups.	- 58 -

SUMMARY

Neuroblastoma (NB), a childhood tumor arising from immature sympathetic nervous system cells, is a heterogeneous disease with prognosis ranging from excellent long-term survival to high-risk with fatal outcome. In order to determine the most appropriate treatment modality, patients are stratified into risk groups at the time of diagnosis, based on combinations of clinical and biological parameters, namely age of the patient, tumor stage, histology, grade of differentiation, *MYCN* oncogene amplification, chromosome 11q aberration and DNA ploidy. However, use of this risk classification system has shown that accurate assessment of NB prognosis remains difficult and that additional prognostic markers are warranted. Therefore, we aimed to identify prognostic tumor DNA methylation biomarkers for NB.

To find new biomarkers, we profiled the primary tumor DNA methylome using methyl-CpG-binding domain (MBD) sequencing, i.e. massively parallel sequencing of methylation-enriched DNA fractions, captured using the high affinity of MBD to bind methylated cytosines. As proof of principle, we applied this technology to 8 NB cell lines, and in combination with mRNA expression studies, this led to a first selection of 43 candidate biomarkers. Next, methylation-specific PCR (MSP) assays were designed, to allow candidate-specific methylation analysis in a primary tumor cohort of 89 samples. As such, we identified new prognostic DNA methylation biomarkers, and delineated the technological aspects and data analysis pipeline to set up a more extended biomarker study. In this follow-up study, the DNA methylome of 102 primary tumors, selected for risk classification and survival, was characterized by MBD sequencing. Differential methylation analyses between the prognostic patient groups put forward 78 top-ranking biomarker candidates, which were subsequently tested on two independent cohorts of 132 and 177 samples, adopting the high-throughput MSP pipeline of our pilot study. Multiple individual MSP assays were prognostically validated and through the implementation of a newly developed statistical framework, a robust 58-marker methylation signature predicting overall and event-free survival was established. This study represents the largest DNA methylation (biomarker) study in NB so far.

The MBD sequencing data were shared with the research community through the format of a data descriptor. As such, these data are fully available to others, ensuring its reusability for other research purposes. To illustrate how these data can be applied to gain new insights into the NB pathology, we characterized the DNA methylome of stage 4S NB, a special type of NB found in infants with widespread metastases at diagnosis that paradoxically is associated with an excellent outcome due to its remarkable capacity to undergo spontaneous regression. More specifically, we compared promoter methylation levels between stage 4S, stage 1/2 (localized disease with favorable prognosis) and stage 4 (metastatic disease with dismal prognosis) tumors, and showed that specific chromosomal locations are enriched in stage 4S differentially methylated promoters and that specific subtelomeric promoters are hypermethylated in stage 4S. Furthermore, genes involved in important oncogenic pathways, in neural crest development and differentiation, and in epigenetic processes are differentially methylated and expressed in stage 4S.

In conclusion, by exploring the DNA methylome of NB, we have not only demonstrated that DNA methylation patterns are intimately related to NB biology, but also found additional clinically relevant prognostic biomarkers.

SAMENVATTING

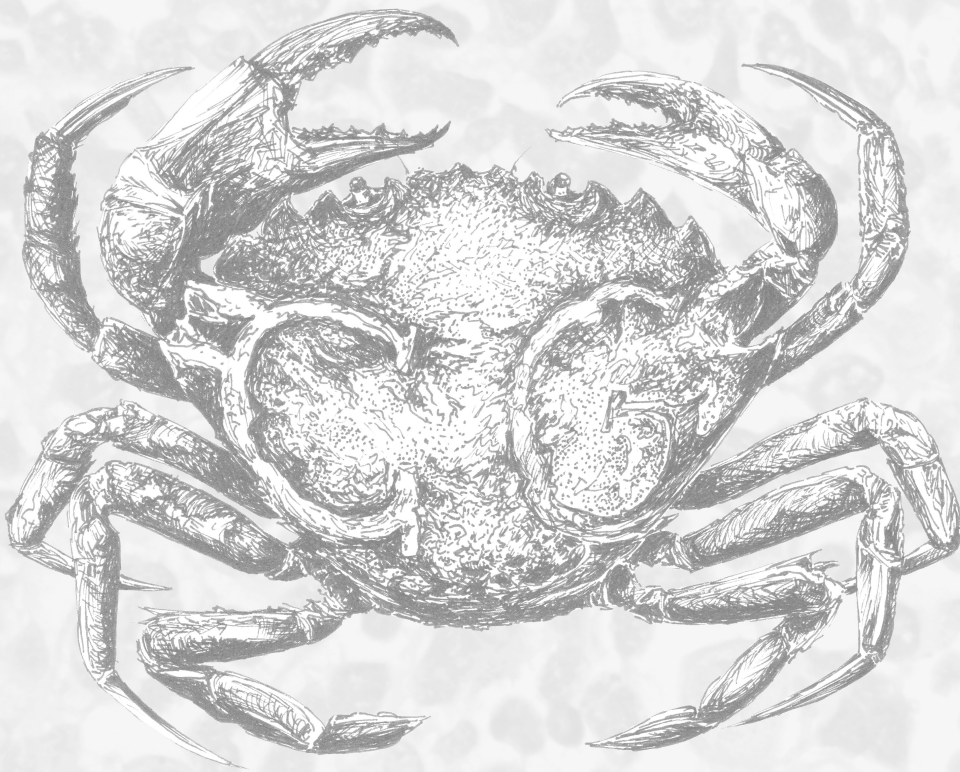
Neuroblastoom (NB) is een kindertumor die ontstaat uit immature sympathische zenuwcellen en wordt gekenmerkt door een grote klinische heterogeniteit met een prognose variërend van zeer gunstig tot fataal. Bij diagnose worden de patiënten ingedeeld in een specifieke risicogroep, een classificatie die gebaseerd is op combinaties van specifieke klinische en genetische karakteristieken (leeftijd van de patiënt, tumorstadium, histologie, differentiatiegraad, *MYCN*-oncogenamplificatie, chromosoom 11q-afwijkingen en DNA-ploidie), om zo voor elke patiënt de meest geschikte therapie te bepalen. Follow-up studies hebben echter aangetoond dat exacte prognosebeoordeling voor NB moeilijk blijft en dat bijkomende prognostische merkers noodzakelijk zijn. Vandaar dat in dit onderzoek werd getracht prognostische tumor-DNA-methylatiebiomerkers te identificeren.

Hiertoe werden de DNA-methylatiepatronen van primaire NB-tumoren op een genomwijde manier in kaart gebracht met behulp van methyl-CpG-binding domain (MBD)-sequencing. Deze techniek is gebaseerd op de verrijking van gemethyleerd DNA, gebruikmakend van de hoge affiniteit van MBD om te binden met gemethyleerde cytosines, gevolgd door sequenceringsanalyse van deze verrijkte DNA-fractie. Ter optimalisatie van de analyseprocedure werd deze techniek initieel toegepast op 8 NB-cellijnen, wat in combinatie met mRNA-expressiestudies leidde tot een eerste selectie van 43 kandidaatbiomerkers. Vervolgens werd methylatie-specifieke PCR (MSP) geïmplementeerd om deze merkers verder te evalueren in 89 primaire NB-tumoren. Zo werden verscheidene nieuwe prognostische DNA-methylatiebiomerkers geïdentificeerd en werden de technologische aspecten en data-analysepijplijn voor een meer uitgebreide biomarkerstudie vastgelegd. In deze vervolgstudie werd het DNA-methyloom van 102 primaire NB-tumoren, die geselecteerd werden op basis van risicoclassificatie en overlevingsstatus, geprofileerd met MBD-sequencing. Differentiële methylatieanalyses tussen de prognostische patiëntengroepen lieten toe 78 kandidaatbiomerkers te prioriteren, die vervolgens op twee onafhankelijke cohortes van 132 en 177 stalen werden getest met behulp van de MSP-pijplijn uit de pilootstudie. De prognostische waarde van verscheidene individuele MSP assays werd gevalideerd en via de implementatie van een nieuw ontwikkeld statistisch kader werd een robuuste prognostische methylatiesignatuur bestaande uit 58 merkers ontwikkeld en gevalideerd. Dit is de grootste DNA-methylatie(biomarker)studie in NB dusver.

De MBD-sequenceringsdata werd publiek beschikbaar gemaakt onder de vorm van een data descriptor. Hierdoor kan de data ook voor andere onderzoeksdoeleinden worden aangewend en kunnen zo nieuwe mechanismes in de NB-pathologie worden ontrafeld. Ter illustratie hiervan werd het DNA-methyloom van stadium 4S-NB gekarakteriseerd. Dit tumorstadium komt enkel voor bij jonge kinderen en wordt getypeerd door metastasen die doorgaans spontaan regresseren, met een gunstige prognose tot gevolg. Differentiële promotormethylatieanalyses tussen stadium 4S, stadium 1/2 (gelokaliseerde ziekte met gunstige prognose) en stadium 4 (metastatische ziekte met ongunstige prognose) toonden aan dat specifieke chromosomale regio's rijk zijn aan promotoren differentieel gemethyleerd in stadium 4S, en dat dit tumorstadium gekenmerkt wordt door hypermethylatie van specifieke promotoren in subtelomeren. Verder bleek dat genen die een belangrijke rol spelen in kanker, in neurale kamontwikkeling en -differentiatie, en in epigenetische processen differentieel gemethyleerd zijn in stadium 4S.

Samengevat kon door exploratie van het NB-DNA-methyloom niet alleen worden aangetoond dat DNA-methylatieprofielen de NB-biologie bepalen, maar konden eveneens bijkomende klinisch relevante prognostische biomerkers worden gevonden.

1



Introduction

1.1 DNA methylation: inherited information beyond the DNA sequence

Deoxyribonucleic acid (DNA) occurs as a double helix of which each strand is made up of a sequence of nucleotides consisting of a phosphate and a sugar linked to one of the four following bases: cytosine (C), thymine (T), adenine (A) or guanine (G). Hydrogen bonds between the A-T and C-G base pairs (bp) hold the two strands together. The DNA is further intimately complexed with various specialized proteins, which together form chromatin. Nucleosomes form the fundamental repeating units of chromatin and consist of 147 bp of DNA wrapped around a histone octamer (Figure 1). Additional levels of higher-order chromatin organization lead to further packaging into chromatin fibers, larger looped chromatin domains and eventually to the formation of chromosomes. The human genome is organized into two sets of 23 chromosomes, of which one set is inherited from each parent.

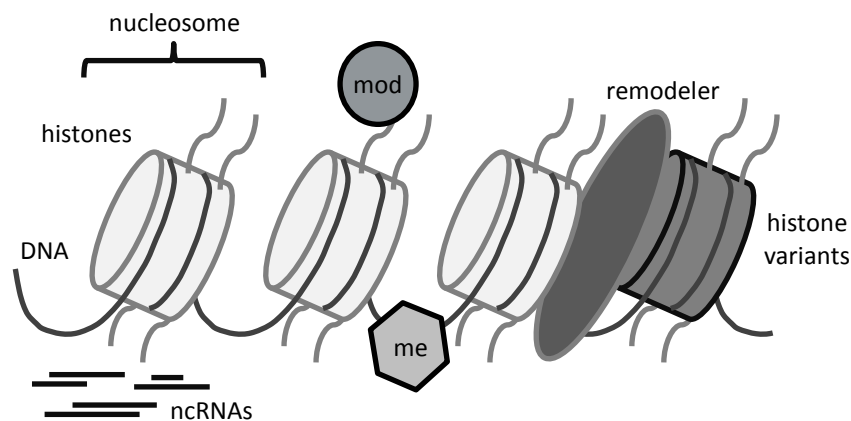


Figure 1. Various epigenetic mechanisms contribute to modulation of the chromatin structure. me: DNA methylation; mod: histone modification; ncRNA: non-coding RNA. Based on [1].

Chromatin is a dynamic molecule existing in many configurations, ranging from highly condensed chromatin (heterochromatin) to a less compacted type (euchromatin) where genes are typically expressed. These chromatin variations are established through various so-called epigenetic mechanisms (Figure 1), including the dynamic shuffling of histone variants in and out of nucleosomes, chromatin remodeling resulting in nucleosome repositioning, the addition of chemical flags to histone proteins, the targeting role of non-coding ribonucleic acids (ncRNAs) and DNA methylation. In concert with transcription factor regulation, these epigenetic mechanisms collectively effectuate different patterns of gene expression and silencing from the same genome, creating different cell types [1]. Importantly, this epigenetic machinery needs to be well orchestrated to avoid that cellular identity gets lost, leading to malignancies such as cancer. Epigenetics have therefore heralded a new era of cancer research, in which it has become clear that something beyond the DNA sequence can be inherited and disrupted. These epigenetic alterations can be identified and used both in a diagnostic and therapeutic setting.

1.1.1 Principles and functions of DNA methylation

1.1.1.1 DNA methylation and demethylation

DNA methylation is the addition of a methyl group (CH_3) to the C5 position of cytosines in CpG dinucleotides (the 'p' refers to the phosphodiester bond between the cytosine and the guanine) and occurs in 60-80% of the 56 million CpG dinucleotides in the human genome, which corresponds to 4-6% of all cytosines. CpG dinucleotides are underrepresented in relation to other dinucleotides, because of the hydrolytic deamination of 5-methylcytosine to thymine during molecular evolution, and are unequally distributed across the genome. Depletion of CpGs is observed in intergenic and intragenic regions, whereas repetitive DNA and CpG islands are CpG-rich [2]. CpG islands are defined as regions of at least 200 bp that have a GC content of 50% and an observed/expected CpG ratio of more than 0.6 [3]. For comparison, the observed/expected CpG ratio in the bulk of the genome is 0.1-0.2. DNA methylation is achieved by DNA methyltransferases (DNMTs), that catalyse the transfer of the methyl group from S-adenosyl-L-methionine (SAM) to the cytosine C5 position. The initial DNA methylation pattern is set during embryonic development (Figure 2) by the *de novo* DNMT3A and DNMT3B DNA methyltransferases. After each round of DNA replication, hemimethylated DNA is obtained, in which only the parental strand carries methylation marks. The newly synthesized daughter strand, which is initially devoid of methylation, then gets methylated by the maintenance DNA methyltransferase DNMT1. The preference of DNMT1 for hemimethylated sites prevents methylation of previously unmethylated sites and thereby preserves the DNA methylation pattern through cell division [2].

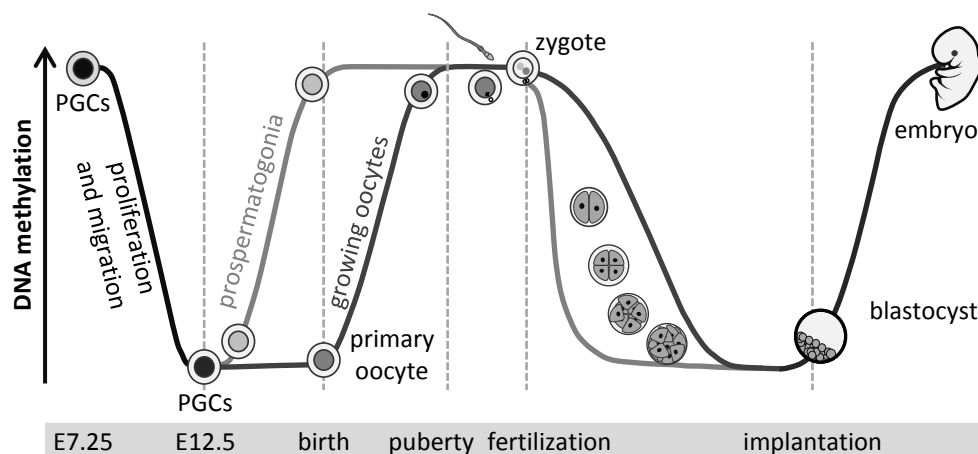


Figure 2. DNA methylation levels change dynamically during human development. Primordial germ cells (PGCs) emerge in embryos at embryonic day 7.25. Concomitant with their proliferation and migration towards the genital ridge, DNA methylation is globally erased. Following sex determination, new DNA methylation landscapes are established in the germ cell precursors of the male or female embryo. In males, this *de novo* DNA methylation is completed before birth. In females, *de novo* DNA methylation is established after birth, during the follicular/oocyte growth phase. Following fertilization, a new wave of DNA demethylation takes place that is distinct on the parental genomes. In the zygote, DNA methylation of the paternal genome is rapidly erased by an active mechanism. Demethylation of the maternal genome is slower and dependent on DNA replication (passive demethylation). During blastocyst implantation and cell lineage determination, new methylation landscapes are established, associated with cellular differentiation. Based on [4].

Although DNA methylation has been viewed as a stable epigenetic mark, recent studies have indicated that it is more dynamic than originally thought. DNA demethylation has been observed in

specific contexts, for example during specific developmental stages (Figure 2), and can occur through passive or active mechanisms. If DNMT1 is inhibited or absent when cells divide, the newly synthesized DNA strands will not be methylated and successive rounds of cell division will result in passive DNA demethylation. Additionally, several modes of active DNA demethylation have been described, which include enzymatic DNA demethylation, (deamination of 5-methylcytosine to thymine followed by) base excision repair, nucleotide excision repair and mechanisms influencing the availability or biochemical features of SAM [5]. Key players in the enzymatic DNA demethylation processes are the Ten-eleven translocation proteins TET1, TET2 and TET3. These enzymes generate 5-hydroxymethylcytosine from existing 5-methylcytosine, which they can further process to 5-formylcytosine and 5-carboxylcytosine (Figure 3) [6]. Subsequently, these oxidation products are recognized by thymine-DNA glycosylase and excised through base excision repair to install an unmodified cytosine [7, 8]. Alternatively, 5-hydroxymethylcytosine can also be converted to 5-hydroxymethyluracil by the AID/APOBEC family of cytidine deaminases, and 5-hydroxymethyluracil is then excised through base excision repair (Figure 3) [9, 10]. Apart from their role as intermediates in this active DNA demethylation pathway, several studies have also indicated direct functions of 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine. As a consequence, these epigenetic marks are increasingly being studied, also in cancer [6, 9, 11–16].

1.1.1.2 DNA methylation regulates gene expression

Gene expression is regulated by a sophisticated interplay between transcription factors (reviewed elsewhere [17]) and numerous epigenetic actors, in which the precise role of DNA methylation has proved to be challenging to unravel, as its function seems to vary with genomic context [18]. Initial DNA methylation studies focused on CpG islands at promoters and postulated that active promoters, characterized by nucleosome-depleted regions (NDRs) upstream of their transcription start sites (TSSs), normally lack DNA methylation. These promoters are further marked by the presence of specific histone modifications, such as H3 trimethylation at lysine 4 (H3K4me3) and lysine acetylation, and histone variant H2A.Z. On the other hand, genes of which the promoter lacks H2A.Z and has nucleosomes positioned over the TSS harbouring repressive histone marks (H3K27me3/H3K9me2/H3K9me3), can be stably silenced by DNA methylation [19]. Nevertheless, most promoters remain unmethylated and the transcriptional level of these genes is regulated by transcription factors [17]. Importantly, about 40% of human genes does not show CpG islands at their promoters, and because of the long-standing focus on CpG islands, the role of promoter methylation in controlling transcription of these genes is not yet fully understood. Recent studies further indicate that methylation of CpG island shores and shelves, i.e. regions of relatively low CpG density that flank CpG islands, as well as methylation of gene bodies and regulatory sites, such as enhancers and insulators, also play a role in fine-tuning gene expression. Although the majority of the gene bodies is CpG-poor, they are extensively methylated and, in contrast to promoter methylation, this is generally positively correlated with gene expression. However, the detailed mechanism by which DNA methylation of these different genomic regions orchestrates transcription regulation remains to be elucidated [18]. Systematic investigation of DNA methylation patterns at genome-wide level has been hampered by the presence of a large number of epigenomes; each individual has essentially one genome, but each cell type in each individual is believed to have a distinct epigenome. Therefore, in analogy to the Human Genome Project, several large-scale epigenomics projects have been initiated (Box 1), to produce a public resource of epigenome data

that includes maps of DNA methylation, histone modifications, chromatin accessibility and RNA expression of cell types and tissues relevant to important biological processes and diseases. These genome-wide data will indisputably improve our understanding of gene expression regulation and the role of DNA methylation herein [20, 21].

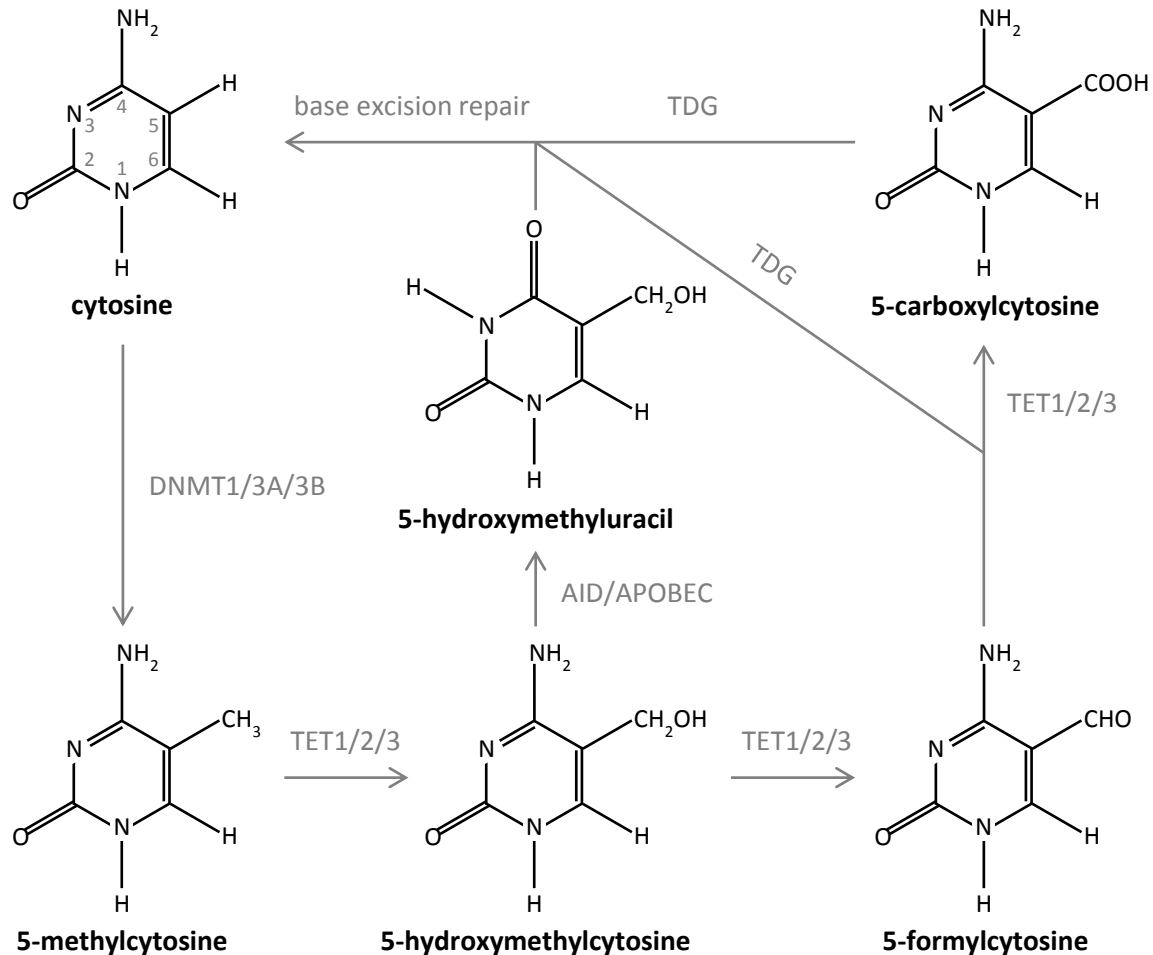


Figure 3. Enzymatic DNA demethylation occurs via multiple intermediates. DNMT proteins are responsible for the methylation of cytosines. Two pathways of active demethylation have been described. Most evidence exists for a pathway in which TET proteins convert 5-methylcytosine into 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine through three consecutive oxidation reactions. Subsequently, 5-formylcytosine and 5-carboxylcytosine are recognized by thymine-DNA glycosylases (TDGs) which activate the base excision repair pathway. In addition, evidence exists for a pathway in which AID/APOBEC proteins deaminate 5-hydroxymethylcytosine to 5-hydroxymethyluracil, followed by TDG-mediated base excision repair. Based on [9].

Box 1. Several large-scale epigenomics projects have been initiated [20, 21].

The *International Human Epigenome Consortium (IHEC)* was launched in 2010 to offer a forum for coordination of other initiatives, with the objective of avoiding redundant research efforts, implementing high data quality standards and thus maximizing efficiency among scientific research centers. Several projects have been initiated, each contributing to the IHEC's major goal - mapping 1,000 epigenomes by 2020 (www.ihec-epigenomes.org). The *BLUEPRINT* project is a European initiative focusing on distinct types of hematopoietic cells from healthy individuals and their malignant leukemic counterparts (www.blueprint-epigenome.eu). The *German epigenome project DEEP* analyzes cell types connected to complex diseases, such as metabolic diseases and inflammatory diseases of the joints and the intestine (www.deutsches-epigenom-programm.de). The following categories of human cell and tissue types are studied by the *Canadian Epigenetics, Environment and*

Health Research Consortium (CEEHRC): stem and differentiated cell types relevant to complex diseases (including cancer), placental tissue and tissues relevant to disorders of genomic imprinting, central nervous system tissues, tissues relevant to cancer heterogeneity, breast cancer, tissues relevant to cardiovascular diseases and tissues related to human immune function (www.cihr-irsc.gc.ca/e/43602.html). The *Core Research for Evolutional Science and Technology (CREST)/IHEC, Team Japan* aims to produce reference epigenomes of gastrointestinal epithelial cells, vascular endothelial cells and cells of reproductive organs (www.crest-ihec.jp/english/index.html). The *National Institutes of Health (NIH) Roadmap Epigenomics Program* in the US, which was already launched in 2008, now also contributes to the IHEC (www.roadmapepigenomics.org), and also the *Encyclopedia of DNA Elements (ENCODE) Consortium* maps epigenomes (www.encodeproject.org).

1.1.2 DNA methylation in normal cellular processes

1.1.2.1 Subtelomeric and pericentromeric repeat silencing

Telomeres are nucleoprotein structures at the ends of chromosomes and consist of TTAGGG tandem repeats bound by a multiprotein complex known as shelterin. Telomeres are essential for chromosome end protection and chromosomal stability. Subtelomeres, located proximal to telomeres, are also enriched in repetitive DNA and are further characterized by a low gene density and the presence of CpG dinucleotides, that are absent in telomeres. Pericentromeres flank chromosome centromeres and are involved in kinetochore function and sister chromatid cohesion during cell division. These three types of repeat-rich chromatin regions exhibit histone modifications commonly found in heterochromatin, such as H3K9me3 and H4K20me3, and are enriched in heterochromatin protein HP1. Additionally, subtelomeres and pericentromeres are also typified by DNA methylation. These heterochromatic marks have important functions in chromosome segregation, act as a negative regulator of homologous recombination in these regions, and suppress telomere elongation [22].

1.1.2.2 Transposon silencing

The human genome is strewn with transposons, also called transposable elements, DNA sequences that can migrate ('transpose') within the genome. Depending on their mechanism of transposition, transposons are grouped into two categories: DNA transposons and retrotransposons. While DNA transposons move around by a cut-and-paste mechanism, retrotransposons are spread by a copy-and-paste mechanism, which involves transcription into an RNA intermediate and integration of the reverse-transcribed cDNA copy at a new locus. Active transposons are highly mutagenic and can influence neighbouring genes by altering splicing and polyadenylation patterns, or by functioning as enhancers or promoters. To prevent transposon-induced damage to the genome, defense mechanisms have been developed, including suppression of transcriptional activity of transposons by DNA methylation and chromatin modifications, such as H3K9 methylation [23, 24].

1.1.2.3 X chromosome inactivation

Dosage compensation for X-linked gene products between the sexes is achieved by X-inactivation in females, which involves multiple levels of epigenetic modifications. X-inactivation is initiated at the X-inactivation center (XIC) which contains the long non-coding RNA X-inactive-specific-transcript (*XIST*). During early differentiation, *XIST* is monoallelically expressed and binds to high affinity sites on the

chromosome from which it is transcribed, resulting in *XIST* RNA coating of the future inactive X chromosome. Although the precise mechanisms underlying the repressive effect of *XIST* RNA are unknown, *XIST* RNA coating is followed by loss of euchromatic histone marks (H3K4me2/3, H3K9Ac and H4Ac). During this time window, X-linked gene silencing initiates by accumulation of the PRC2 and PRC1 complex proteins, causing enrichment of repressive histone modifications (H3K27me3, H3K9me2, H2Aub1 and H4K20me1). At the final differentiation stage, in the maintenance phase, the PRC2 and PRC1 complexes no longer appear to be present and macroH2A becomes associated with the imprinted X chromosome. Finally, DNA methylation marks the promoters of X-linked genes [25].

1.1.2.4 Genomic imprinting

Genomic imprinting is an epigenetic phenomenon that results in monoallelic gene expression according to parental origin. Most of the approximately 80 known imprinted genes (e.g. *DLK1*, *GNAS*, *GRB10*, *IGF2* and *SNURF*) occur in clusters that contain 2 to 15 genes and that vary in size from less than 100 kilobase to several megabases [26]. The parent-specific expression of these gene clusters is under the overall control of a *cis*-acting imprinting control region (ICR), which shows parental allele-specific DNA methylation and chromatin modifications. DNA methylation of the ICR is acquired in either maternal or paternal germ cells and is resistant to the extensive reprogramming of the genome that occurs in the embryo after fertilization (Figure 2). Imprinted genes have major effects on prenatal and postnatal development, survival and growth, as well as on metabolism, and neural and behavioural processes. Alterations of DNA methylation in imprinted regions may lead to disease. For example, lack of maternal *GNAS* methylation imprinting results in pseudohypoparathyroidism type 1b [26].

1.1.3 Cancer: when DNA methylation goes awry

It is now generally accepted that epigenetic aberrations, next to genetic lesions, contribute to cancer initiation and progression. Cancer cells present a profoundly distorted epigenetic landscape in which multiple epigenetic players can be affected, including DNA methylation [27]. Since DNA methylation alterations are commonly observed in benign neoplasms and early-stage tumors, epigenetic deregulation has been considered an early event in tumorigenesis which may precede the classical genetic changes [28]. Of note, DNA methylation patterns are directly affected by a person's diet and exogenous stimuli such as viruses and bacteria, which may initiate cancer formation [29]. However, recent evidence indicates that deregulation of epigenetic and genetic mechanisms are not separate events in cancer, but that they intertwine and take advantage of each other; genetic alterations of the epigenome may contribute to tumorigenesis and epigenetic processes may cause point mutations and disable DNA repair functions [30].

1.1.3.1 Hypomethylation

One of the first epigenetic alterations found in human cancer is the presence of massive global loss of DNA methylation. This global DNA hypomethylation contributes in several ways to the development of cancer (Figure 4). It mainly occurs at repetitive sequences of subtelomeric and pericentromeric regions, promoting chromosomal instability and rearrangements, mitotic recombination and aneuploidy. Additionally, DNA hypomethylation can lead to reactivation of

transposons. An example hereof is the *LINE-1* retrotransposon, which is silenced in normal cells, but becomes hypomethylated and transcriptionally reactivated in breast, lung, bladder and liver tumors. Also, loss of imprinting due to DNA hypomethylation increases the risk of cancer. Loss of imprinting of *IGF2*, for example, is accompanied by an increased risk for colorectal cancer and contributes to the development of Wilm's tumor. Finally, DNA hypomethylation can activate aberrant expression of alternative transcripts from a gene, due to loss of DNA methylation in coding regions and introns, and can induce the expression of oncogenes [27, 29, 31].

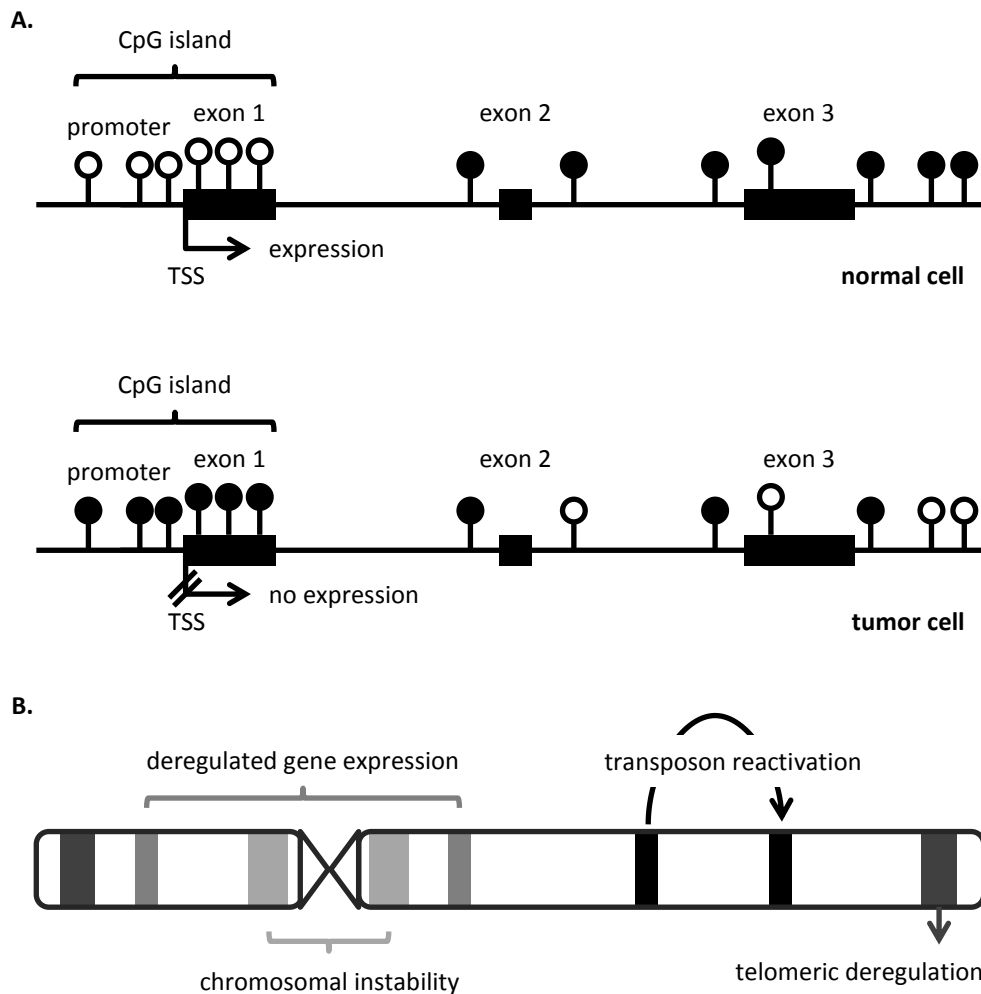


Figure 4. DNA methylation alterations contribute to carcinogenesis via several mechanisms. A. Lollipop diagram of a normal and tumor cell. DNA methylation states are indicated as white (unmethylated) or black (methylated) lollipops (CpG dinucleotides). **B.** These alterations can lead to deregulated expression (of onco- and tumor suppressor genes), transposon reactivation, chromosomal instability and telomeric deregulation. TSS: transcription start site. Based on [32].

1.1.3.2 Hypermethylation

Another common epigenetic feature of human cancers is hypermethylation at CpG islands of promoters (Figure 4). Remarkably, only specific promoters become hypermethylated, while others remain unmethylated, and this hypermethylation pattern is tumor-specific. Transcriptional inactivation caused by promoter hypermethylation frequently affects tumor suppressor genes and genes involved in the main cellular pathways: DNA repair (*MLH1*, *MGMT*, *WRN*, *BRCA1*), vitamin

response (*RARB*, *CRBP1*), RAS signaling (*RASSF1*, *RASSF5*), cell cycle control (*CDKN2A*, *CDKN2B*, *RB1*), TP53 network (*TP73*, *HIC1*) and apoptosis (*PYCARD*, *DAPK1*, *WIF1*, *SFRP1*). Recent findings further indicate that aberrant DNA methylation of CpG island shores (e.g. at *HOXA2* and *GATA2*) also frequently occurs in cancer [27]. Importantly, monoallelic DNA methylation silencing may act as one of the two hits in addition to a genetic lesion to fulfill Kundson's two-hit hypothesis, which states that disruption of gene function requires loss of both copies of the involved gene. An example of this close epigenetic-genetic cooperation is observed in the colon cancer cell line HCT116, in which one allele of *MLH1* and *CDKN2A* is genetically mutated, whereas the other is silenced by DNA methylation. Finally, DNA methylation can induce the generation of disease-causing mutations, due to the spontaneous hydrolytic deamination of 5-methylcytosine to thymine [30].

1.1.3.3 Genetic lesions of the DNA methylation machinery

Next to mutations in histone-modifying enzymes, also enzymes involved in the DNA methylation machinery can be genetically perturbed in cancer. *DNMT3A* mutations have been reported in acute myeloid leukemia and are associated with poor prognosis, and *DNMT1* mutations are detected in colorectal cancer [30].

1.1.4 A plethora of technologies to illuminate DNA methylomes

As 5-methylcytosine and cytosine roughly have the same base pairing characteristics, standard sequence detection technologies cannot discriminate between them. To overcome this, several DNA modification and preparation steps have been introduced to enable DNA methylation analysis. The most frequently applied methodologies are based on the use of (methylation state) restriction enzymes, precipitation of methylated DNA fragments and/or DNA bisulfite treatment (Table 1).

Restriction enzyme-based DNA methylation analyses make use of methylation-sensitive restriction enzymes, of which the endonuclease activity is influenced (in most cases inhibited) by methylation of the CpG(s) in the recognition site of the enzyme. Subsequent analysis of the restriction pattern then reveals DNA methylation information. The most widely used methylation-sensitive restriction enzymes are *HpaII* (recognition sequence C^yCGG) and *SmaI* (recognition sequence CCC^yGGG), because they each have an isoschizomer (*MspI* for *HpaII*) or neoschizomer (*XmaI* for *SmaI*) that is not inhibited by CpG methylation. The very first analysis methods used gel electrophoresis or Southern blots, but later on, more advanced technologies, such as (quantitative) polymerase chain reaction ((q)PCR), arrays and sequencing, were combined with methylation-sensitive restriction enzymes (Table 1) [33]. It should be noted that some methylation-sensitive restriction enzymes are also sensitive to hydroxymethylation (§1.1.1.1), possibly confounding DNA methylation analyses [34].

Bisulfite conversion-based methods rely on the detection of chemically induced methylation-specific nucleotide changes in the DNA sequence, making use of a difference in reaction kinetics between cytosine and 5-methylcytosine with bisulfite. Deamination by bisulfite proceeds by three steps (Figure 5). The process begins with the nucleophilic addition of HSO_3^- to the C6 of the pyrimidine ring (sulfonation step), which can only occur at (5-methyl)cytosines not involved in base-pairing, as in double-stranded DNA the bases are locked in the *anti*-conformation in which C6 is sterically impeded. The amino group at position 4 is destabilized by this sulfonation, so that hydrolytic

Table 1. Multiple DNA methylation detection methodologies, making use of restriction enzymes, affinity enrichment and/or bisulfite conversion, have been developed.

method	ref.	year of publication	Google Scholar citations (d.d. 01/07/2016)	restriction enzyme	analysis strategy		detection technology		
					bisulfite conversion	affinity enrichment	PCR and/or gel/ electrophoresis	array	sequencing
RLGS	[35]	1991	388	✓			✓		
MS-AP-PCR	[36]	1997	229	✓			✓		
MSRF	[37]	1997	149	✓			✓		
MCA-RDA	[38]	1999	513	✓			✓		
AIMS	[39]	2002	121	✓			✓		
MS-MLPA	[40]	2005	314	✓			✓		
DMH	[41]	1999	470	✓				✓	
<i>Mcr</i> BC/array	[42]	2004	67	✓				✓	
MethylScope	[43]	2005	119	✓				✓	
MMASS	[44]	2006	43	✓				✓	
MSNP	[45]	2006	60	✓				✓	
MCAM	[46]	2007	128	✓				✓	
CHARM	[47]	2008	337	✓				✓	
HELP assay	[48]	2006	348	✓				✓	
high-resolution HELP assay	[49]	2009	113	✓				✓	✓
MSDK	[50]	2006	29	✓					✓
Methyl seq	[51]	2009	211	✓					✓
MSCC	[52]	2009	617	✓					✓
Methyl-MAPS	[53]	2010	102	✓					✓
MRE seq	[54]	2010	809	✓					✓

(continues)

Note. This overview is limited to methods making use of restriction enzymes, bisulfite conversion and/or affinity enrichment, in combination with PCR and/or gel/electrophoresis, array and/or sequencing. Technologies indicated with * are discussed in more detail in §1.1.4.1, §1.1.4.2, §1.1.4.3 and §1.1.4.4. AIMS: amplification of intermethylated sites; CHARM: comprehensive high-throughput arrays for relative methylation; DMH: differential methylation hybridization; HELP: *Hpa*II tiny fragment enrichment by ligation-mediated PCR; MCAM: methylated CpG island amplification microarray; MCA-RDA: methylated CpG island amplification with representational difference analysis; Methyl-MAPS: methylation mapping analysis by paired-end sequencing; MMASS: microarray-based methylation assessment of single samples; MRE seq: methylation-sensitive restriction enzyme sequencing; MS-AP-PCR: methylation-sensitive arbitrarily primed PCR; MSCC: methyl-sensitive cut counting; MSDK: methylation-specific digital karyotyping; MS-MLPA: methylation-specific multiplex ligation-dependent probe amplification; MSNP: methylation-sensitive single nucleotide polymorphism chip analysis; MSRF: methylation-sensitive restriction fingerprinting; PCR: polymerase chain reaction; RLGS: restriction landmark genomic scanning.

Table 1. Multiple DNA methylation detection methodologies, making use of restriction enzymes, affinity enrichment and/or bisulfite conversion, have been developed.
(continued)

method	ref.	year of publication	Google Scholar citations (d.d. 01/07/2016)	restriction enzyme	analysis strategy		detection technology		
					bisulfite conversion	affinity enrichment	PCR and/or gel/ electrophoresis	array	sequencing
COBRA	[55]	1997	1139	✓	✓		✓		
RRBS*	[56]	2005	436	✓	✓				✓
scRRBS*	[57]	2013	79	✓	✓				✓
RRMAB seq	[58]	2016	0	✓	✓				✓
MB-PCR	[59]	2006	57	✓		✓	✓		
MSP*	[60]	1996	6032		✓		✓		
MS-DGGE	[61]	1999	131		✓		✓		
MethyLight*	[62]	2000	1151		✓		✓		
MS-MCA	[63]	2001	168		✓		✓		
Ms-SNuPE	[64]	2002	65		✓		✓		
MS-SSCA	[65]	2002	18		✓		✓		
HeavyMethyl*	[66]	2004	131		✓		✓		
MS-HRM	[67]	2007	316		✓		✓		
MethyLight ddPCR	[68]	2015	4		✓		✓		
HT-TREBS	[69]	2014	6		✓		✓		✓
BisPCR ²	[70]	2015	2		✓		✓		✓
MSO	[71]	2002	334		✓			✓	
GoldenGate assay*	[72]	2006	510		✓			✓	
Illumina 27 K array*	[73]	2009	338		✓			✓	

(continues)

Note. This overview is limited to methods making use of restriction enzymes, bisulfite conversion and/or affinity enrichment, in combination with PCR and/or gel/electrophoresis, array and/or sequencing. Technologies indicated with * are discussed in more detail in §1.1.4.1, §1.1.4.2, §1.1.4.3 and §1.1.4.4. COBRA: combined bisulfite restriction analysis; ddPCR: digital droplet PCR; HT-TREBS: high-throughput targeted repeat element bisulfite sequencing; MAB seq: methylation-assisted bisulfite sequencing; MB-PCR: methyl-binding PCR; MS-DGGE: methylation-specific denaturing gradient gel electrophoresis; MS-HRM: methylation-sensitive high-resolution melting; MS-MCA: methylation-specific melting curve analysis; MSO: methylation-specific oligonucleotide; MSP: methylation-specific PCR; Ms-SNuPE: methylation-sensitive single nucleotide primer extension; MS-SSCA: methylation-sensitive single strand conformation analysis; PCR: polymerase chain reaction; RRBS: reduced representation bisulfite sequencing; RRMAB seq: reduced-representation MAB seq; scRRBS: single cell RRBS.

Table 1. Multiple DNA methylation detection methodologies, making use of restriction enzymes, affinity enrichment and/or bisulfite conversion, have been developed.
(continued)

method	ref.	year of publication	Google Scholar citations (d.d. 01/07/2016)	restriction enzyme	analysis strategy		detection technology		
					bisulfite conversion	affinity enrichment	PCR and/or gel/ electrophoresis	array	sequencing
Illumina 450 K array*	[74]	2011	533		✓			✓	
TAB-array	[75]	2014	8		✓			✓	
MSBE	[76]	2005	8		✓				✓
BS seq	[77]	2008	1211		✓				✓
scBS seq*	[78]	2014	112		✓				✓
scWGBS*	[79]	2015	35		✓				✓
μWGBS*	[79]	2015	35		✓				✓
MethylC seq*	[80]	2008	1430		✓				✓
BSPP	[81]	2009	396		✓				✓
LHC-BS	[82]	2011	15		✓				✓
LD-BSP	[83]	2015	1		✓				✓
MAB seq	[58]	2016	0		✓				✓
MeDIP chip	[84]	2005	1335			✓		✓	
MeCP2 chip*	[85]	2006	1129			✓		✓	
MBD2 chip*	[86]	2006	166			✓		✓	
MIRA	[87]	2010	40			✓		✓	✓
MeDIP seq	[88]	2008	474			✓			✓
MethylCap seq*	[89]	2010	134			✓			✓
MiGS*	[90]	2010	239			✓			✓

Note. This overview is limited to methods making use of restriction enzymes, bisulfite conversion and/or affinity enrichment, in combination with PCR and/or gel/electrophoresis, array and/or sequencing. Technologies indicated with * are discussed in more detail in §1.1.4.1, §1.1.4.2, §1.1.4.3 and §1.1.4.4. BSPP: bisulfite sequencing padlock probes; BS seq: bisulfite sequencing; LD-BSP: limiting dilution bisulfite pyrosequencing; LHC-BS: liquid hybridization capture-based bisulfite sequencing; MAB seq: methylation-assisted bisulfite sequencing; MBD2: methyl-CpG-binding domain protein 2; MeCP2: methyl-CpG-binding protein 2; MeDIP: methylated DNA immunoprecipitation; MethylC seq: cytosine methylome sequencing; MiGS: MBD-isolated genome sequencing; MIRA: methylated CpG island recovery assay; MSBE: methylation-specific single base extension; scBS seq: single cell BS seq; scWGBS: single cell WGBS; TAB-array: TET-assisted bisulfite conversion with array analysis; WGBS: whole-genome bisulfite sequencing; μWGBS: WGBS in very small cell populations.

liberation of NH_3 takes place, the rate-limiting key deamination step. On treatment with alkali, HSO_3^- is then again released to regenerate the 5,6-double bond (desulfonation step). In this way, both cytosine and 5-methylcytosine can be converted to either uracil or thymine, respectively, but the fact that the deamination reactions of sulfonated cytosine and 5-methylcytosine proceed at very different reaction rates, such that the deamination of cytosine will be complete before substantial 5-methylcytosine deamination has occurred, allows distinguishing these two bases in DNA. Namely, under specific reaction conditions, bisulfite treatment only converts cytosines to uracils, while methylated cytosines remain unchanged. These methylation-specific induced single nucleotide changes can subsequently be analyzed using various molecular techniques (Table 1) [91, 92]. Importantly, bisulfite-based methods cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine, as 5-hydroxymethylcytosine is also resistant to bisulfite conversion [34, 93].

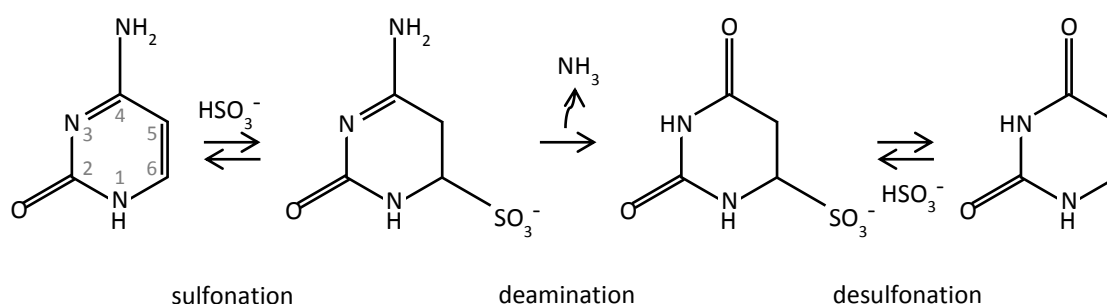


Figure 5. Bisulfite converts cytosine to uracil. The conversion process proceeds in three steps: sulfonation, deamination and desulfonation.

Methodologies making use of affinity enrichment are based on enrichment of methylated DNA using antibodies specific for 5-methylcytosine or using methyl-binding proteins with affinity for methylated genomic DNA. The most frequently used proteins to precipitate methylated DNA are methyl-CpG-binding domain protein 2 (MBD2) and methyl-CpG-binding protein 2 (MeCP2). After enrichment, the DNA fraction is analyzed using array or sequencing technology (Table 1).

1.1.4.1 Methylation-specific PCR

Historically seen, methylation-specific PCR (MSP) was developed to encounter disadvantages of earlier established DNA methylation detection methodologies, such as Southern blotting and PCR-based approaches after digestion of the DNA with methylation-sensitive restriction enzymes. Namely, these methods could only shed light on CpG methylation in the restriction sites of the used enzymes and only assess the overall methylation status of CpG islands. In 1996, Herman et al. introduced the first MSP method, applicable to any block of CpG sites in the genome, allowing site-specific methylation analysis. Principally, their method relies on the bisulfite-mediated conversion of cytosines to uracils, followed by PCR with specific primer pairs designed on the resultant modified DNA, and polyacrylamide gel electrophoresis [60]. In order to render a more quantitative format suitable for clinical settings where high-throughput is required, many other researchers soon started to adapt this MSP methodology. The most important adjustments made are the use of fluorescence-based real-time PCR technology (MethylLight) [62], and methylation-specific oligonucleotide blockers and probes for the analysis of very low concentrations of methylated DNA (HeavyMethyl) [66] (Figure 6). Although careful attention should be paid to MSP primer design, the method is relatively easy to perform.

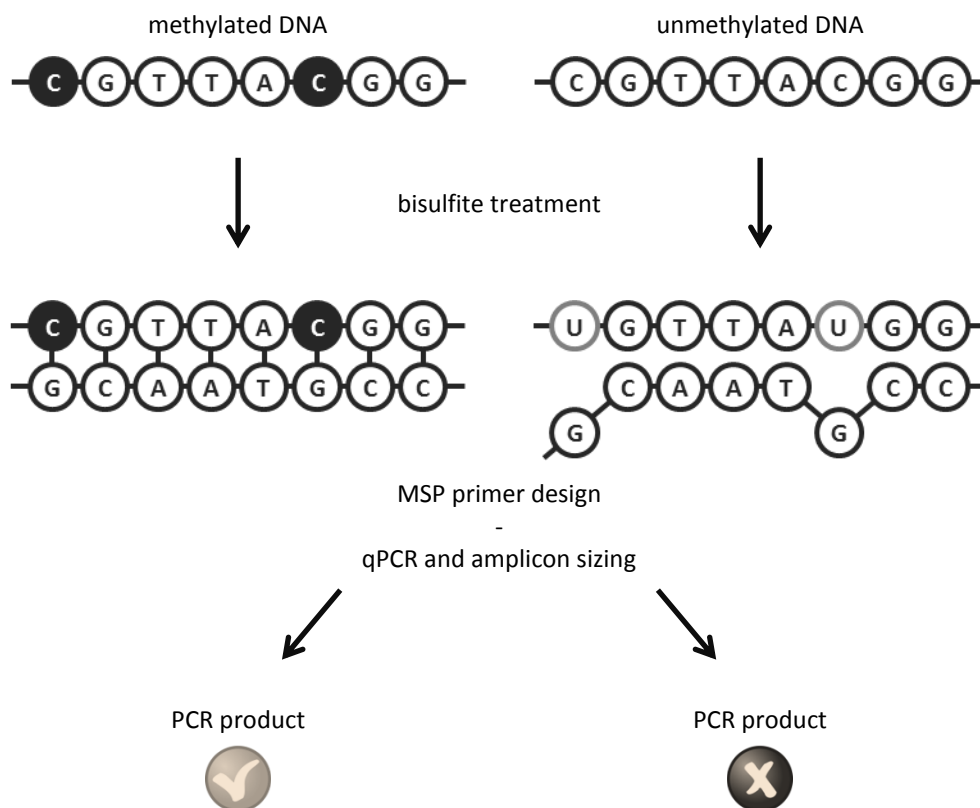


Figure 6. Methylation-specific PCR (MSP) allows site-specific DNA methylation analysis. First, the DNA is treated with bisulfite, converting unmethylated cytosines (white) to uracil. Methylated cytosines (dark) remain as cytosines. MSP primer pairs are designed to only amplify the bisulfite-converted target region and do not anneal to genomic DNA. For simplicity, only one primer and its template are depicted. As each primer contains at least two CpG sites, this means that only if the template is methylated, a PCR product will be generated. qPCR: quantitative polymerase chain reaction.

1.1.4.2 Methylation microarrays

Over the past decade Illumina has developed several DNA methylation analysis technologies that are based on genotyping bisulfite-converted DNA using arrays. First, the GoldenGate Assay for Methylation, implemented on a BeadArray platform, was introduced to investigate the methylation status of 1,536 specific CpG sites in 96 samples simultaneously [72]. Later on, this technology was replaced by the Infinium HumanMethylation27 BeadChip Kit (Illumina 27 K array [73]), that allows a more genome-wide screening of DNA methylation patterns by interrogating approximately 27,000 CpG sites. This kit makes use of a pair of bead-bound probes (Infinium I Assay) per CpG locus to detect the presence of an 'unmethylated' T or 'methylated' C by hybridization, followed by single-base extension with a labelled nucleotide. Then, the array is fluorescently stained, scanned and the intensities of the methylated and unmethylated bead types measured. Afterwards, the methylation status of each CpG site is quantified in a β -value, which represents the ratio of the intensity of the methylated bead type to the combined locus intensity (0 is completely unmethylated; 1 is completely methylated) [74]. Subsequently, the Infinium HumanMethylation450 BeadChip (Illumina 450 K array [74]) and Infinium MethylationEPIC BeadChip Kit were developed, expanding the number of interrogated CpG sites to approximately 480,000 and 850,000, respectively, by making use of an additional assay type (Infinium II Assay). This assay type uses only one bead-bound probe per CpG

locus, of which the 3' terminus complements the base directly upstream of the queried CpG site, allowing hybridization to both 'methylated' and 'unmethylated' alleles. Here, the single-base extension depends on the methylation status of the hybridized allele and results in the addition of a labeled G or A, complementary to either the 'methylated' C or 'unmethylated' T, respectively. As the G and A are differently labeled, β -values can then be calculated by using a dual-color readout for intensity measurement [74, 94].

1.1.4.3 Methyl-CpG-binding domain precipitation

Affinity enrichment-based strategies making use of methyl-CpG-binding domain precipitation rely on the natural ability of methyl-binding proteins, such as MBD2 and MeCP2, to bind methylated CpGs. Methodologically, fragments of sonicated double-stranded DNA are exposed to MBD capture. Methylated fragments are bound, precipitated and subsequently analyzed using array (MeCP2 chip [85], MBD2 chip [86]) or sequencing technology (MethylCap seq [89], MiGS [90]; Figure 7). Since regions with a high density of methylated CpGs are preferentially or more effectively captured, these DNA methylation profiling approaches are biased towards CpG-dense regions, such as CpG islands [89, 90]. As such, when comparing DNA methylation levels of different genomic regions, MBD sequencing is less accurate for methylation quantification, as it measures the relative enrichment of methylated DNA rather than absolute DNA methylation levels. Another disadvantage is that it does not provide single-nucleotide resolution. Nevertheless, MBD sequencing has a higher genomic coverage (approximately 18%) than the Illumina 27 K and 450 K arrays (less than 4%) and can be used to identify differentially methylated regions (DMRs) between different samples (Box 2) [95, 96].

1.1.4.4 Reduced representation and whole-genome bisulfite sequencing

For reduced representation bisulfite sequencing (RRBS), the genomic DNA is first digested using a methylation-insensitive restriction enzyme, usually *MspI* which recognizes the sequence C^YCGG, followed by size selection using gel electrophoresis and band excision to select genomic regions with moderate to high CpG density [97, 98]. This reduced representation (or library) of the genome is then profiled using bisulfite sequencing. RRBS allows single-base resolution, but covers less than 4% of all CpG dinucleotides [96, 99]. In contrast, whole-genome bisulfite sequencing (WGBS) provides single-base resolution and whole-genome coverage. The first human methylomes, originating from embryonic stem cells and fetal lung fibroblasts [100], were reported in 2009 (MethylC seq) and later on also cancer methylomes were determined [101, 102]. Yet, the number of studies using WGBS is still relatively limited, likely due to the high sequencing cost, as sufficient sequencing depth across the full genome is needed to accurately quantitate methylation of individual CpG sites, as well as due to the technical expertise and downstream computation requirements [96]. For example, accurate mapping of bisulfite sequencing reads is challenging, due to the lower sequence complexity and reduced GC content of bisulfite-converted DNA, and requires specific alignment tools [103]. Very recently, RRBS and WGBS that enables DNA methylation mapping in very small cell populations (μ WGBS) and single cells (scBS seq, scWGBS, scRRBS) has been developed [57, 78, 79].

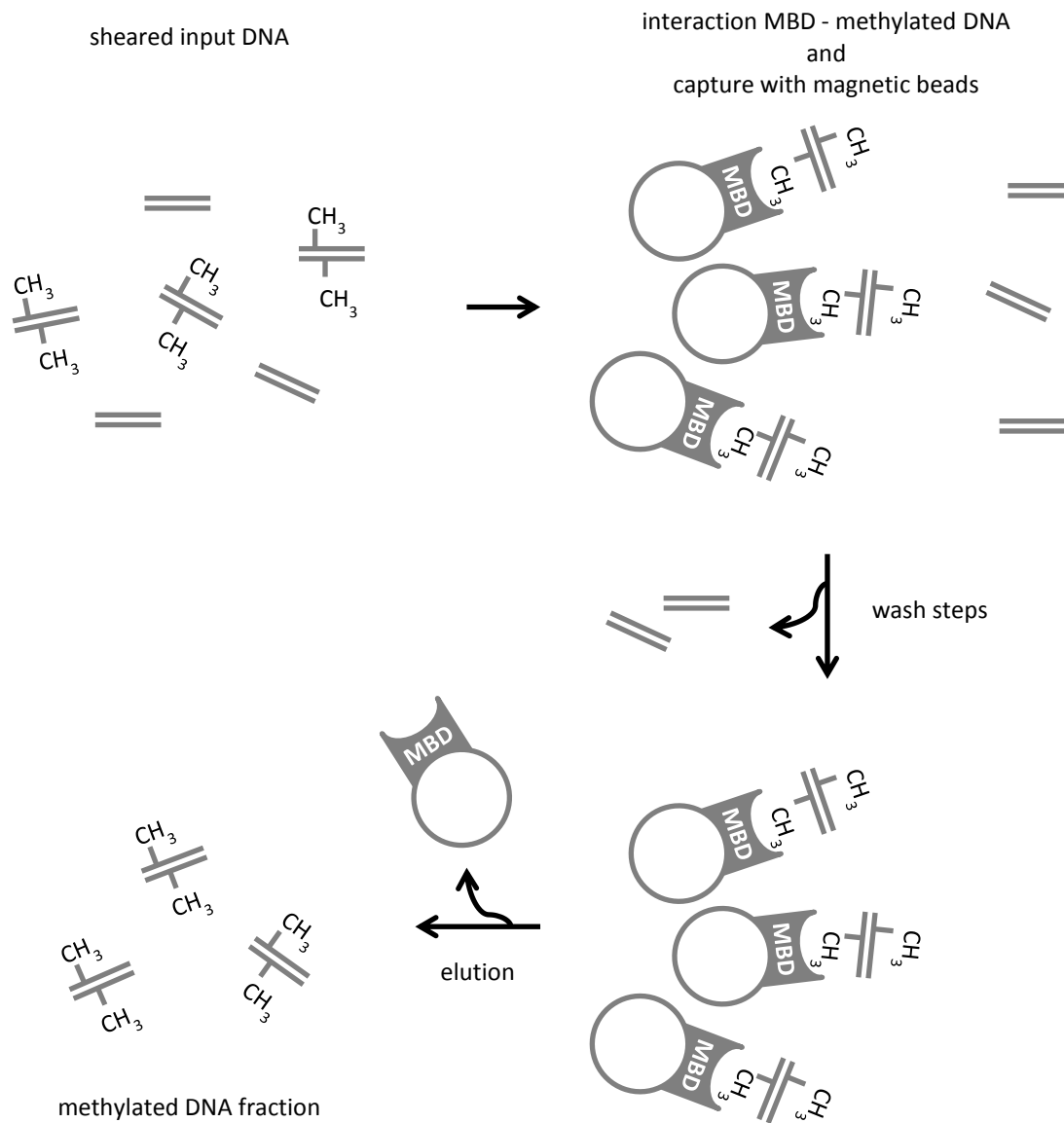


Figure 7. Methyl-CpG-binding domain (MBD) sequencing allows genome-wide DNA methylation analysis. Prior to MBD-based capture, the DNA is sheared or sonicated to create (methylated and unmethylated) DNA fragments. These DNA fragments are subsequently exposed to MBD proteins, which bind methylated DNA. Upon capture with magnetic beads, wash steps are performed to remove unbound (unmethylated) DNA fragments. Finally, the captured DNA fragments are eluted from the beads and can be sequenced. Based on the MethylCap Kit of Diagenode.

Box 2. Differential methylation analyses based on MBD sequencing data make use of several bioinformatics tools.

Differential methylation analyses based on MBD sequencing data make use of a bioinformatics pipeline, consisting of multiple consecutive steps: (1) quality control on the raw sequencing data, (2) read mapping, (3) peak calling, (4) data visualization, and (5) differential methylation analysis. For each of these steps, different bioinformatics tools can be used [104, 105].

First, some *quality control checks on the raw sequencing data* are performed, for example by using FASTQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), in order to detect potential sequencing errors. For example, the per base sequence quality scores and per sequence GC content can be evaluated. The per base sequence quality score reflects the probability that the corresponding base has been called incorrectly,

and can thus be used to check if the base calls are of good quality [106]. For MBD sequencing, evaluation of the per sequence GC content enables a first check of the presence of enrichment towards CpG-dense sequences.

The following step is alignment of the sequences (reads) to the reference genome of the organism of interest, called *read mapping*. A wide variety of mapping tools exists, each with its own specific features. An extensive overview of the most commonly used alignment algorithms and software tools is given by Li et al. [104] and Su et al. [107]. Important aspects to take into account when selecting a mapping tool for MBD sequencing are the performance (speed/accuracy balance), the use of the quality information of the reads and the ability to map short paired-end sequencing reads. An example of a frequently used mapping tool suited for MBD sequencing analysis is Bowtie2 [108].

Once the sequencing reads are mapped to the reference genome, genomic regions significantly enriched upon MBD capture need to be identified. The process of converting mapped sequencing reads to coverage vectors and the detection of enriched regions (peaks) is referred to as *peak calling*. Wilbanks et al. [105] have compared the performance of the most frequently used peak calling programs, including MACS [109], and measured their sensitivity, accuracy and usability. In addition to the location of the identified peaks and their significance scores, MACS also outputs files for *data visualization*, as exemplified in Decock et al. [110].

Differential methylation analyses between sample groups can be performed by using for example the R package DESeq [111], which uses raw counts of sequencing reads as input. For each region of interest included in the analysis, DESeq then yields the mean normalized counts per sample group, as well as the fold change between the sample groups, describing how much the methylation level differs between the two groups, and a p-value for the statistical significance of this difference. Also p-values adjusted for multiple testing with the Benjamini-Hochberg procedure, which controls false discovery rate (FDR), are reported [112].

1.1.5 DNA methylation as clinically useful biomarker

Biomarkers are molecular targets of which quantitative or qualitative analysis enables personalized tailoring of patient management. To be clinically applicable, biomarkers must be specific, sensitive (Box 3 and Table 2) and preferably detectable in specimens obtained through minimally invasive procedures. Several studies have shown that tumor-specific DNA methylation aberrations can be detected in blood or in body fluids that have been in physical contact with the site of the tumor, such as in urine for bladder cancer and in sputum for lung cancer. Additionally, several DNA methylation detection methodologies are applicable to formalin-fixed paraffin-embedded tumor samples. As such, DNA methylation might become the biomarker of choice for the clinical management of cancer patients. Until a few years ago, the search for DNA methylation biomarkers has mainly been focused on promoter-associated CpG islands, where methylation is in most cases inversely correlated to the transcriptional activity of the corresponding gene. However, also methylation of CpGs outside promoter regions should be investigated as potential biomarkers, given the increasing evidence on their involvement in cancer (§1.1.3) [113].

DNA methylation biomarkers can be used in various clinical applications. For example, as changes in the DNA methylation pattern frequently occur early in tumorigenesis (§1.1.3), DNA methylation biomarkers can contribute to the detection of early-stage neoplasia [113]. A study by Palmisano et al. has demonstrated that aberrant DNA methylation of *CDKN2A* in smokers is detectable up to three years before the individuals are diagnosed with squamous cell lung carcinoma, and Scesnaite et al. showed that also never-smokers exposed to second-hand tobacco have a tendency of *CDKN2A* hypermethylation [113–115]. Besides early detection, DNA methylation biomarkers can also be used to characterize and classify cancers, as well as for prognosis prediction. One of the most quintessential examples hereof is the CpG island methylator phenotype (CIMP), described as the

Box 3. Several statistical metrics can be used to evaluate biomarker performance [113, 116].

The performance of a biomarker can be evaluated by analyzing the test outcome in a population containing both individuals with and without the tested condition (e.g. the presence of cancer). As such, four groups of patients can be distinguished: cancer patients with a positive test (*true positives*; TP), cancer patients with a negative test (*false negatives*; FN), cancer-free patients with a positive test (*false positives*; FP) and cancer-free patients with a negative test (*true negatives*; TN). Using these four population parameters, following statistical metrics can be calculated (for calculations see Table 2):

- * The *specificity* or true negative rate is the percentage of cancer-free patients with a negative test.
- * The *sensitivity* or true positive rate or recall is the percentage of cancer patients with a positive test.
- * The *negative predictive value* is the percentage of patients with a negative test result that is truly cancer-free.
- * The *positive predictive value* or precision is the percentage of patients with a positive test result that truly has cancer.

Other commonly used statistical tools are the accuracy or balanced accuracy (BAC; for imbalanced population cohorts) and the area under the Receiver Operating Characteristic (ROC) curve. This curve can be plotted as the sensitivity against 1-specificity, in which each point in the curve represents the fraction of cancer patients with a biomarker measurement above a specific threshold versus the corresponding fraction of cancer-free patients above the same threshold. The area under the ROC curve then represents the probability that a randomly chosen cancer patient is ranked as more likely to have cancer than a randomly chosen cancer-free patient. This value is a useful way to describe the performance of a biomarker with a continuous output variable, regardless of the threshold level.

Table 2. Biomarkers can be evaluated by calculating several statistical metrics.

		condition		statistical metrics
		positive	negative	
test outcome	positive	true positive TP	false positive FP	positive predictive value PPV = $TP/(TP+FP)$
	negative	false negative FN	true negative TN	negative predictive value NPV = $TN/(TN+FN)$
statistical metrics	statistical metrics	true positive rate TPR = $TP/(TP+FN)$	true negative rate TNR = $TN/(FP+TN)$	accuracy ACC = $(TP+TN)/(TP+FP+FN+TN)$
		false negative rate FNR = $FN/(TP+FN)$	false positive rate FPR = $FP/(FP+TN)$	

Note. Specificity is TNR; sensitivity or recall is TPR; precision is PPV. Balanced accuracy (BAC) is the arithmetic mean of sensitivity and specificity.

aberrant and concordant hypermethylation of multiple promoter CpG islands, which can distinguish subgroups of colorectal cancer with different clinical, pathological and biological characteristics [113, 117]. DNA methylation biomarkers can also help in predicting and monitoring a patient's response to treatment, as illustrated by the DNA methylation markers *ESR1* and *ARH1* which are predictive

markers of response to tamoxifen treatment in breast cancer patients [113, 118]. Most of the studies so far mainly focused on promoter-associated CpG islands of single genes. However, it should be noted that apart from CIMP, also other panels of DNA methylation biomarkers have been described. In esophageal adenocarcinoma, methylation of four or more genes among *APC*, *DAPK*, *CDH1*, *ESR1*, *MGMT*, *CDKN2A* and *TIMP3* was associated with higher risk for early tumor recurrence and poor survival when taken together as a panel, while none of the single genes was prognostic when used as a single marker [113, 119]. Additional examples of clinically implemented DNA methylation biomarkers are given in Figure 8.

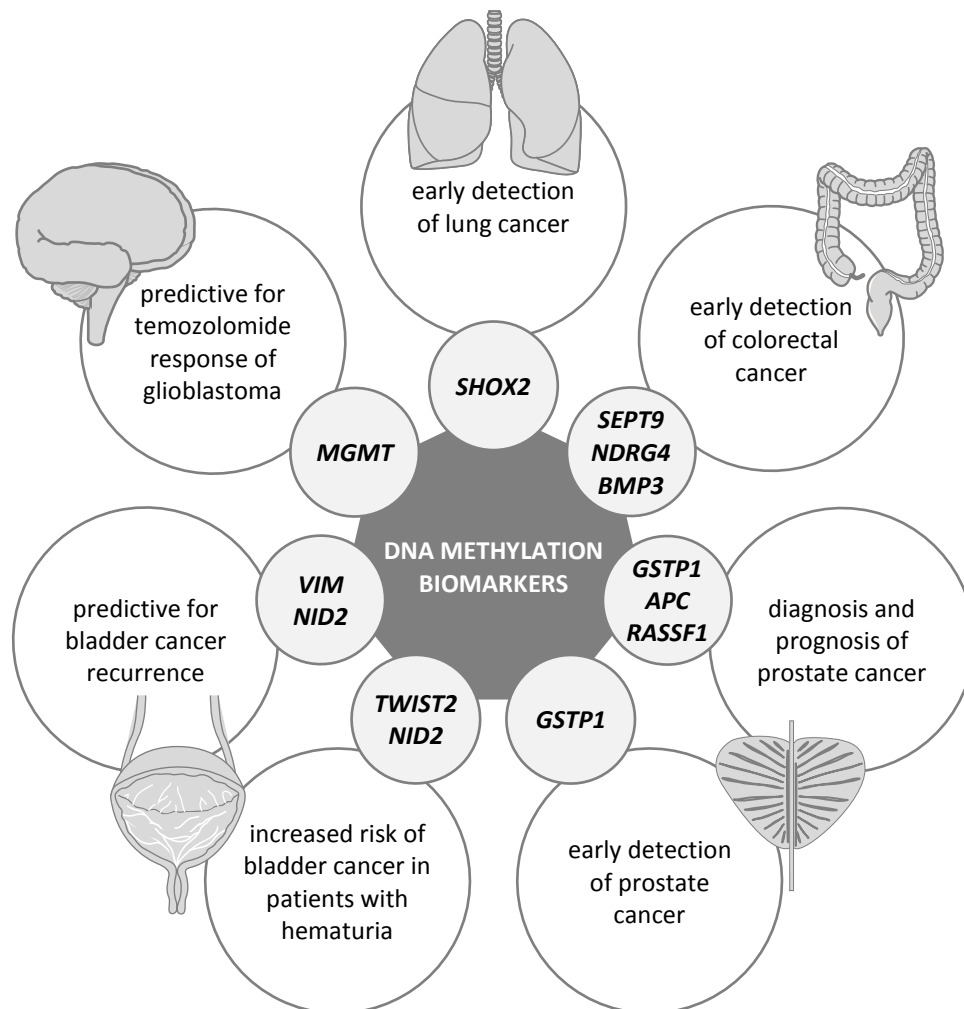


Figure 8. DNA methylation biomarkers are used in many clinical applications. For the depicted markers, commercially available test have been developed and clinically implemented. Based on [120].

1.1.6 DNA methylation as therapeutic target

Considering the crucial role of epigenetic alterations in cancer pathogenesis, there has been a growing interest in the utility of these changes in the development of strategies for cancer treatment. A plethora of epigenetic modulators of both DNA methylation and histone marks has been investigated and clinically implemented. The most important DNA methylation modulators are described in detail in the following sections. An overview of histone modulators is given in Box 4.

Box 4. Several classes of histone modulators have been developed and clinically implemented [121, 122].

An overview of the most important classes, subclasses and examples of histone modulators is given below. For detailed descriptions of their mechanisms of action and clinical trials using these drugs, the reader is referred to Nebbioso et al. and Juo et al. [121, 122].

- * histone deacetylase (HDAC) inhibitors
 - *short-chain and aromatic fatty acids*
valproic acid, phenylbutyrate, pivaloyloxymethyl butyrate
 - *hydroxamic acids and derivatives*
trichostatin A, pyroxamide, vorinostat, panobinostat, CHR-3996, tefinostat, pracinostat, givinostat, belinostat, JHJ-26481585, dacinostat
 - *benzamides*
entinostat, mocetinostat, chidamide
 - *cyclic peptides*
trapoxin A and B, romidepsin, apicidin
 - *sirtuin inhibitors*
splitomicin, tenovins, AGK2, sirtinol, suramin, EX-257, salermide, UVI5008
- * HDAC activators
 - *sirtuin activators*
resveratrol
- * histone acetyl transferase (HAT) inhibitors
 - *p300/CBP inhibitors*
anacardic acid, garcinol, curcumin
- * HAT activators
 - *p300 activators*
N-(4- chloro-3-trifluoromethyl-phenyl)-2-ethoxy-6-pentadecyl-benzamide
- * histone methyltransferase (HMT) inhibitors
 - *lysine methyltransferase (KMT) G9a inhibitors*
chaetocin, BIX-01294, UNC0224
 - *EZH2 inhibitors*
3-deazaneplanocin A
 - *protein arginine methyltransferase (PRMT) inhibitors*
AMI-1
- * histone demethylase (HDM) inhibitors
 - *LSD1 inhibitors*
tranilcypromine

1.1.6.1 Nucleoside analogues

The most archetypal examples of epigenetic cancer drugs are azacitidine (Vidaza) and decitabine (Dacogen). After their cellular uptake, these drugs are metabolized to 5-aza-2'-deoxycytidine-triphosphate (azacytosine), which is then incorporated into replicating DNA, where it substitutes for the naturally occurring cytosine. Azacytosine-guanine dinucleotides are recognized by the DNMTs as natural substrate and the enzymes will initiate the methylation reaction by a nucleophilic attack. This results in the establishment of a covalent bond between the C6 of the cytosine pyrimidine ring and the enzyme. Normally, this bond is then resolved by β -elimination through the C5 atom of cytosine (§1.1.4), but with azacytosine, where the C5 is substituted by N, this reaction is blocked and the DNMT is trapped. As a consequence, active DNMTs are depleted after several cell cycles, leading to loss of DNA methylation [123]. As azacitidine also gets incorporated into RNA, it also causes ribosomal disassembly, defective tRNA function and inhibited protein synthesis. Azacitidine exhibits greater cytotoxicity during S-phase, supporting the greater importance of its DNA effects [121]. The

Food and Drug Administration (FDA) approved azacitidine and decitabine for the treatment of the leukemia predisposition disorder myelodysplastic syndrome [123]. Importantly, these drugs need to be administered at nanomolar concentrations, as they become cytotoxic at higher doses. More recently, low doses of azacitidine/decitabine have also been tested in the management of solid tumors, both as a single drug treatment and in combination therapy, especially with HDAC inhibitors. In adults with metastatic non-small cell lung cancer, tumor responses improved gradually and progressively over several months of treatment with low doses of azacitidine in combination with the HDAC inhibitor entinostat, and this effect was sustained even after cessation of epigenetic therapy [124]. Azacitidine is also used to prime solid tumors for response to subsequent cytotoxic therapy [124, 125]. For example, it was shown that azacitidine partially reverses platinum resistance in patients with ovarian cancer [125].

Another nucleoside analogue is zebularine, which also inhibits cytidine deaminase, the enzyme responsible for inactivation of azacitidine and decitabine. Zebularine is more stable and less toxic compared to azacitidine and decitabine. This allows oral administration of the drug and simplifies continuous low-dose therapy. Importantly, this drug also shows a higher selectivity for tumor cells than azacitidine and decitabine [126]. Also of note is SGI-110, a derivative of decitabine with improved pharmacokinetics and metabolic stability, which results from a decreased degradation by cytidine deaminase [122].

1.1.6.2 Small molecule inhibitors

In efforts to circumvent the toxicity and instability of nucleoside DNMT inhibitors, non-nucleoside small molecule inhibitors of DNMTs have been developed, such as hydralazine, procainamide, RG108, SGI-1027 and MG98.

Hydralazine and procainamide are FDA approved drugs for the treatment of hypertension and cardiac arrhythmia, respectively. Procainamide specifically inhibits the maintenance methyltransferase activity of DNMT1, mainly by reducing the affinity of the enzyme for both DNA and SAM [127]. The demethylating action of hydralazine is still under investigation, but some evidence indicates that it binds to CpG-rich sequences and interferes with translocation of DNMTs along the DNA strand [121]. Although both compounds are considered weak DNA methylation inhibitors, a clinical trial on patients carrying solid tumors refractory to conventional therapy has demonstrated that hydralazine in combination with the HDAC inhibitor valproate causes DNA demethylation and overcomes chemotherapy resistance [127, 128]. Further pharmacological exploitation of these drugs might lead to the development of more potent DNA methylation inhibitors, as currently high levels of these drugs are needed to induce DNA demethylation [129].

Small molecule DNMT inhibitors that block the active site of DNMTs also have been identified. Based on a three-dimensional model of the human DNMT1 catalytic domain, a small-molecule database was screened *in silico* in order to find candidate DNMT inhibitors. This led to the identification of 2-(1,3-dioxo-1,3-dihydro-2*H*-isoindol-2-yl)-3-(1*H*-indol-3-yl)propanoic acid, later on renamed to RG108, which was predicted to strongly interact with the DNMT1 active site. *In vitro* work on HCT116 colon cancer cells further demonstrated that this compound efficiently blocked DNA methylation without the need to be incorporated into DNA and that the central carboxyl group of RG108 is responsible for

the specificity in the interaction between the inhibitor and the DNMT1 active site. These features make RG108 a promising DNMT inhibitor [130]. Another promising class of demethylating agents that inhibit DNMT activity represents synthetic small molecule inhibitors based on quinoline, a heterocyclic aromatic compound. Datta et al. have shown that one of these compounds, designated SGI-1027, effectively reactivates silenced tumor suppressor genes in colon cancer cells by demethylation of their respective CpG islands. SGI-1027 treatment also resulted in selective degradation of DNMT1 and exhibited minimal or no cytotoxic effect in rat hepatoma cells [131].

Also antisense oligonucleotides (ASOs), synthetic nucleic acids that inhibit translation of specific messenger RNAs (mRNAs), or degrade them, by binding to a target region within the mRNA, are used to inhibit DNA methylation. The ASO MG98 specifically inhibits human DNMT1 mRNA, resulting in reduced DNMT1 protein levels and re-expression of silenced tumor suppressor genes. Clinical trials on patients with advanced solid tumors have demonstrated that MG98 treatment has antitumor activity and is generally well tolerated [132, 133].

1.1.6.3 Natural compounds

The DNMT inhibitory potency of naturally occurring molecules, such as psammaplin A and (–)-Epigallocatechin-3-gallate (EGCG), have also been topic of investigation.

Literature on the DNMT inhibitory effect of psammaplin A is controversial. Early studies indicated that this natural product isolated from marine sponges was an extremely potent inhibitor of both HDAC and DNMT, and showed cytotoxicity against human lung, ovarian, skin and colon cancer cell lines. Although more recent studies confirmed the antitumor activity of this compound, evidence of DNMT activity inhibition could not be demonstrated in these studies [134].

EGCG is a green tea polyphenol that has been shown to have antitumorigenic properties. Nandakumar et al. demonstrated that treatment of skin cancer cells with EGCG reduced the levels of DNA methylation and DNMT activity, resulting in re-expression of tumor suppressor genes. It also inhibited HDAC activity and increased levels of acetylated histones [135]. Additionally, also EGCG analogues have been developed. These molecules are currently being tested for different applications in cancer management [136].

1.2 The quest for prognostic neuroblastoma markers

In 1864, in a study on hyperplasia of the pineal and adrenal glands, Virchow described for the first time a child with an abdominal tumor which he denominated as glioma. Further examination of Marchand in 1891 showed histological similarities between tumors of the adrenal gland and the sympathetic nervous system [137]. Later on, in 1901, Pepper described a group of enigmatic adrenal sarcomas in infants associated with massive liver dissemination, of which the clinical and pathologic features closely resembled one another and a few years later, in 1907, Hutchison similarly reported on a separate group of adrenal sarcomas in older children with orbital and skeletal metastases [138, 139]. In 1910, Wright pointed out that the adrenal tumors described by Pepper and Hutchison were mainly composed of an identical cell type, regardless their distinct patterns of spread. These cells had the same morphology as the cells from which the sympathetic nervous system and the medulla of the adrenal developed, and were regarded by embryologists as arising from migrated primitive nerve cells. Owing to their characteristic primitive neural cell origin, Wright collectively named these types of tumors 'neuroblastoma' [140, 141].

Today, neuroblastoma (NB) is appointed as a disease of the sympathicoadrenal lineage of the neural crest, a transient population of cells during embryonal development that arises at the border between the neural plate and the non-neural ectoderm (Figure 9 and Table 3). Precursors with the potential to form neural crest cells initially are contained within the dorsal portion of the neural tube. Subsequently, these premigratory neural crest cells emerge from the neural tube and start migrating along characteristic pathways to give rise to diverse and numerous derivatives. These migratory pathways are regionalized according to the original position of the neural crest cells along the neural axis, such that cells from a given axial level give rise to a characteristic array of progeny and follow distinct pathways from those arising at other axial levels (Table 3). For example, ventrally migrating trunk neural crest cells give rise to the peripheral nervous system (PNS) of the trunk, including the chain of the sympathetic ganglia and dorsal root ganglia, as well as chromaffin cells of the adrenal medulla. In addition to these neurons, these neural crest cells also generate glia of the peripheral ganglia, non-neural cells supporting and protecting neurons, including Schwann cells that ensheath and myelinate peripheral axons. Cessation of neural crest cell migration and subsequent neurogenesis in the developing peripheral nervous system (PNS) is not yet completely understood, but local environmental signals seem to critically control neural and glial cell fate. For example, bone morphogenetic proteins (BMPs) are essential for the development of the sympathetic nervous system. An interesting feature of the developing PNS is that migrating neural crest cells proliferate rapidly as they move. Even after exhibiting defined neural characteristics, some neural crest derivatives continue to divide. For example, in the developing sympathetic ganglia, neural-crest derived cells express neurotransmitters and other proteins characteristic of sympathetic neurons, but remain actively mitotic [142].

Clearly, the processes involved in the development of the sympathetic nervous system are very complex and alterations herein can initiate disease pathogenesis. Although the molecular aberrations that cause NB are not yet fully unraveled, it has been shown that these tumors develop from immature sympathetic nervous system cells, called neuroblasts. As such, NB can develop anywhere in this system, which contributes to the most important hallmark of this type of tumor: heterogeneity. The early reports of NB by Pepper and Hutchison, which today would have been

staged as MS and M respectively, perfectly illustrate this clinical heterogeneity. Stage MS and M tumors both represent metastatic disease, but where patients with stage M often show relentless progression, stage MS tumors usually regress spontaneously [143]. These extreme differences in disease course have triggered a quest for markers that allow prognosis prediction at the time of diagnosis, a still persisting effort with the ultimate goal to fine-tune the treatment protocol for each patient.

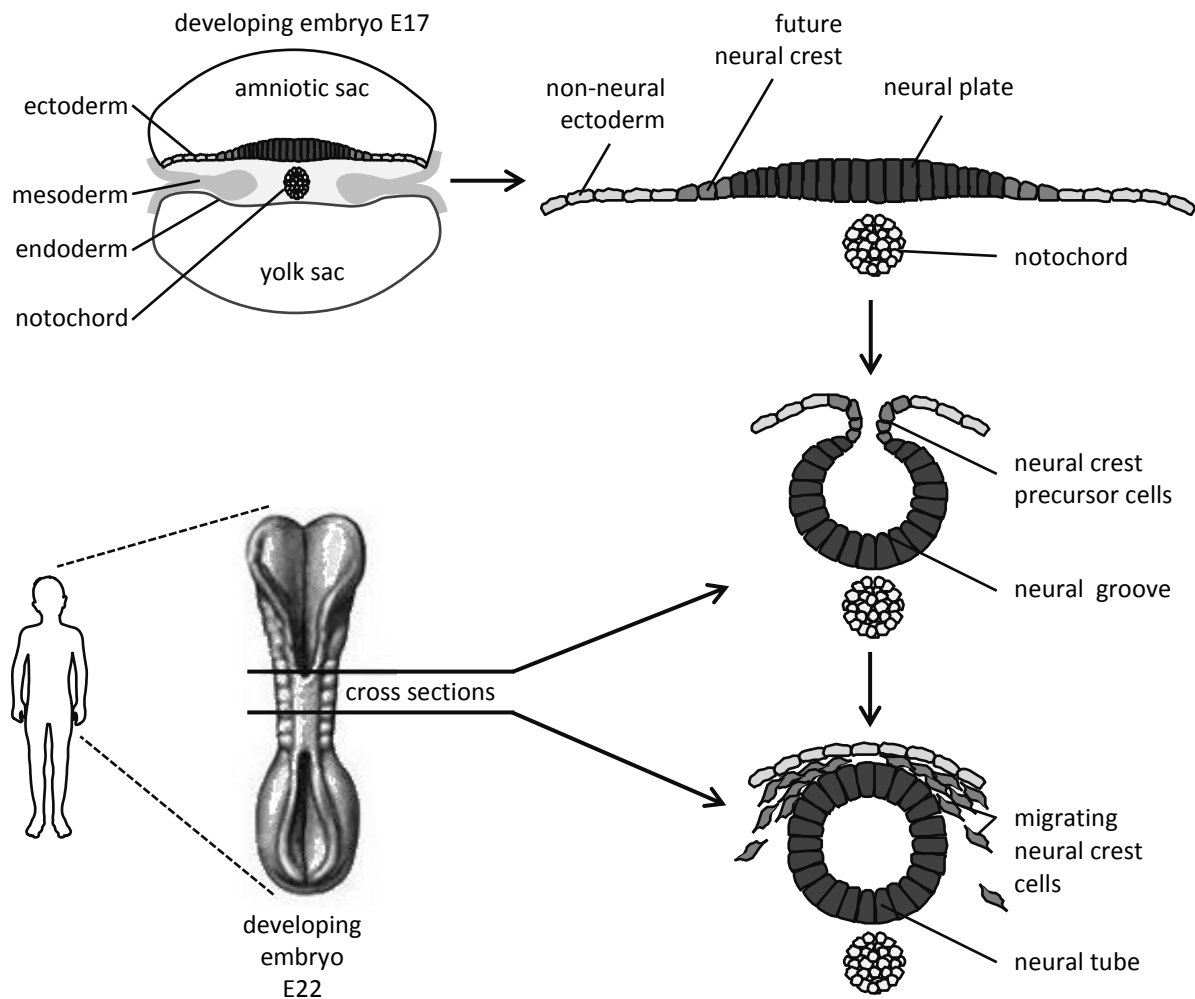


Figure 9. The neural crest is a transient population of cells during embryonal development. At embryonic day 17, the developing embryo consists of three primary germ layers (ectoderm, mesoderm and endoderm; top left). The notochord defines the body axis, i.e. the midline that divides the left and right side of the body, and is the site of the future vertebral column. The amniotic sac is filled with fluid and buffers the developing embryo from physical shock until the time of birth. The yolk sac will form part of the future digestive system. As the notochord develops, it signals the overlying ectoderm to start forming the spinal cord and brain, an event called neurulation (close-up views of the ectoderm at the right of the figure). Specifically, the ectoderm in the dorsal midline thickens into a neural plate, and then starts to fold inwards as a neural groove. This groove deepens until a hollow neural tube is pinched off into the body (at embryonic day 22). Complete closure of the neural tube occurs by the end of week 4. Neural crest cells originate from ectodermal cells on the lateral ridges of the neural plate, and are pulled into the body along with the invaginating neural tube. According to the original position of the neural crest cells along the neural axis, they will give rise to specific body structures (Table 3). Based on [144].

Table 3. Different populations of neural crest cells arising along the neural axis give rise to distinct cell types.

neural crest cell population	derivatives
cranial	cranial sensory ganglia parasympathetic ciliary ganglia cartilaginous elements of the facial skeleton
vagal	enteric nervous system
trunk	melanocytes sympathetic ganglia dorsal root ganglia chromaffin cells of the adrenal medulla glia Schwann cells
lumbosacral	enteric nervous system

Note. Neuroblastoma is appointed as a disease of the sympathoadrenal lineage of the neural crest, a specific subset of trunk neural crest cells. Based on [142].

The prognostic value of a great diversity of parameters has already been investigated, including the primary tumor site and volume [145, 146], the metastatic site [147], the ratio of the amount of the catecholamine metabolite markers vanillylmandelic acid (VMA) and homovanillic acid (HVA) in urine [148, 149], the serum neuron-specific enolase (NSE) [150], ferritin [149, 151], lactate dehydrogenase (LDH) [149, 151], chromogranin A [152] and neuropeptide Y level [153], and many others. However, the best studied and widely used prognostic factors are the age of the patient at diagnosis, tumor histology and stage, *MYCN* oncogene amplification, DNA index and specific chromosome gains and losses. Additionally, during the last decades, the development of high-resolution and high-throughput (epi)genome and transcriptome profiling methods has provoked the advent of new molecular prognostic markers, such as mRNA and microRNA (miRNA) signatures, as well as DNA methylation biomarkers. Widely used prognostic parameters have been combined into a pretreatment risk classification system [154].

1.2.1 Widely used prognostic markers

1.2.1.1 Age of the patient at diagnosis

One of the most remarkable characteristics of NB is its age-specific incidence. According to the Surveillance, Epidemiology, and End Results (SEER) Program 1975-2012 of the National Cancer Institute, the incidence rate of NB is 10.1 per million children under 15 years of age, but this rate increases to 20.5 per million in the 1-4 years of age group and even to 49.2 per million during infancy (grouped category data of NB and ganglioneuroblastoma (GNB), a more differentiated variant of NB (§1.2.1.2)) [155]. Although NB is one of the most common cancers of infancy, it is also diagnosed in adolescents and adults. However, these cover less than 10% of all cases, as 90% is diagnosed by the age of 6 years (median age of diagnosis is approximately 20 months) [156].

It has long been observed that NB prognosis varies markedly with the age of the patient at diagnosis. More than 40 years ago, Breslow and McCann [157] found that survival probabilities (Box 5) decreased with increasing age of the patient during the first two years and that the age effect tapered off thereafter. These results suggest that days of age should be used for risk stratification as a continuous variable, but as this is not clinically practical to tailor therapy, a convenient age cutoff of

12 months was applied. Hereby, patients older than 12 months receive more therapy than younger patients [158]. In 2003, Schmidt et al. [159] showed that the survival rates of some patients slightly older than 12 months (between 12 and 18 months) were significantly better than those of older patients, although they received the same therapy. Two years later, in order to maximally lower the burden of therapy (and accompanying late-term side effects) for these patients, London et al. retrospectively analyzed the influence of age on outcome in 3,666 patients to identify a better, statistically optimal age cutoff. Their findings indicate that the optimal age cutoff was between 15 and 19 months [158].

Box 5. Survival probabilities can be estimated using the Kaplan-Meier method [160].

Kaplan-Meier curves and estimates of survival data are frequently used to analyze survival, as this method can be performed on patient data with differing survival times, i.e. times-to-event, as well as censored data. The time-to-event variable can be described as the clinical-course duration or follow-up time of a patient and the event may be any event of interest, for example death (when studying overall survival (OS)) or progression/relapse (when studying event-free survival (EFS)) of a specific disease. Survival studies are often performed to determine whether different patient groups, stratified according to a nominal variable of interest (for example age group of the patient), show statistically different survival rates. In such studies, each patient is characterized by three parameters: its clinical-course duration (time-to-event), its event status at the end of its clinical-course duration (i.e. event occurred or patient is censored), and its study group. Using these parameters, Kaplan-Meier curves can be constructed, as exemplified in Figure 10. Kaplan-Meier curves consist of horizontal lines, representing clinical-course duration intervals, and vertical lines, showing changes in the cumulative survival probability as the curve advances. Each event in the data set ends an interval and begins another interval. Censored patients do not demarcate intervals and their times-to-event are indicated with ticks. Importantly, these patients strongly effect survival rates and for this reason, it is recommended to indicate patient numbers at risk below the curves. The cumulative survival probability for an interval defines the probability of survival at the beginning and throughout the interval, and is calculated by multiplying the interval survival rates up to that interval. Kaplan-Meier curves of multiple study groups can be statistically compared by the log-rank test or by calculating Cox proportional hazards. The log-rank test calculates the chi-square for each event time for each group and sums the results. The summed results for each group are added to derive the ultimate chi-square to compare the full curves of each group. Cox proportional hazards show the increased rate of having an event in one curve versus the other [160].

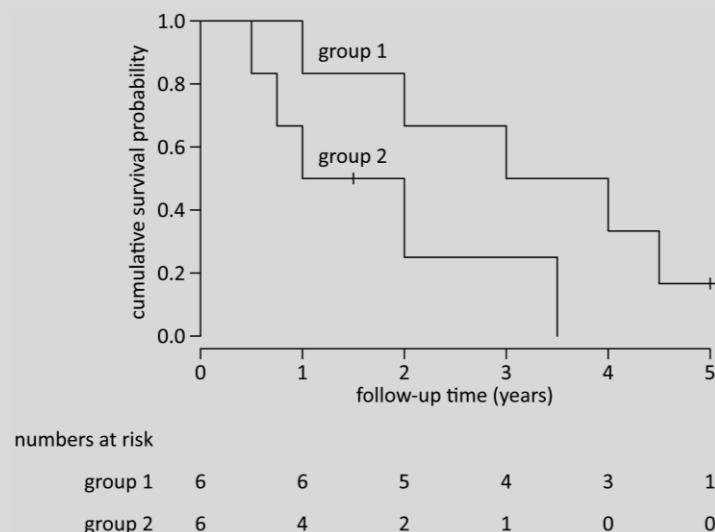


Figure 10. Kaplan-Meier curves are used to analyze survival data. Horizontal lines represent clinical-course duration intervals. Vertical lines represent changes in the cumulative survival probability. Censored patients are indicated with ticks and patient numbers at risk are shown below the Kaplan-Meier curves. Based on [160].

1.2.1.2 Tumor histology and differentiation grade

Tumors that belong to the heterogeneous group of neuroblastic tumors, such as NB, are typified by different morphologic characteristics. During the past, these differences in histological appearance have led to the development of several histopathologic grading systems that showed to be of prognostic importance [161]. The initial grading systems, proposed by Beckwith and Martin [162], Mäkinen [163], and Hughes [164], were based on the degree of differentiation of the neuroblasts into ganglion cells and Schwann cells (§1.2) by evaluating histologic signs of maturation. It was generally acknowledged that differentiated tumors had a more favorable prognosis than undifferentiated tumors. Shimada et al. took a new approach with their age-linked classification and divided neuroblastic tumors into Schwannian stroma-rich and -poor tumors. Further subdivisions depended on the grade of differentiation and nuclear morphology of the neuroblasts. The latter was quantified in the mitosis-karyorrhexis index (MKI), which reflected the percentage of neuroblasts in mitosis (demonstrating cell proliferation) and in the process of karyorrhexis (i.e. nuclear fragmentation, demonstrating cell death) [161, 165, 166]. Also Joshi et al. proposed an age-linked grading system based on tumor calcification and the mitotic rate, later on replaced by the MKI [167, 168]. In 1994, the International Neuroblastoma Pathology Committee (INPC) was formed, which aimed at testing the prognostic significance of the different morphologic features and their combination. Based on detailed definitions of these features and statistical analyses of pathology review data, the International Neuroblastoma Pathology Classification was proposed, which largely adopted the Shimada classification (Table 4). In this classification, NB tumors are defined as neuroblastic Schwannian stroma-poor tumors (i.e. the proportion of tumor tissue with stroma-rich histology does not exceed 50%) and are further categorized in undifferentiated, poorly differentiated and differentiating subtypes. In the undifferentiated subtype (Figure 11A), the tumor tissue is composed of undifferentiated neuroblasts without identifiable neuropil or Homer Wright rosettes (Box 6). In order to establish the diagnosis, supplementary tests such as immunohistochemistry, electron microscopy, and/or molecular/cytogenetic analyses are usually required. Diagnosis of the poorly differentiated subtype (Figure 11B) is relatively easy, because of the presence of varying amount of neuropil and/or rosettes. Most of the tumor cells in this subtype are undifferentiated: less than 5% of the population has morphological evidence of differentiation. Tumors of the differentiating subtype (Figure 11C) usually have abundant neuropil. At least 5% of the tumor cells are differentiating neuroblasts, which are characterized by synchronous differentiation of the nucleus (enlarged, eccentrically located with vesicular chromatin pattern, and a single prominent nucleolus) and of the cytoplasm (eosinophilic/amphophilic (having affinity for both acid and basic dyes) with a diameter two or more times larger than the nucleus) [169].

1.2.1.3 Tumor stage

The practice of dividing NB tumors into stages arose from the fact that patient survival rates differed between patients with localized disease and patients in which the disease had extended beyond the site of origin. As a result, multiple NB tumor staging systems have been proposed in the past to aid in estimating prognosis [170, 171]. These include that of James [172], Pinkel [173, 174], Cohen [175], and Thurman and Donaldson [176], which are based on the extent of the disease, the surgical resectability, the pattern of metastatic spread and, in some cases, the degree of histologic differentiation. The first internationally accepted staging schema though, was that of Evans (and the

Table 4. The International Neuroblastoma Pathology Classification (Shimada System) stratifies neuroblastomas into favorable and unfavorable tumors, using the age of the patient and tumor differentiation grade.

International Neuroblastoma Pathology Classification				Shimada Classification
NB (Schwannian stroma-poor)	favorable	< 1.5 years	poorly differentiated or differentiating and low or intermediate MKI tumor	stroma-poor (favorable)
		1.5-5 years	differentiating and low MKI tumor	
	unfavorable	< 1.5 years	undifferentiated tumor high MKI tumor	stroma-poor (unfavorable)
		1.5-5 years	un- or poorly differentiated tumor intermediate or high MKI tumor	
		≥ 5 years	all tumors	
GNB, intermixed (Schwannian stroma-rich)				stroma-rich, intermixed (favorable)
GN (Schwannian stroma-dominant)	maturing			well differentiated (favorable)
	mature			GN (favorable)
GNB, nodular (composite Schwannian stroma-rich/stroma-dominant and stroma-poor)				stroma-rich, nodular (unfavorable)

Note. GN: ganglioneuroma; GNB: ganglioneuroblastoma; MKI: mitosis-karyorrhexis index; NB: neuroblastoma [161].

Box 6. The presence or absence of Homer Wright rosettes is used to determine the neuroblastoma differentiation grade.

Rosettes consist of a halo or spoke-wheel arrangement of cells surrounding a central lumen or hub (Figure 11B). Several types of rosettes are described in pathology literature. Typically, in NB the halo is formed by neuroblasts which enclose a central lumen or hub with a fibrillary collection of primitive neurites (neuropil). Although this type of rosettes was first described by Wright in NB, it also emerges in medulloblastoma, primitive neuroectodermal tumor (PNET) and pineoblastoma. The mechanism for the formation of the characteristic rosette pattern is not completely understood. The cell populations exhibiting neural differentiation are believed to secrete glycoproteins and glycolipids, which mediate cell-to-cell recognition and adhesion. One hypothesis is that these sticky cell surface markers cause the developing cell bodies to cluster or aggregate and their primitive neurites to tangle. As the cells grow, the neurite tangle remains centrally located and the cell bodies are squeezed to the periphery, explaining the rosette pattern [177].

Children's Cancer Study Group (CCSG) [178, 179]. This staging system was based on lymph node involvement and used the anatomic midline of the body as a reference to describe the extent of the disease. Evans et al. also devised a special stage IV (IV-S) category, as it was recognized that a certain constellation of ominous findings did not have the grave prognosis usually predicted for other cancers under such circumstances (stage IV). More precisely, it was restricted to patients who would otherwise have been stage I or II (with localized primary tumors), but who had remote disease confined to liver, skin or bone marrow, and any combination of these, without involvement of the skeleton. Over the following years, two other major staging systems evolved in the management of NB: the tumor-node-metastases (TNM) system from the International Union against Cancer (UICC) and the American Joint Committee on Cancer (AJCC) [180, 181], and the system used by the St Jude Children's Research Hospital (SJCRH) and the Pediatric Oncology Group (POG) [182, 183].

Additionally, modifications of these systems have been proposed by the Italian Cooperative Working Group [184] and the Malignant Tumor Committee of the Japanese Society of Pediatric Surgeons [171, 185].

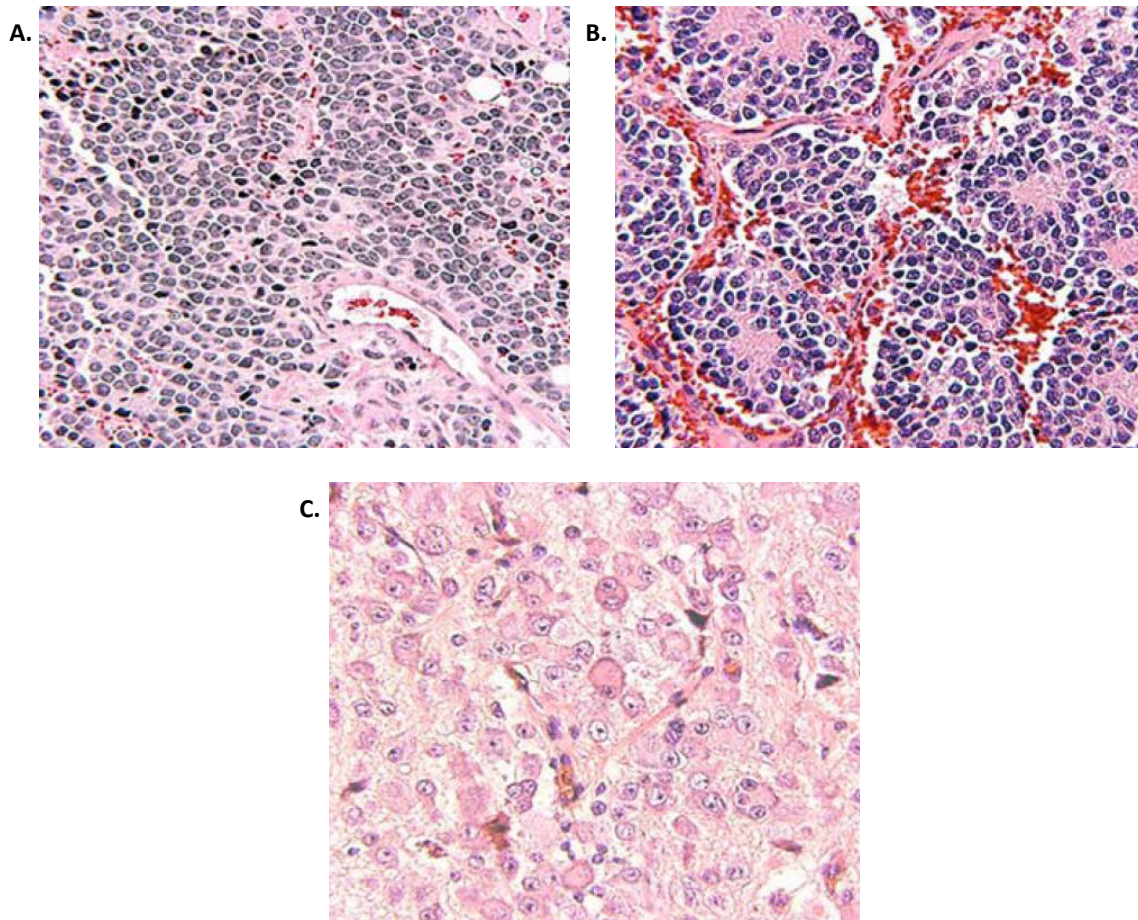


Figure 11. Neuroblastoma tumors are categorized into undifferentiated, poorly differentiated and differentiating subtypes. A. Undifferentiated subtype. B. Poorly differentiated subtype. C. Differentiating subtype. Adapted from [169].

Although each of these staging systems has its strengths, the plethora of diversiform systems made it difficult to compare the results of clinical trials and biological studies from different groups and countries. To aid in the development of a consensus system, a meeting was held in 1986 to address the problem of standardizing definitions for staging in NB. The resulting International Neuroblastoma Staging System (INSS) was published in 1988 [185] and revised (Table 5) in 1993 [186]. However, a major drawback of this system is that staging is based on the degree of surgical resection of the tumor and lymph node involvement, criteria that highly depend on the thoroughness and expertise of the physician. In 2009, the International Neuroblastoma Risk Group (INRG) Task Force [154] therefore proposed a new system, the INRG Staging System (INRGSS; Table 6), based on radiological risk factors for surgery, i.e. imaging characteristics (determined via computed tomography (CT), magnetic resonance imaging (MRI), iodine-123 meta-iodobenzylguanidine (^{123}I -mIBG) scintigraphy and/or technetium-99m-methylene diphosphonate ($^{99\text{m}}\text{Tc}$ -MDP) bone scintigraphy (Box 9)) which are associated with an increased risk of surgical complications. Staging based on these so-called image-defined risk factors (IDRFs) is not only more robust and reproducible than one based on surgical findings, but also allows pretreatment risk stratification [143].

Table 5. The International Neuroblastoma Staging System (INSS) stratifies neuroblastoma tumors into several tumor stages, based on the degree of surgical tumor resection and lymph node involvement.

stage	description
1	localized tumor with complete gross excision, with or without microscopic residual disease; representative ipsilateral lymph nodes negative for tumor microscopically (nodes attached to and removed with the primary tumor may be positive)
2A	localized tumor with incomplete gross excision; representative ipsilateral non-adherent lymph nodes negative for tumor microscopically
2B	localized tumor with or without complete gross excision, with ipsilateral non-adherent lymph nodes positive for tumor; enlarged contralateral lymph nodes must be negative microscopically
3	unresectable unilateral tumor infiltrating across the midline*, with or without regional lymph node involvement; or localized unilateral tumor with contralateral regional lymph node involvement; or midline tumor with bilateral extension by infiltration (unresectable) or by lymph node involvement
4	any primary tumor with dissemination to distant lymph nodes, bone, bone marrow, liver, skin and/or other organs (except as defined for stage 4S)
4S	localized primary tumor (as defined for stage 1, 2A or 2B), with dissemination limited to skin, liver, and/or bone marrow** (limited to infants < 12 months of age)

Note. Multifocal primary tumors (for example bilateral adrenal primary tumors) should be staged according to the greatest extent of disease, as defined above, followed by subscript letter M. *The midline is defined as the vertebral column. Tumors originating on one side and crossing the midline must infiltrate to or beyond the opposite side of the vertebral column. **Marrow involvement in stage 4S should be minimal, that is, < 10% of total nucleated cells identified as malignant on bone marrow biopsy or on marrow aspirate. More extensive marrow involvement would be considered to be stage 4. The MIBG scan (if performed) should be negative in the marrow [186].

Table 6. The International Neuroblastoma Risk Group Staging System (INRGSS) stratifies neuroblastoma tumors into several tumor stages, based on image-defined risk factors.

stage	description
L1	localized tumor not involving vital structures as defined by the list of IDRFs and confined to one body compartment
L2	locoregional tumor with presence of one or more IDRFs
M	distant metastatic disease (except stage MS)
MS	metastatic disease in children < 18 months of age, with metastases confined to skin, liver, and/or bone marrow

Note. Detailed criteria and list of IDRFs in Box 7 and Box 8, respectively. Patients with multifocal primary tumors should be staged according to the greatest extent of disease as defined in the table. IDRFs: image-defined risk factors [143].

Box 7. The International Neuroblastoma Risk Group Staging System (INRGSS) stratifies neuroblastoma tumors into several tumor stages, based on image-defined risk factors [143].

Stage L1 tumors are localized tumors that do not involve vital structures as defined by the list of IDRFs. The tumor must be confined within one body compartment, neck, chest, abdomen, or pelvis. The isolated finding of intraspinal tumor extension that does not fulfill the criteria for an IDRF is consistent with stage L1.

Stage L2 tumors are locoregional tumors with one or more IDRFs. The tumor may be ipsilaterally continuous within body compartments (i.e., a left-sided abdominal tumor with left-sided chest involvement should be considered stage L2). However, a clearly left-sided abdominal tumor with right-sided chest (or vice versa) involvement is defined as metastatic disease.

Stage M is defined as distant metastatic disease (i.e., not contiguous with the primary tumor) except as defined for MS. Non-regional (distant) lymph node involvement is metastatic disease. However, an upper abdominal tumor with enlarged lower mediastinal nodes or a pelvic tumor with inguinal lymph node involvement is

considered locoregional disease. Ascites and a pleural effusion, even with malignant cells, do not constitute metastatic disease unless they are remote from the body compartment of the primary tumor.

Stage MS is metastatic disease in patient younger than 18 months (547 days) with metastases confined to skin, liver, and/or bone marrow. Bone marrow involvement should be limited to less than 10% of total nucleated cells on smears or biopsy. MIBG scintigraphy must be negative in bone and bone marrow. Provided there is MIBG uptake in the primary tumor, bone scans are not required. The primary tumor can be L1 or L2 and there is no restriction regarding crossing or infiltration of the midline.

Box 8. The International Neuroblastoma Risk Group Staging System (INRGSS) makes use of a list of image-defined risk factors (IDRFs) to determine the neuroblastoma tumor stage [143].

ipsilateral tumor extension within two body compartments

neck-chest, chest-abdomen, abdomen-pelvis

neck

tumor encasing carotid and/or vertebral artery and/or internal jugular vein
tumor extending to base of skull
tumor compressing the trachea

cervico-thoracic junction

tumor encasing brachial plexus roots
tumor encasing subclavian vessels and/or vertebral and/or carotid artery
tumor compressing the trachea

thorax

tumor encasing the aorta and/or major branches
tumor compressing the trachea and/or principal bronchi
lower mediastinal tumor, infiltrating the costo-vertebral junction between T9 and T12

thoraco-abdominal

tumor encasing the aorta and/or vena cava

abdomen/pelvis

tumor infiltrating the porta hepatis and/or the hepatoduodenal ligament
tumor encasing branches of the superior mesenteric artery at the mesenteric root
tumor encasing the origin of the coeliac axis, and/or of the superior mesenteric artery
tumor invading one or both renal pedicles
tumor encasing the aorta and/or vena cava
tumor encasing the iliac vessels
pelvic tumor crossing the sciatic notch

intraspinal tumor extension

whatever the location provided that more than one third of the spinal canal in the axial plane is invaded and/or the perimedullary leptomeningeal spaces are not visible and/or the spinal cord signal is abnormal

infiltration of adjacent organs/structures

pericardium, diaphragm, kidney, liver, duodeno-pancreatic block, and mesentery

conditions to be recorded, but not considered IDRFs

multifocal primary tumors
pleural effusion, with or without malignant cells
ascites, with or without malignant cells

Box 9. Radiological imaging is used to evaluate the presence or absence of image-defined risk factors.

Computed tomography (CT) and/or magnetic resonance imaging (MRI) with three-dimensional measurements and of sufficient quality to address image-defined risk factors (IDRFs) is mandatory for imaging the primary tumor. The presence or absence of each individual IDRF should be evaluated and recorded (Box 8). When possible, metastatic sites should also be measured by CT and/or MRI, as this information may be needed to evaluate treatment response [143].

The iodine-123 meta-iodobenzylguanidine (^{123}I -mIBG) scintigraphy imaging procedure is essential for the detection of metastases and is based on the intravenous administration of the radiopharmaceutical ^{123}I -mIBG, followed by recording the distribution of the radioactivity using a camera, here a device used to image gamma radiation emitting radioisotopes. Being an analogue of noradrenalin, ^{123}I -mIBG enters cells expressing the noradrenalin transporter and is stored in neurosecretory granules, which results in a concentration of radioactivity in these cells. Therefore, ^{123}I -mIBG scintigraphy is useful to image tumors of neuro-endocrine origin, particularly those of the neuro-ectodermal (sympatho-adrenal) system, such as NB [187]. Occasionally, false-positive readings may occur because of uptake in mature ganglioneuroma or other neuro-endocrine tumors, or because of physiological uptake that may be mistaken for tumor in the adrenal gland, salivary gland, nasopharynx, brown fat or excretion through renal pelvis and bladder. False-negative scans may be observed in approximately 10% of NB tumors that do not concentrate ^{123}I -mIBG, owing to low expression of the noradrenalin transporter or owing to blood-brain barrier or large areas of scar or necrosis. In addition, very small amounts of bone marrow tumor will often not be detected, and therefore ^{123}I -mIBG scintigraphy must be supplemented with bilateral bone marrow biopsy [188].

Technetium-99m-methylene diphosphonate ($^{99\text{m}}\text{Tc}$ -MDP) bone scintigraphy is used to trace metastatic NB in the skeleton. The radiopharmaceutical $^{99\text{m}}\text{Tc}$ -MDP is a marker of bone perfusion and turnover, as it is adsorbed to the crystalline structure of hydroxyapatite, the mineral component of bone built of crystals containing mainly calcium and phosphate. Therefore, visualization of tumor lesions in the bone using $^{99\text{m}}\text{Tc}$ -MDP and gamma radiation imaging is possible, as the presence of the tumor in the bone causes changes in the bone remodelling activity [189]. $^{99\text{m}}\text{Tc}$ -MDP bone scintigraphy is only required when ^{123}I -mIBG positivity of the primary tumor cannot be confirmed [143].

1.2.1.4 MYCN amplification

Double minutes (DMs) and homogeneously staining regions (HSRs) are karyotypic abnormalities frequently seen in tumor cells and point towards amplification of cellular genes (Box 10). In this way, an increase in gene dosage is caused and amplification of proto-oncogenes leads to activation of their oncogenic potential [190]. In 1983, upon the discovery of amplification of *MYC* (the human V-Myc Avian Myelocytomatosis Viral Oncogene Homolog) in colon carcinoma, Schwab et al. demonstrated that NB tumors showed cytogenetic evidence for amplification of a DNA domain that exhibited homology to *MYC* [191]. One year later, the locus of this gene, designated as the V-Myc Avian Myelocytomatosis Viral Oncogene Neuroblastoma Derived Homolog (*MYCN*), was mapped to the short arm of chromosome 2 (2p24). *MYCN* amplification in NB occurs in approximately 20-30% and copy numbers range from 10 to more than 500, but 50 to 100 copies are generally observed. Lower numbers are indicated as *MYCN* copy gain [190]. Normally, *MYCN* is only expressed during embryogenesis in pre-B cells, kidney, forebrain, hindbrain and intestine, with the highest expression in the developing brain. After embryonic development, the transcription factor is downregulated and in adult tissues it is no longer significantly expressed. The essential developmental role of *MYCN* is underscored by embryonic lethality of *MYCN*-null mice [192]. Additionally, it has been shown that enhanced expression of *MYCN* is a factor in tumorigenic cell conversion and that *MYCN* has multiple direct and indirect targets, including other important players in the NB pathogenesis such as *ALK* [193–195]. More recent mice experiments signify *MYCN* as an oncogenic driver gene in NB [196].

Box 10. Several technologies can be used to assess DNA gains and losses.

Frequently applied technologies in NB molecular diagnostics are interphase fluorescence *in situ* hybridization (I-FISH), PCR, array comparative genomic hybridization (aCGH) and multiplex ligation-dependent probe amplification (MLPA) [197].

The *MYCN* amplification status is frequently determined via I-FISH, using a fluorophore-labeled probe in 2p24, and presents as extrachromosomal double minutes (DMs) or intrachromosomal homogeneously staining regions (HSRs; Figure 12) [198]. Usually, two color I-FISH is used, so that the *MYCN* signal can be compared with that of a reference probe located on chromosome 2q [197]. Other detection technologies are qPCR [199], and comparative genomic hybridization (CGH). For CGH, test DNA and normal reference DNA are hybridized simultaneously to normal chromosome spreads. As the hybridization is detected with two different fluorochromes, regions of gain or loss of DNA sequences are seen as changes in the ratio of the intensities of the two fluorophores along the target chromosomes [200]. Later on, these chromosome spreads were replaced by arrayed oligonucleotides, allowing fast and high-resolution measurements of DNA copy number changes. Furthermore, a NB-specific PCR-based MLPA kit was developed to analyze the prognostic impact of certain genomic changes in NB [201, 202].

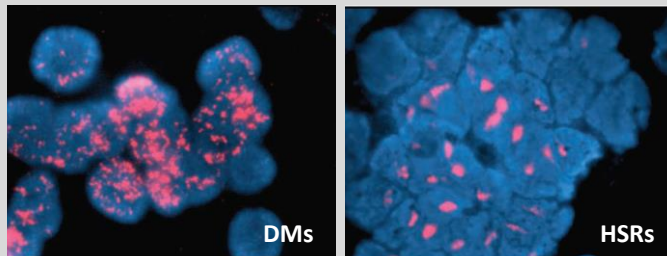


Figure 12. Fluorescence *in situ* hybridization of neuroblastoma cells depicts *MYCN* amplification as double minutes or intrachromosomal homogeneously staining regions. DMs: double minutes; HSRs: homogeneously staining regions. Adapted from [198].

Soon after its discovery, it was shown that *MYCN* amplification correlates with advanced disease stage and is associated with rapid progression and worse prognosis [203, 204]. Given its pronounced effect on survival rates, *MYCN* amplification became a key determinant in risk estimation and therapy stratification [154]. Also the mRNA and protein levels were tested on their prognostic significance, as it was noticed that tumors without *MYCN* amplification can still be characterized by high *MYCN* expression. Though, in literature controversy exists whether these factors have prognostic value or not [205–207]. Of note, in the absence of *MYCN* amplification, high-stage NB tumors frequently show high expression of MYC, and activated MYC signaling is associated with poor outcome [208, 209].

1.2.1.5 Chromosome aberrations

1.2.1.5.1 DNA hyperdiploidy

Next to amplification of *MYCN* (§1.2.1.4), NB tumors often show a hyperdiploid chromosomal constitution. This means that the tumor cells gained extra chromosomal copies resulting in a higher DNA content (triploid, tetraploid,...) compared to normal cells (diploid). The ploidy status is clearly associated with the age of the patient at diagnosis and *MYCN* copy number. By flow cytometric determination of the DNA index, George et al. showed that the frequency of hyperdiploid tumors was the highest in infants and that a diploid DNA content was associated with *MYCN* amplification in children younger than 24 months [210]. It has been shown that a diploid DNA index and *MYCN*

amplification confer a very high risk of treatment failure in infants with disseminated NB, while hyperdiploidy with non-amplified *MYCN* is associated with a very favorable outcome in infants younger than 18 months, even if the disease is widely disseminated at diagnosis. In older children, DNA ploidy measurements have no prognostic value [210].

1.2.1.5.2 Numerical and segmental chromosome aberrations

Thoroughly screening the entire genome for DNA copy number gains and losses only became possible with the advent of comparative genomic hybridization (CGH; Box 10). NB tumors are typically characterized by gains and losses of various chromosomes and chromosome parts, coined numerical (NCAs) and segmental chromosomal aberrations (SCAs), respectively. Based on CGH data of 231 primary NB tumors, Vandesompele et al. demonstrated the existence of three clinicogenetic subgroups with distinct genomic aberration patterns and clinical variables. The first subgroup (subtype 1) represents predominantly near-triploid low-stage tumors with NCAs and favorable histology from infants with excellent outcome. The other two subgroups (subtype 2A and 2B) mainly contain near-diploid/tetraploid high-stage tumors with SCAs and unfavorable histology. Typically, subtype 2B is featured by *MYCN* amplification and 1p deletions (§1.2.1.5.2.1), while most of the tumors of subtype 2A have non-amplified *MYCN* and 11q loss (§1.2.1.5.2.2). Importantly, it was further shown that OS probabilities are worse for patients with subtype 2A or B tumors (with SCAs) compared to patients with subtype 1 tumors (without SCAs) [211]. Later on, several other studies confirmed that a genomic profile characterized by SCAs is associated with a higher risk of relapse [212–214]. In the study of Schleiermacher et al. analyses were focused on tumors of infants with *MYCN* non-amplified localized unresectable/disseminated NB, in order to determine whether profiling of SCAs is useful for therapeutic stratification in this specific patient group. As this study clearly showed that the presence of SCAs is associated with worse progression-free survival, it was decided to use SCA profiling results for therapy stratification in the SIOPEN Low and Intermediate Neuroblastoma European Study (LINES), in an attempt to lower treatment burden in this patient group [214].

1.2.1.5.2.1 Chromosome 1p

One of the first discovered SCAs in NB is deletion of chromosome 1p [215]. In search of tumor suppressor genes, loss of heterozygosity (LOH) studies have mapped the smallest region of overlapping deletions (SRO) between tumors to 1p36 [216]. It occurs in approximately 25% of the cases and is associated with other adverse prognostic factors (age 12 months or older, INSS stage 4, *MYCN* amplification, unfavorable histology and diploidy). Attiyeh et al. showed that 1p36 LOH is highly associated with poor outcome and is independently predictive of worse EFS in patients without *MYCN* amplification [217].

1.2.1.5.2.2 Chromosome 11q

LOH of chromosome band 11q23 is detected in about 34% of NB tumors. These tumors are characterized by LOH of the entire chromosome 11 or by LOH of 11q with retention of 11p material (unbalanced 11q LOH). 11q23 LOH primarily occurs in tumors without *MYCN* amplification and chromosome 1p deletion, and is associated with the presence of INSS stage 4 and unfavorable

histology, as well as with the favorable prognostic factor of hyperdiploidy [217]. Of note, this aberration frequently appears with concurrent loss of 3p and gain of 17q [218]. Survival analyses indicate that 11q23 LOH is linked with a decreased probability of event-free survival (EFS). In patients without *MYCN* amplification, unbalanced 11q LOH leads to a decrease in both EFS and overall survival (OS) [217].

1.2.1.5.2.3 Chromosome 17q

Gain of material from chromosome 17 is the most frequent genetic abnormality of NB and may consist of an entire chromosome 17 or only the distal segment of 17q (17q21-qter; in approximately 54% of NB tumors). The principal mechanism underlying gain of 17q is an unbalanced translocation with a variety of partner chromosomes (e.g. 1p). Gain of 17q is strongly associated with INSS stage 4, age of 12 months or more at diagnosis, 1p deletion, amplification of *MYCN* and diploidy or tetraploidy [219]. Vandesompele et al. have shown that patients with tumors with a normal chromosome 17 status (in conjunction with other defects) or with a 17q gain have a worse prognosis compared to patients whose tumors bear whole chromosome 17 gain [211].

1.2.2 mRNA and miRNA-based prognostic markers

More recently, tumor mRNA expression profiling by microarray and qPCR analyses identified single gene prognostic markers (e.g. *NTRK1* [220]), as well as prognostic multimarker mRNA signatures. For example, Oberthuer et al. designed a customized oligonucleotide microarray covering a high percentage of transcripts previously related to NB tumor behavior to profile 251 tumors, which allowed them to construct a prognostic 144-gene signature [221]. Vermeulen et al. developed a qPCR assay for 59 prognostic genes, identified by re-analysis of previously published studies, and showed that this signature is an accurate, independent predictor of outcome by profiling 579 NB tumors [222]. Additionally, also miRNAs, i.e. non-coding RNAs regulating mRNA expression, have been source of prognostic biomarker research and several miRNA-based signatures have been proposed [223, 224]. Importantly, these RNA-based prognostic signatures require independent validation and performance evaluation in prospective studies before they can be introduced into clinical settings.

1.2.3 DNA methylation-based prognostic markers

An overview of DNA methylation-based prognostic markers in NB is given in the following review paper, which summarizes the most important findings up until a few years ago.

review paper

**Neuroblastoma epigenetics:
from candidate gene approaches to
genome-wide screenings**

Anneleen Decock, Maté Ongenaert, Jo Vandesompele and Frank Speleman

published in Epigenetics 6(8):962-970, 2011

impact factor 2011: 4.318

times cited (d.d. 11/09/2016; Google Scholar): 29

REVIEW PAPER: NEUROBLASTOMA EPIGENETICS: FROM CANDIDATE GENE APPROACHES TO GENOME-WIDE SCREENINGS

Anneleen Decock*, Maté Ongenaert*, Jo Vandesompele and Frank Speleman

**Contributed equally. Correspondence to franki.speleman@ugent.be.*

Center for Medical Genetics, Ghent University Hospital, Medical Research Building, De Pintelaan 185, Ghent, Belgium.

Contribution of AD: In a joint effort, AD participated in collecting and structuring data, as well as in drafting the manuscript, mainly focusing on the candidate gene case study on apoptosis and methylation in NB.

Abstract

Neuroblastoma (NB) is a childhood tumor originating from sympathetic nervous system cells. Although recently new insights into genes involved in NB have emerged, the molecular basis of NB development and progression still remains poorly understood. The best-characterized genetic alterations include amplification of the proto-oncogene *MYCN*, *ALK* activating mutations or amplification, gain of chromosome arm 17q and losses of 1p, 3p and 11q. Epigenetic alterations have been described as well: caspase 8 (*CASP8*) and RAS association domain family 1 isoform A (*RASSF1A*) DNA methylation are important events for the development and progression of NB. In total, about 75 genes are described as epigenetically affected in NB cell lines and/or NB primary samples. These epigenetic alterations were either found using a candidate gene approach or based on the analysis of genome-wide screening techniques. This review gives an extensive overview of all epigenetic changes described in NB as of today, with a main focus on both prognostic use and the potential of genome-wide techniques to find epigenetic prognostic biomarkers in NB. We summarize the key findings so far and the state-of-the-art of the upcoming methods at a unique time frame in the transition towards combined genome-wide chromatin immunoprecipitation (ChIP) and DNA sequencing techniques.

Keywords: epigenetics, apoptosis, DNA methylation, histone modifications, chromatin modification, neuroblastoma, review

Introduction and background

Neuroblastoma (NB) is a childhood tumor originating from sympathetic nervous system cells. The molecular basis of NB development and progression is still poorly understood. The best-characterized genetic alterations include amplification of the proto-oncogene *MYCN*, amplification and mutation of the *ALK* gene, gain

of chromosome arm 17q and losses of 1p, 3p and 11q. Classical risk factors include the age at diagnosis, *MYCN* amplification status and stage of the disease. The last decade, DNA methylation research has been conducted in NB, revealing that silencing of caspase 8 (*CASP8*) and RAS association domain family 1 isoform A (*RASSF1A*) are important in the development and progression of the disease. Both genes are often

found to be methylated in primary NB samples and the methylation status of these genes is significantly associated with survival.

The use of genome-wide screening techniques, such as re-expression analysis after treatment with 5-aza-2'-deoxycytidine (DAC), promoter assays after affinity-based capture (for instance using a 5-methylcytidine antibody), methylation microarrays after bisulfite treatment, and next-generation techniques, has led to an immense increase in both throughput and genome coverage. In NB, as of today, about 75 different DNA methylation biomarkers are described in scientific literature using various detection techniques. A selection of these research papers has shown the potential to use epigenetic biomarkers for prognostic purposes (survival, risk classification). This review summarizes all available epigenetics data in NB, with a focus on the use of DNA methylation biomarkers in predicting prognosis. We present this review at a unique time frame in this post-genomic era where next-generation sequencing technologies have become feasible to be used for whole-genome analysis at a reasonable cost. We will thus emphasize the potential of such genome-wide detection technologies in the NB epigenetics perspective.

DNA methylation markers in neuroblastoma: methodological approach

A total of 61 publications on methylation in NB were included in this review. In these studies, DNA methylation is detected using (1) methylation-sensitive restriction enzymes or library enrichment towards methylated DNA using restriction enzymes recognizing CpG-rich sequences, (2) bisulfite treatment, converting unmethylated C to T or (3) affinity-based enrichment, using the 5-methylcytosine antibody (methylated DNA immunoprecipitation (MeDIP)) or proteins containing a methyl-CpG-binding domain (MBD), such as MBD2 and MeCP2. In order to detect histone modifications,

specific antibodies can be used.

In addition, several studies make use of a pharmacologic unmasking strategy on NB cell lines. Using DAC, a cytosine analogue which will replace cytosine after several cell divisions, DNA methylation is inhibited, causing a demethylating effect. Combined with trichostatin-A (TSA), an inhibitor of histone deacetylases (HDACs) inhibiting the deacetylation of histone tails, this causes transcriptionally silenced genes (by DNA methylation and/or histone deacetylation) to become active again. These re-expression or reactivation events can be identified by comparing expression levels before and after treatment with DAC and/or TSA by making use of qPCR or expression microarrays. The latter approach can thus be used to identify possibly epigenetically silenced regions in NB cell lines in a genome-wide way. In contrast, most analyses on primary NB samples are candidate-based: the methylation status of the promoter regions of a rather limited number of genes (usually less than 10-15) is tested. These candidates either arise based on the reactivation studies in NB cell lines or are chosen as they have shown to be involved in NB. The candidates are for instance reported as being significantly lower expressed in high-risk tumors versus low-risk tumors or are important in the NB biology.

In total, about 75 different genes have been described as methylated (in varying degrees) in NB. These genes can be grouped based on their main function (cell cycle control, cell invasion and architecture, apoptosis-related genes, etc.). An overview of all reported genes and some summarizing information (range of methylation degree in cell lines and primary samples and associations with patient characteristics such as age at diagnosis, INSS stage and *MYCN* amplification status) is given in Supplemental Table 1. Very detailed information on each of these 75 genes, reported as being epigenetically modified in NB, is given in the supplements. All of these papers discuss DNA methylation; only

two papers describe histone modifications (*PTGER2* and *NSD1*) as detected by ChIP [1,2].

Candidate gene case study: apoptosis and methylation in neuroblastoma

As a case study to illustrate DNA methylation research in NB, the pathway of death receptors (DRs), decoy receptors (DcRs) and RASSF1A, and their relationship with apoptosis and caspases, is chosen and is depicted in Figure 1. Although methylation of caspase 8 (*CASP8*) in NB was first described in 2000 [3], only four years later it was already considered a key event in the biology of NB, next to *MYCN* amplification and other genetic factors [4]. From that first discovery on, other researchers soon started to investigate the methylation status of *CASP8* and other genes implied in the apoptotic pathway in their NB sample sets. This so called candidate gene approach, where a limited number of genes are tested for their methylation status based on prior evidence, led to the further discovery of various methylated genes related to the extrinsic and intrinsic apoptotic route. In this perspective, methylation of *RASSF1A* was discovered and found to be frequently methylated in NB [5].

Caspase 8 pathway

In most cases, the apoptotic process is characterized by the proteolytic activity of caspases. The two main apoptotic pathways in which caspase activation occurs, are the intrinsic mitochondrial pathway and the extrinsic receptor-mediated pathway [6].

The extrinsic pathway is initiated by binding of extracellular death signals with DRs on the surface of the target cells. These DRs belong to the tumor necrosis factor receptor (TNFR) superfamily and members of this family include TNFR1 (DR1, CD120a, p55 or p60), CD95 (DR2, APO1 or FASR), DR3 (also TNFRSF25, APO3, LARD, TRAMP or WSL1), TRAILR1 (DR4, APO2 or TNFRSF10A), TRAILR2 (DR5, KILLER, TRICK2 or

TNFRSF10B), DR6 (TNFRSF21), ectodysplasin A receptor (EDAR) and nerve growth factor receptor (NGFR). These receptors are characterized by a specific cytoplasmatic domain of approximately 80 amino acids, the so called death domain (DD). By homotypic interaction, this DD recruits specific adaptor proteins depending on the type of the stimulated receptor. For CD95 for example, FAS-associated via death domain (FADD) is recruited. FADD carries a death effector domain (DED) which associates with the DED of procaspase 8. This clustering of proteins results in the formation of the death-inducing signaling complex (DISC), which leads to activation of procaspase 8. Once caspase 8 is activated, the execution phase of apoptosis is triggered.

The intrinsic pathway is induced by cytotoxic signals such as DNA damage. These signals cause changes in the mitochondrial membrane, leading to an opening of the mitochondrial permeability transition pore (PTP), loss of the mitochondrial transmembrane potential and release of pro-apoptotic proteins which are normally sequestered in the intermembrane space of the mitochondria such as cytochrome C. In the presence of (deoxy)adenosine triphosphate, the cytoplasmic cytochrome C clusters with the apoptotic protease activating factor 1 (APAF1). This adaptor protein contains a caspase recruitment domain (CARD) that allows binding with the CARD of procaspase 9. The formed apoptosome complex ultimately activates caspase 9, which then cleaves other targets.

The extrinsic and intrinsic apoptotic pathways are also interlinked. Caspase 8 for example cleaves the BID protein, resulting in truncated BID (tBID) which inactivates BCL-2. This protein is an anti-apoptotic member of the B cell lymphoma 2 family which regulates the permeabilization of the mitochondria. Both pathways eventually merge into the same degrading execution phase. Here, execution caspases, such as caspase 3, are activated and cleave several cytoskeletal and nuclear proteins

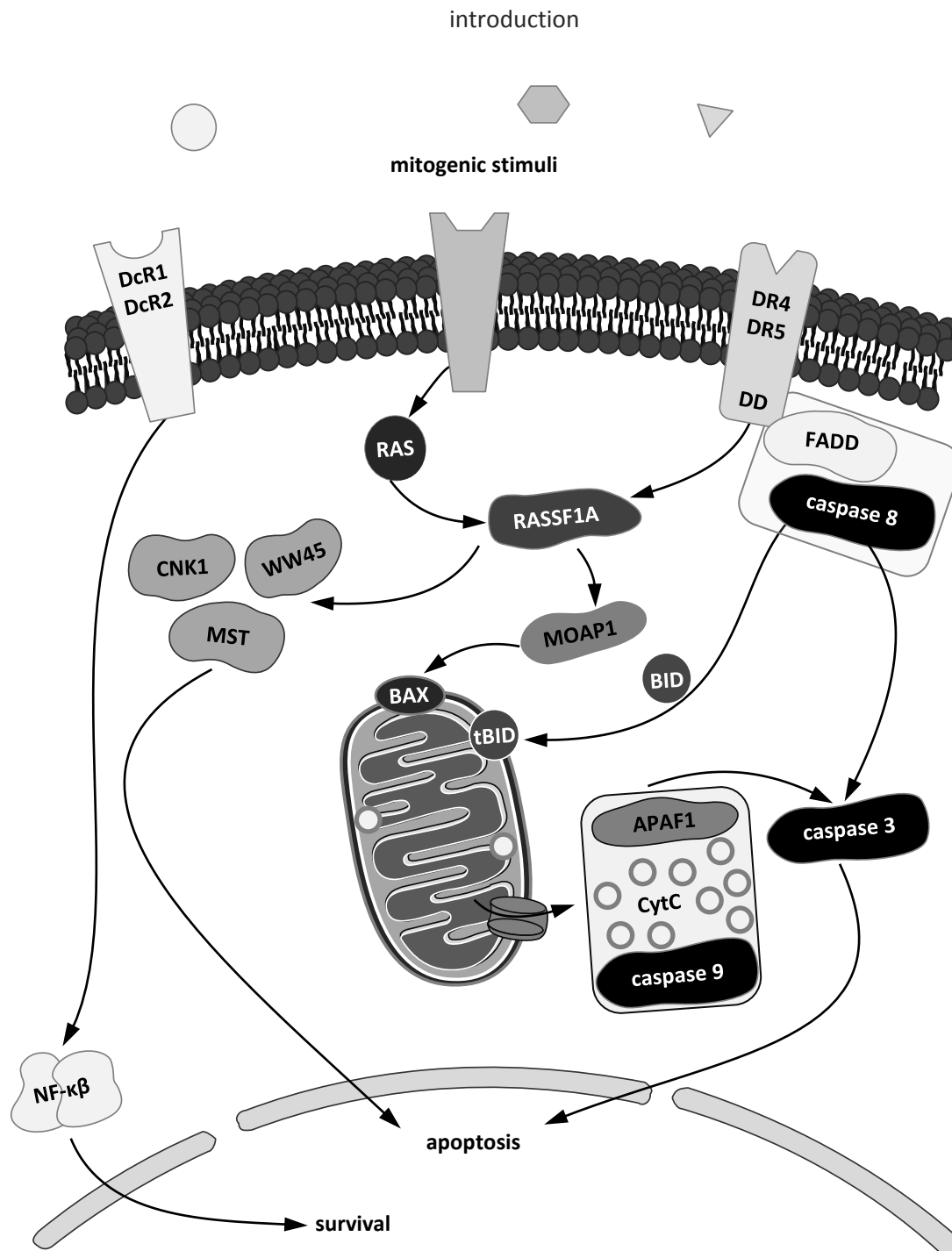


Figure 1. Simplified schematic representation of the death receptors, decoy receptors and RASSF1A pathways related to apoptosis. Death receptor (DR) stimulation leads to the formation of the death-inducing signaling complex (DISC; indicated by the rectangle near the cell membrane). The intrinsic apoptotic pathway is characterized by the formation of the apoptosome (indicated by the rectangle near the mitochondrion). The genes coding for DcR1, DcR2, DR4, DR5, caspase 8 and RASSF1A are reported to be methylation biomarkers in neuroblastoma. APAF1: apoptotic protease activating factor 1; BAX: BCL-2 associated X-protein; BID: BH3 interacting death domain; CNK1: connector enhancer of kinase suppressor of RAS; CytC: cytochrome C; DcR1: decoy receptor 1, TRAILR3, TNFRSF10C or TRID; DcR2: decoy receptor 2, TRAILR4, TNFRSF10D or TRUND; DD: death domain; DR4: TRAILR1, APO-2 or TNFRSF10A; DR5: TRAILR2, KILLER, TRICK2 or TNFRSF10B; FADD: FAS-associated via death domain; MOAP-1: modulator of apoptosis-1; MST: mammalian sterile 20 like kinase; NF κ B: nuclear factor kappabeta; RASSF1A: RAS association domain family 1 isoform A; tBID: truncated BID; WW45: human Salvador homolog.

(structural and signaling proteins or kinases) like GDID4, PARP, GAS2 and Lamin-A, and thus cause apoptosis.

Death ligands also interact with DcRs such as TRAILR3 (DcR1, TNFRSF10C or TRID), TRAILR4 (DcR2, TNFRSF10D or TRUNDD), DcR3 (TNFRSF6B, TR6 or M68) and osteoprotegerin (OPG). These receptors belong to the TNFR superfamily, but completely lack a DD or contain a truncated, non-functional DD. They can bind the extracellular death signals, but are unable to associate with the intercellular signaling molecules of apoptosis. They thus counteract the DRs. TRAILR3 and TRAILR4 compete with DR4 and DR5 for binding of APO2L/TRAIL. DcR3 competes with CD95 for binding of FASL and with DR3 for binding of TL1A. DcR stimulation can activate the transcription factor NFκB that directs the transcription of anti-apoptotic genes, such as C-FLIP, BCL-XL and IAPs, promoting cell survival.

In these pathways, the DRs *TRAILR1* and *TRAILR2*, the DcRs *TRAILR3* and *TRAILR4*, *CASP8* and *APAF1* are described to be methylated in NB cell lines and primary tumors. Thus, DNA methylation affects the apoptotic pathways both at the receptor and downstream signaling level, and thereby inhibits apoptosis. The downregulation of the DcRs in NB is a puzzling feature, because it renders cancer cells more susceptible to TRAIL-induced apoptosis and, thus, would counteract tumorigenesis. According to van Noesel et al. this could be seen as a protective response against tumor formation or progression. In this view, TRAILR3 and TRAILR4 downregulation represents a physiological response of the (pre)cancerous cell to a cellular state in which a higher level of apoptotic sensitivity is warranted [7]. Furthermore, in NB, increased levels of the anti-apoptotic BCL-2 and BCL-XL have been observed, and correlate with decreased apoptosis [8,9] and poor prognostic factors like *MYCN* amplification and unfavorable histology [10]. The methylation status of a selection of the reported genes is frequently

described as being associated with prognostic factors (stage, survival, age, risk) and is described in prognostic methylation biomarker signatures. *TRAILR2* is only described as being methylated in NB cell lines [7]. The fraction of NB patients that show methylation in specific regions of these genes and the relationship with methylation and clinical risk factors is given in Table 1. The gene coding for TMS1 (PYCARD), a member of a superfamily that mediates assembly of large signaling complexes in the inflammatory and apoptotic signaling pathways via the activation of caspases, is also reported methylated in NB and its methylation state is associated with *MYCN* amplification, stage and risk [11,12].

RAS association domain family 1 isoform A pathway

In addition to *CASP8*, another frequently methylated gene in NB is *RASSF1A*, a member of the RAS association domain family of proteins. Due to its interaction with multiple partners, RASSF1A influences a diversity of signaling pathways. Although this complex signaling network complicates unraveling the precise functional and biological relevance of RASSF1A, it is known to be implicated in the regulation of cell proliferation and apoptosis. As RASSF1A is characterized by a RAS association domain, one of the most obvious interaction partners of RASSF1A is the membrane-bound GTPase RAS. The interaction probably arises through heterodimerization of RASSF1A with RASSF5 which directly binds activated RAS [23]. RASSF1A also interacts with the scaffold protein connector enhancer of kinase suppressor of RAS (CNK1) and the pro-apoptotic mammalian sterile 20 like kinase (MST), through which these also contribute to the pro-apoptotic signaling initiated by activated RAS [24]. The MST kinases take part in a pro-apoptotic tumor suppressor kinase cascade via coupling with the adaptor protein WW45, which is also recruited by

Table 1. Methylation of death receptors, decoy receptors and *CASP8* in neuroblastoma cell lines, primary and relapsed tumors.

gene	samples	methylation number	%	clinical features	reference
<i>CASP8</i>	cell lines	5/9	56		[7]
	cell lines	8/10	80		[11]
	cell lines	11/12	92		[13]
	primary tumor	24/60	40		[5]
	primary tumor	6/44	14		[14]
	primary tumor	25/41	60		[13]
	primary tumor	39/70	56	survival; signature with <i>TNFRSF10D</i>	[15]
	primary tumor	17/45	38	stage	[16]
	relapsed tumor	6/17	35		[16]
	primary tumor	20/36	56		[17]
	primary tumor	52/70	74	survival	[18]
	primary tumor	10/11	91		[19]
	bone marrow	5/11	55		[19]
	primary tumor		52	stage; infiltrated bone marrow; risk; survival; signature with other apoptotic genes	[12]
<i>TNFRSF10D</i> (<i>DCR2</i>)	cell lines	6/9	67		[7]
	cell lines	10/10	100		[11]
	cell lines	14/14	100	signature with <i>CASP8</i>	[15]
	primary tumor	7/28	25		[7]
	primary tumor	13/31	42	<i>MYCN</i> ; signature with <i>RASSF1A</i> ; <i>SFN</i> and <i>TP73</i>	[20]
	primary tumor	31/70	44	survival	[15]
	primary tumor	11/45	25		[16]
	relapsed tumor	5/17	29		[16]
	primary tumor	24/86	28	stage; <i>MYCN</i> ; survival	[21]
<i>TNFRSF10C</i> (<i>DCR1</i>)	cell lines	6/9	67		[7]
	cell lines	8/10	80		[11]
	primary tumor	6/28	21		[7]
	primary tumor	5/45	11		[16]
	relapsed tumor	3/17	18		[16]
	primary tumor	6/11	55		[19]
	bone marrow	4/11	36		[19]
	primary tumor		50	stage; age; risk	[12]
<i>TNFRSF10A</i> (<i>DR4</i>)	cell lines	4/9	44		[7]
	cell lines	8/10	80		[11]
	cell lines	11/14	79		[15]
	primary tumor		51		[11]
	primary tumor	6/11	55		[19]
	bone marrow	4/11	36		[19]
	primary tumor		50	stage; age; risk	[12]
<i>TNFRSF10B</i> (<i>DR5</i>)	cell lines	2/9	22		[7]
<i>APAF1</i>	primary tumor	6/11	55		[19]
	bone marrow	4/11	46		[19]
	primary tumor		29		[12]

(continues)

Note. The number and percentage of samples with methylation reported is given, next to the number of tested samples. Associations with clinical risk factors are given when reported significant.

Table 1. Methylation of death receptors, decoy receptors and *CASP8* in neuroblastoma cell lines, primary and relapsed tumors.

(continued)

gene	samples	methylation		clinical features	reference
		number	%		
<i>PYCARD</i> (<i>TMS1</i>)	cell lines	8/10	80		[11]
	cell lines	13/14	93		[15]
	primary tumor		31	<i>MYCN</i>	[11]
	primary tumor	32/70	46		[15]
	primary tumor	3/18	17		[22]
	primary tumor	3/11	27		[19]
	bone marrow	4/11	36		[19]
	primary tumor		25	stage; <i>MYCN</i> ; risk	[12]

Note. The number and percentage of samples with methylation reported is given, next to the number of tested samples. Associations with clinical risk factors are given when reported significant.

RASSF1A [25]. All these interactions are obvious examples that RASSF1A functions as a scaffold for assembly of an apoptotic complex. Furthermore, RASSF1A is also tightly linked with the extrinsic apoptotic pathway. Upon DR stimulation RASSF1A binds modulator of apoptosis 1 (MOAP1), which enables the association of MOAP1 with BAX, and subsequent BAX activation leads to permeabilization of the outer mitochondrial membrane and apoptosis [26].

RASSF1A is reported to be methylated in a very high fraction of NB patients. Table 2 shows the methylation status of *RASSF1A* in NB samples and the association of *RASSF1A* methylation with clinical risk factors. Only about 7 kb upstream of *RASSF1A*, another promoter region frequently reported to be methylated in NB is zinc finger, myeloid, nervy and DEAF1-type containing 10 (*ZMYND10* (*BLU*)), which is related to stage, age and risk [12,16,27].

Genome-wide DNA methylation profiling in neuroblastoma

The candidate gene approach described in the above section is thus based on prior evidence: decreased expression in high-risk patients versus low-risk patients, reported methylation in other cancer types or involvement of the gene in NB biology. The number of promoter regions to be tested by such an approach is limited (usually

less than 10-15). In order to identify DNA methylation biomarkers that can be used for better prognosis prediction, it may be needed to perform a whole-genome screening, followed by an analysis strategy to prioritize potential DNA methylation biomarkers and validation of the top-scoring candidates.

During the last decades, a number of whole-genome DNA methylation detection methodologies have been developed: on one hand re-expression analysis after a demethylating treatment using DAC, on the other hand promoter arrays or sequencing after affinity-based capture using MeDIP, and methylation-specific microarrays or deep-sequencing after bisulfite treatment. So far, in NB research, whole-epigenome sequencing technologies have not been described.

Re-expression experiments after treatment with DAC

This pharmacologic unmasking strategy exists of expression profiling (using expression microarrays) before and after treatment with DAC. DAC is a cytosine analogue which cannot become methylated by the DNA methyltransferases (DNMTs). After several cell divisions, this causes a global demethylation of the genome. DAC can thus be considered a demethylating agent and an inhibitor of the DNMTs, which yields comparable results

compared to DNMT double knockout models, such as the DKO HCT-116 colorectal cancer cell line [30]. In addition, some studies use TSA in combination with DAC. TSA is an inhibitor of

HDACs and will thus cause open chromatin structures. Genes that are highly re-expressed after treatment with DAC and/or TSA (reactivated) may be silenced due to DNA

Table 2. Methylation of *RASSF1A* and *ZMYND10* in neuroblastoma cell lines, primary and relapsed tumors, and serum.

gene	samples	methylation		clinical features	reference
		number	%		
<i>RASSF1A</i>	cell lines	10/10	100		[11]
	cell lines	12/12	100		[13]
	primary tumor	37/67	55		[5]
	primary tumor	39/56	70	age; risk; survival	[28]
	primary tumor	51/67	76		[5]
	primary tumor	26/31	84	<i>MYCN</i> ; signature with <i>SFN</i> ; <i>TP73</i> ; <i>TNFRSF10D</i>	[20]
	primary tumor	34/41	83		[13]
	primary tumor	63/70	90		[15]
	primary tumor	42/45	93		[16]
	relapsed tumor	17/17	100		[16]
	primary tumor	29/41	71		[17]
	primary tumor	64/68	94		[29]
	serum	17/68	25	stage; <i>MYCN</i>	[29]
	primary tumor	11/11	100		[19]
	bone marrow	8/11	73		[19]
	primary tumor		75	stage; 1p deletion; infiltrated in bone marrow	[12]
<i>ZMYND10</i>	cell lines	6/7	86		[27]
	cell lines	7/11	66		[13]
	cell lines	7/10	70		[11]
	cell lines	13/14	93		[15]
	primary tumor	20/49	41	stage	[27]
	primary tumor	25/67	37		[7]
	primary tumor	3/35	8		[13]
	primary tumor	38/70	54		[15]
	primary tumor	15/45	34	stage	[16]
	relapsed tumor	4/17	24		[16]
	primary tumor	6/40	15		[17]
	primary tumor	8/11	73		[19]
	bone marrow	10/11	91		[19]
	primary tumor		35	age; risk	[12]

Note. The number and percentage of samples with methylation reported is given, next to the number of tested samples. Associations with clinical risk factors are given when reported significant.

methylation/histone modifications. This approach can only be applied on cell lines as demethylation occurs after several cell divisions. In NB, treatment with DAC and/or TSA is applied in more than half of the papers (over 30) describing epigenetics. However, only four publications perform a genome-wide expression study after a pharmacologic unmasking

screening strategy [28,31,32,34]. In all four publications, this led to the discovery of previously undescribed DNA methylation biomarkers in NB. In all other research papers, reactivation is shown by (quantitative) PCR for a limited set of genes of interest, in order to demonstrate the link between DNA methylation and transcriptional silencing in NB cell lines.

MeDIP chip

While the treatment with DAC and/or TSA leads to re-expression, this strategy is an indirect method to detect DNA methylation as the influence of the demethylating effect is measured at the transcriptional level. The last years, genome-wide detection techniques have been developed that are able to directly measure DNA methylation without the detour around (silenced) expression.

One of these techniques is the use of the 5-methylcytosine antibody (MeDIP) and combining this methylation-enriching immunoprecipitation with microarrays (MeDIP chip; tiling arrays or focused arrays such as promoter regions or CpG islands). This strategy is used by Murphy et al. [33] where the authors demonstrate that a large portion of methylated genes indeed shows reduced expression. In addition, the authors performed a ChIP experiment using *MYCN* and were able to define specific regions where *MYCN*-binding and DNA hypermethylation colocalize.

Methylation-specific arrays

Next to affinity-based methodologies to detect DNA methylation, bisulfite conversion has been used in all gold standard techniques such as methylation-specific PCR (MSP) and bisulfite sequencing. A frequently used high-throughput platform, making use of different probe types for methylated versus unmethylated bisulfite-treated DNA is the Illumina Infinium HumanMethylation27 BeadChip. This technique with over 27,000 methylation-specific probes has recently been applied on NB primary samples [34] and has led to the identification of several novel methylation biomarkers that show potential for prognostic use.

Next-generation sequencing techniques

All three earlier described detection principles

are also used in combination with next-generation sequencing platforms (such as Roche 454 and Illumina GAIIx) to detect DNA methylation in a genome-wide way. The complete honey bee and *Arabidopsis thaliana* methylomes have been determined by bisulfite sequencing [35,36]. As this technique would not be cost-efficient for larger genomes such as humans, one needs to decrease the fraction of the genome to be covered by sequencing. In order to enrich this limited library that is sequenced, restriction enzymes were used, making sure the retained fractions (based on sizing) cover the most dense CpG islands or regions, followed by deep bisulfite sequencing. Another methodology is to make use of the high affinity towards methylated cytosines of MBD-containing proteins such as MBD2 and MeCP2 [37,38] or by applying MeDIP [39], followed by (paired-end) sequencing.

Some research groups, including our group, are at this moment in the process of using sequencing techniques to determine the methylome and/or to detect histone modifications in NB samples (cell lines and primary patients). The first results indicate the potential of epigenetic sequencing techniques: at a feasible cost, it becomes feasible to investigate the methylome or epigenome of a sample in great detail. Their findings will reveal the complexity of the epigenetic changes during development and progression of NB and eventually would allow assessment of the effect on the epigenome of diverse treatments, and which patients would be eligible for which type of treatment (prediction and personalized medicine).

Epigenetics and neuroblastoma prognosis

Despite advances in multimodal anticancer therapies, survival rates for children with NB remain disappointingly low. Although current risk assessment schemes have been significantly approved, inevitably undertreatment or

overtreatment will still occur for certain children. As a consequence, survival rates are suboptimal in the low-risk group, and a number of patients in the so-called high-risk group are

unnecessarily put at risk for potential long term side effects of the toxic therapy. Clearly, a more objective and accurate classifier is needed for improved outcome prediction. Only then,

Table 3. Overview of studies discussing DNA methylation biomarkers and their relationship with prognostic risk factors and survival.

methylation biomarker (signature)	risk factors					survival		reference
	age	stage	MYCN	risk	other	overall	event-free	
<i>RASSF1A</i>	✓	✓		✓		✓		[40]
<i>TNFRSF10D</i>		✓	✓			✓	✓	[21]
<i>HOXA9/RARB</i>						✓		[11]
<i>PCDHA/PCDHB/HLP/CYP26C1</i>	✓	✓	✓		ploidy TrKA	✓		[41]
<i>CASP8</i>						✓		[18]
<i>RASSF1A, SFN, TP73, TNFRSF10D</i>			✓			✓		[20]
<i>EMP3</i>	✓		✓			✓		[42]
<i>PYCARD, MGMT, RARB</i>						✓		[19]
<i>CASP8, TNFRSF10D</i>						✓		[15]
<i>CASP8, ZMYND10</i>						✓		[16]
<i>NR1I2</i>		✓	✓			✓		[43]
<i>SFN, RASSF1A, CYP26C1, TNFRSF10D</i>				✓		✓	✓	[44]
<i>RASSF1A</i>	✓	✓	✓			✓	✓	[29]
<i>SCNN1A, PRKCDBP, KRT19</i>						✓	✓	[34]
apoptotic genes, such as <i>TMS1, APAF1</i> and <i>CASP8</i>				✓		✓	✓	[12]

Note. Overall survival or event-free survival may be disease-free, relapse-free or progression-free. Only significant associations are reported (p-values according to the original publication).

patients will receive the most appropriate therapy, can be monitored more intensively if needed, and become eligible for new experimental therapies.

As briefly demonstrated in the death receptor pathway and for *RASSF1A*, several methylation biomarkers in NB are described to be associated with classical risk factors: *MYCN* amplification; age at diagnosis and stage. Several publications describe the use of (a combination of) DNA methylation biomarkers to show the effect of methylation on disease-free survival or overall survival. The power and potential of DNA methylation as non-invasive biomarkers is demonstrated by measuring DNA methylation of *RASSF1A* in serum of patients [40]. DNA methylation-based prognosis has the potential to be non-invasive, highly sensitive and specific.

It can be detected using standard PCR or sequencing technologies.

In many studies, the genes discussed in previous sections are used for prognostic purposes. However, more recently genome-wide screening methodologies are becoming more popular and feasible (re-expression analysis after treatment with DAC, promoter chips after capturing with MeDIP, methylation microarrays). This revealed a broader view on the NB methylome and allowed the identification of DNA methylation signatures that can be used for prognostic purposes. Table 3 lists all studies where the DNA methylation state demonstrates prognostic potential. The association with classical risk factors is indicated if significant, as well as the impact on survival. It can be noticed that, next to candidate genes previously described such as

CASP8 and *RASSF1A*, several novel methylation biomarkers that are related to prognosis, are described.

Discussion and future perspectives

As in other cancer types, it clearly has been shown that NB is a disease that is regulated by epigenetic mechanisms. Candidate gene approaches led to the discovery that frequently methylated regions in NB can be found in pathways related to apoptosis. Classical tumor suppressor genes do not show such a clear methylation signal in NB.

Current technologies have the possibility to investigate the epigenetic changes in a genome-wide way, and integrating these data with other data sources, such as mRNA and miRNA expression profiles and proteome data, would reveal the (epi)genomic landscape of NB. Such multidimensional and cross-species integrated information will be exploited using high-end bioinformatic tools and systems biology approaches in order to unravel the various implicated perturbed signaling pathways, complex interactions and cross-talk between critical nodes within these networks. Understanding such regulatory networks at play and in particular possible compensatory interactions may also be crucial for developing appropriate therapeutic strategies and to anticipate to and design strategies against new forms of therapy resistance.

Supplemental materials

Supplemental materials can be found at: www.tandfonline.com/doi/suppl/10.4161/epi.6.8.16516?scroll=top

Supplemental Table 1. DNA methylation in neuroblastoma cell lines and primary tumors.

For each gene reported to be methylated in NB, the percentage of methylated samples is given (according to the original references) and

associations with clinical risk factors are described. Below the table, detailed information (gene function and general information, features of the promoter region, and information related to DNA methylation analyses) on each of these genes can be found.

Acknowledgments

This study was supported by the Emmanuel Van der Schueren Foundation (scientific partner of the Flemish League Against Cancer (VLK)) and by the Fournier-Majoie Foundation (FFM).

References

- [1] Berdasco M et al. (2009). Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. *Proceedings of the National Academy of Sciences of the USA*; 106(51):21830-21835.
- [2] Sugino Y et al. (2007). Epigenetic silencing of prostaglandin E receptor 2 (PTGER2) is associated with progression of neuroblastomas. *Oncogene*; 26(53):7401-7413.
- [3] Teitz T et al. (2000). Caspase 8 is deleted or silenced preferentially in childhood neuroblastomas with amplification of MYCN. *Nature Medicine*; 6(5):529-535.
- [4] van Noesel MM et al. (2004). Pediatric neuroblastomas: genetic and epigenetic "danse macabre." *Gene*; 325:1-15.
- [5] Astuti D et al. (2001). *RASSF1A* promoter region CpG island hypermethylation in pheochromocytomas and neuroblastoma tumours. *Oncogene*; 20(51):7573-7577.
- [6] Elmore S (2007). Apoptosis: a review of programmed cell death. *Toxicologic Pathology*; 35(4):495-516.
- [7] van Noesel MM et al. (2002). Tumor-specific downregulation of the tumor necrosis factor-related apoptosis-inducing ligand decoy receptors DcR1 and DcR2 is associated with dense promoter hypermethylation. *Cancer Research*; 62(7):2157-2161.
- [8] Ikegaki N et al. (1995). Relationship between bcl-2 and myc gene expression in human neuroblastoma. *Cancer Letters*; 91(2):161-168.
- [9] Ikeda H et al. (1995). Bcl-2 oncoprotein expression and apoptosis in neuroblastoma. *Journal of Pediatric Surgery*; 30(6):805-808.

- [10] Castle VP et al. (1993) Expression of the apoptosis-suppressing protein bcl-2, in neuroblastoma is associated with unfavorable histology and N-myc amplification. *American Journal of Pathology*; 143(6):1543-1550.
- [11] Alaminos M et al. (2004). Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma. *Journal of the National Cancer Institute*; 96(16):1208-1219.
- [12] Grau E et al. (2010). Hypermethylation of apoptotic genes as independent prognostic factor in neuroblastoma disease. *Molecular Carcinogenesis*; 50(3):153-162.
- [13] Lazcoz P et al. (2006). Frequent promoter hypermethylation of RASSF1A and CASP8 in neuroblastoma. *BMC Cancer*; 6:254.
- [14] Gonzalez-Gomez P et al. (2003). Aberrant methylation of multiple genes in neuroblastic tumours relationship with MYCN amplification and allelic status at 1p. *European Journal of Cancer*; 39(10):1478-1485.
- [15] Yang Q et al. (2007). Methylation of CASP8, DCR2 and HIN-1 in neuroblastoma is associated with poor outcome. *Clinical Cancer Research*; 13(11):3191-3197.
- [16] Michalowski MB et al. (2008). Methylation of tumor-suppressor genes in neuroblastoma: The RASSF1A gene is almost always methylated in primary tumors. *Pediatric Blood and Cancer*; 50(1):29-32.
- [17] Hoebeeck J et al. (2009). Aberrant methylation of candidate tumor suppressor genes in neuroblastoma. *Cancer Letters*; 273(2):336-346.
- [18] Kamimatsuse A et al. (2009). Detection of CpG island hypermethylation of caspase-8 in neuroblastoma using an oligonucleotide array. *Pediatric Blood and Cancer*; 52(7):777-783.
- [19] Grau E et al. (2010) Epigenetic alterations in disseminated neuroblastoma tumour cells: influence of TMS1 gene hypermethylation in relapse risk in NB patients. *Journal of Cancer Research and Clinical Oncology*; 136(9):1415-1421.
- [20] Banelli B et al. (2005). Distinct CpG methylation profiles characterize different clinical groups of neuroblastic tumors. *Oncogene*; 24(36):5619-5628.
- [21] Yagyu S et al. (2008). Circulating methylated-DCR2 gene in serum as an indicator of prognosis and therapeutic efficacy in patients with MYCN nonamplified neuroblastoma. *Clinical Cancer Research*; 14(21):7011-7019.
- [22] Shahi MH et al. (2010). Expression and epigenetic modulation of sonic hedgehog-GLI1 pathway genes in neuroblastoma cell lines and tumors. *Tumour Biology*; 32(1):113-127.
- [23] Ortiz-Vega S et al. (2002). The putative tumor suppressor RASSF1A homodimerizes and heterodimerizes with the Ras-GTP binding protein Nore1. *Oncogene*; 21(9):1381-1390.
- [24] Rabizadeh S et al. (2004). The scaffold protein CNK1 interacts with the tumor suppressor RASSF1A and augments RASSF1A-induced cell death. *Journal of Biological Chemistry*; 279(28):29247-29254.
- [25] Guo C et al. (2007). RASSF1A is part of a complex similar to the Drosophila Hippo/Salvador/Lats tumor-suppressor network. *Current Biology*; 17(8):700-705.
- [26] Baksh S et al. (2005). The tumor suppressor RASSF1A and MAP-1 link death receptor signaling to Bax conformational change and cell death. *Molecular Cell*; 18(6):637-650.
- [27] Agathangelou A et al. (2003). Epigenetic inactivation of the candidate 3p21.3 suppressor gene BLU in human cancers. *Oncogene*; 22(10):1580-1588.
- [28] Yang Q et al. (2004). Methylation-associated silencing of the heat shock protein 47 gene in human neuroblastoma. *Cancer Research*; 64(13):4531-4538.
- [29] Misawa A et al. (2009). RASSF1A hypermethylation in pretreatment serum DNA of neuroblastoma patients: a prognostic marker. *British Journal of Cancer*; 100(2):399-404.
- [30] Schuebel KE et al. (2007). Comparing the DNA hypermethylome with gene mutations in human colorectal cancer. *PLoS Genetics*; 3(9):1709-1723.
- [31] Margetts CDE et al. (2008). Evaluation of a functional epigenetic approach to identify promoter region methylation in pheochromocytoma and neuroblastoma. *Endocrine-related Cancer*; 15(3):777-786.
- [32] Buckley PG et al. (2010). Genome-wide DNA methylation analysis of neuroblastic tumors reveals clinically relevant epigenetic events and large-scale epigenomic alterations localized to telomeric regions. *International Journal of Cancer*; 128(10):2296-2305.
- [33] Murphy DM et al. (2009). Global MYCN transcription factor binding analysis in neuroblastoma reveals association with distinct E-box motifs and regions of DNA hypermethylation. *PloS One*; 4(12):e8154.
- [34] Caren H et al. (2011). Identification of epigenetically regulated genes that predict patient outcome in neuroblastoma. *BMC Cancer*; 11:66.
- [35] Lyko F et al. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biology*;

- 8(11):e1000506.
- [36] Cokus SJ et al. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*; 452(7184):215-219.
 - [37] Serre D et al. (2010). MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Research*; 38(2):391-399.
 - [38] Rauch T et al. (2006). MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. *Cancer Research*; 66(16):7939-7947.
 - [39] Ruike Y et al. (2010). Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. *BMC genomics*; 11:137.
 - [40] Yang Q et al. (2004). Association of epigenetic inactivation of RASSF1A with poor outcome in human neuroblastoma. *Clinical Cancer Research*; 10(24):8493-8500.
 - [41] Abe M et al. (2005). CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas. *Cancer Research*; 65(3):828-834.
 - [42] Alaminos M et al. (2005). EMP3, a myelin-related gene located in the critical 19q13.3 region, is epigenetically silenced and exhibits features of a candidate tumor suppressor in glioma and neuroblastoma. *Cancer Research*; 65(7):2565-2571.
 - [43] Misawa A et al. (2005). Methylation-associated silencing of the nuclear receptor 112 gene in advanced-type neuroblastomas, identified by bacterial artificial chromosome array-based methylated CpG island amplification. *Cancer Research*; 65(22):10233-10242.
 - [44] Banelli B et al. (2010). Outcome prediction and risk assessment by quantitative pyrosequencing methylation analysis of the SFN gene in advanced stage, high-risk, neuroblastic tumor patients. *International Journal of Cancer*; 126(3):656-668.

1.2.4 Established prognostic markers are combined into a pretreatment risk classification system

Combinations of prognostic variables are used for risk group assignment and treatment stratification. Given the rare occurrence of NB, cooperative pediatric oncology groups have been established in the past to be able to conduct large-scale analyses and in this way to fully optimize the stratification procedures. These include both national and international groups, such as the Children's Oncology Group (COG; North America and Australia), the German Pediatric Oncology and Hematology Group (GPOH), the Japanese Advanced Neuroblastoma Study Group (JANB), the Japanese Infantile Neuroblastoma Co-operative Study Group (JINCS) and the International Society of Pediatric Oncology Europe Neuroblastoma Group (SIOPEN). However, these cooperative groups initially selected different prognostic factors to tailor stratification, still complicating comparison of clinical trials performed throughout the world. In 2004, a task force of investigators from these cooperative groups, the so-called International Neuroblastoma Risk Group (INRG) Task Force, was set up to develop a consensus approach for pretreatment risk stratification of NB [154]. This effort led to the construction of the INRG classification system, next to the INRG Staging System (INRGSS; §1.2.1.3) and recommendation reports on the detection and evaluation of disseminated disease [143, 154, 188, 197, 225].

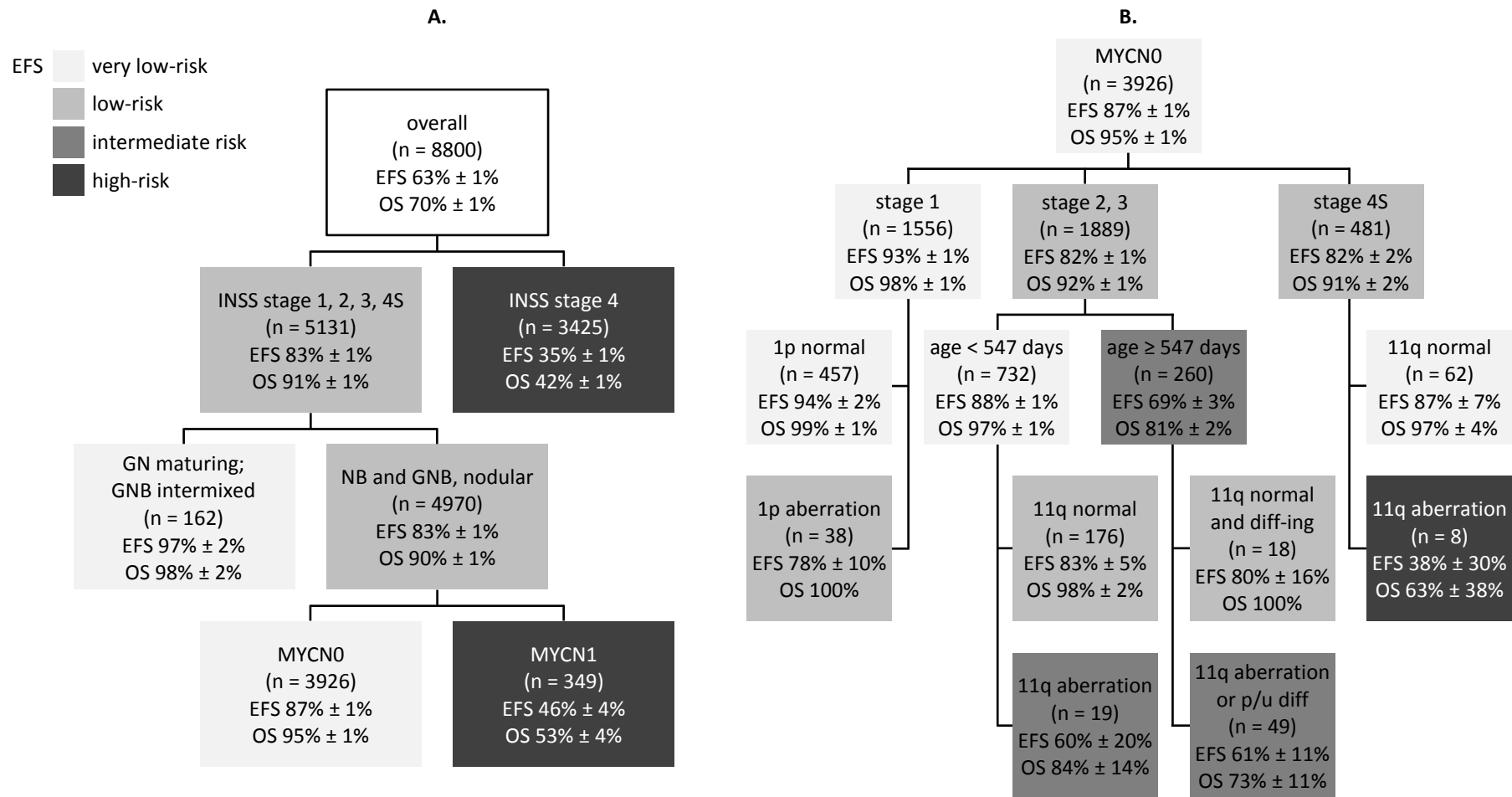
In total, the statistical and clinical significance of 13 potential prognostic factors were analyzed in a cohort of 8,800 patients, and using survival tree regression analysis (Figure 13) the INRG classification system was built. Currently, this system includes seven criteria (the INRG stage (INRGSS), age of the patient at diagnosis, histologic category, grade of tumor differentiation, *MYCN* amplification status, presence/absence of 11q aberrations and tumor cell ploidy) to stratify patients into sixteen pretreatment risk groups (lettered A through R; Table 7). Based on arbitrary cutoffs on the EFS data in the survival tree regression analysis, these pretreatment risk groups were further clustered into four categories: very low-risk (5-year EFS > 85%), low-risk (5-year EFS > 75 to ≤ 85%), intermediate-risk (5-year EFS ≥ 50 to ≤ 75%) and high-risk (5-year EFS < 50%). The proportion of patients grouped in these categories was 28.2%, 26.8%, 9.0% and 36.1%, respectively [154]. Of note, the survival tree regression analysis was performed using the INSS stage (§1.2.1.3), as the sample size of patients with known surgical risk factors (IDRFs that define INRGSS; §1.2.1.3) was too small relative to patients with known INSS stage [143]. However, the INSS stage is not suitable for pretreatment risk classification. For that reason, also posthoc analyses using the INRGSS were performed, which found both INSS stage and INRGSS highly prognostic of survival. This supports the translation of the survival tree regression analysis (in terms of INSS stage) into the INRG classification system (in terms of INRGSS) [143, 154]. To test the predictive ability of age, a consensus cutoff of 18 months was selected based on previous findings (§1.2.1.1). For patients with diploid, *MYCN* non-amplified stage M tumors, clinical justification was used to split patients younger than 12 months from 12 months and older to younger than 18 months of age [154].

Although the INRG Task Force recognized that genome-wide studies had identified powerful predictors of outcome, only *MYCN* amplification, chromosome 11q aberration and DNA ploidy status were included in the INRG pretreatment risk stratification, as microarray analyses of DNA copy number alterations and gene expression were not widely available at the time the classification was established. It is anticipated that the next-generation INRG classification system will incorporate profiles of the NB genome, transcriptome and epigenome to further improve prognostication [226].

Table 7. Using the International Neuroblastoma Risk Group (INRG) consensus pretreatment classification, neuroblastoma patients are stratified into risk groups.

INRG stage	age (months)	histologic category	grade of tumor differentiation	<i>MYCN</i>	11q aberration	ploidy	pretreatment risk group
L1/L2		GN maturing; GNB					A very low
L1		any, except GN maturing or GNB intermixed		MYCN0		B	very low
				MYCN1		K	high
L2	< 18	any, except GN maturing or GNB intermixed		MYCN0	no	D	low
					yes	G	intermediate
	≥ 18	GNB nodular; NB	differentiating	MYCN0	no	E	low
			poorly differentiated or undifferentiated	MYCN0	yes	H	intermediate
				MYCN1		N	high
M	< 18			MYCN0		hyperdiploid	F low
	< 12			MYCN0		diploid	I intermediate
	12 to < 18			MYCN0		diploid	J intermediate
	< 18			MYCN1			O high
	≥ 18					P	high
MS	< 18			MYCN0	no	C	very low
					yes	Q	high
				MYCN1		R	high

Note. For INRG stage, see Table 6. 12 months: 365 days; 18 months: 547 days; blank field: any; diploid: DNA index ≤ 1.0 ; hyperdiploid: DNA index > 1.0 and includes near-triploid and near-tetraploid tumors; EFS: event-free survival; GN: ganglioneuroma; GNB: ganglioneuroblastoma; MYCN1: *MYCN* amplified; MYCN0: *MYCN* non-amplified; NB: neuroblastoma [154].



(continues)

Figure 13. Using cutoffs on the event-free survival data in the survival tree regression analysis of the International Neuroblastoma Risk Group (INRG) cohort, four pretreatment risk group categories were determined: very low-risk, low-risk, intermediate-risk and high-risk. A split occurs for the most highly statistically significant factor as identified using a Cox proportional hazards regression model. **A.** Top levels of the overall tree. **B.** Subtree for NB and GNB-nodular, non-stage 4 MYCN0 patients. The split of stage 2, 3 from stage 4S patients was a clinical decision and not the result of statistical significance. EFS: 5-year event-free survival; GN: ganglioneuroma; GNB: ganglioneuroblastoma; INSS: International Neuroblastoma Staging System; MYCN0: *MYCN* non-amplified; MYCN1: *MYCN* amplified; NB: neuroblastoma; OS: overall survival. Based on [154].

(continued)

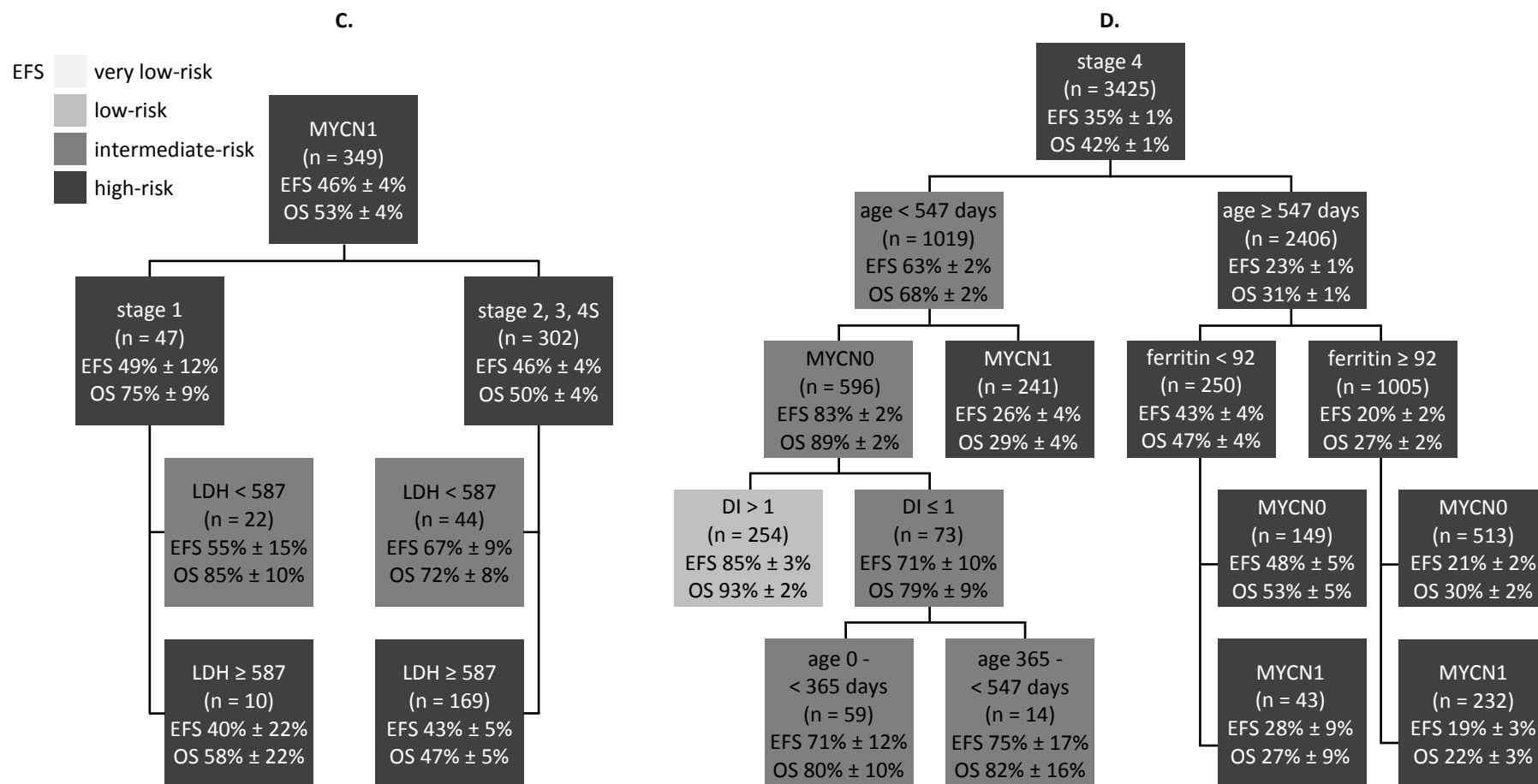
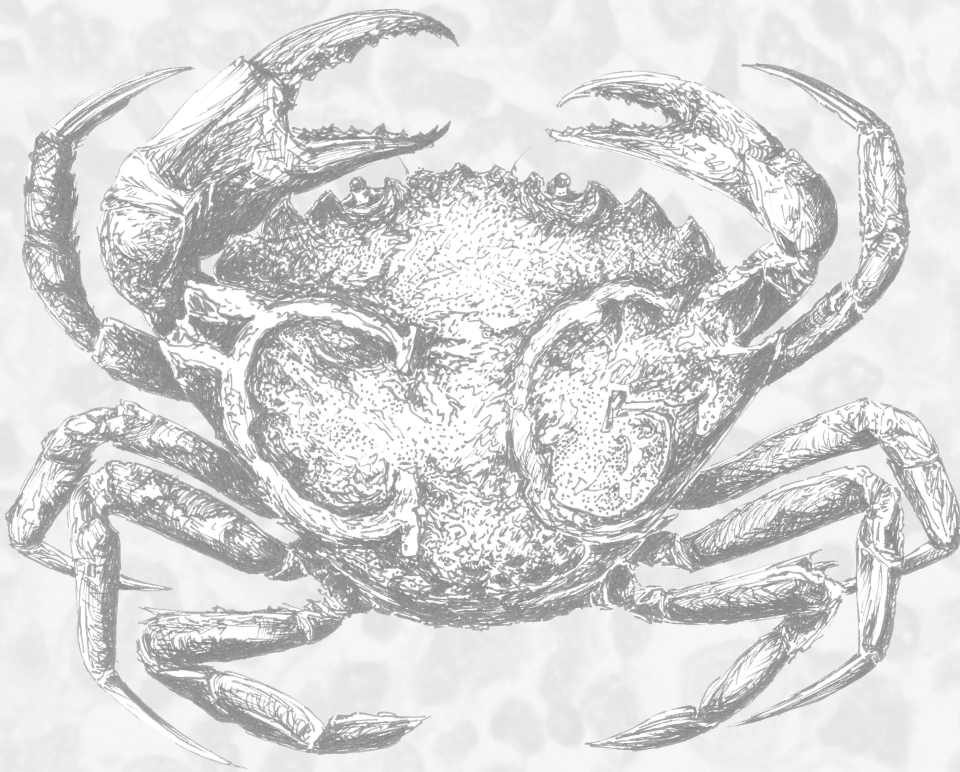


Figure 13. Using cutoffs on the event-free survival data in the survival tree regression analysis of the International Neuroblastoma Risk Group (INRG) cohort, four pretreatment risk group categories were determined: very low-risk, low-risk, intermediate-risk and high-risk. A split occurs for the most highly statistically significant factor as identified using a Cox proportional hazards regression model. **C.** Subtree for NB and GNB-nodular, non-stage 4 MYCN1 patients. The split of stage 1 from stage 2, 3, 4S patients was a clinical decision and not the result of statistical significance. LDH in U/l. **D.** Subtree for INSS stage 4 patients. The split of age 0 - < 365 days from age 365 - < 547 days was a clinical decision and not the result of statistical significance. Ferritin in ng/ml. DI: DNA index; EFS: 5-year event-free survival; LDH: lactate dehydrogenase; MYCN0: *MYCN* non-amplified; MYCN1: *MYCN* amplified; OS: overall survival. Based on [154].

2



Research outline

Unfortunately, the quest for prognostic NB markers has not yet come to an end, as accurate outcome prediction of patients remains challenging (§1.2.4). During the past years, the Center for Medical Genetics Ghent, which acts as the Belgian reference center for the genetic diagnosis and study of NB, has set up several projects to establish additional RNA- and DNA-based biomarkers, based on the study of copy number changes, and mRNA and miRNA tumor levels [211, 222, 224, 227]. In this setting, I have explored DNA methylation as an alternative target of prognostic biomarker research in NB.

Up until a few years ago, the NB tumor DNA methylome was relatively unexplored and few prognostic DNA methylation biomarkers were described, as thus far most studies were candidate gene-based (§review paper; §1.2.3; [228]). Therefore, I started genome-wide screening efforts using methyl-CpG-binding domain (MBD) sequencing (§1.1.4.3) to identify new biomarkers.

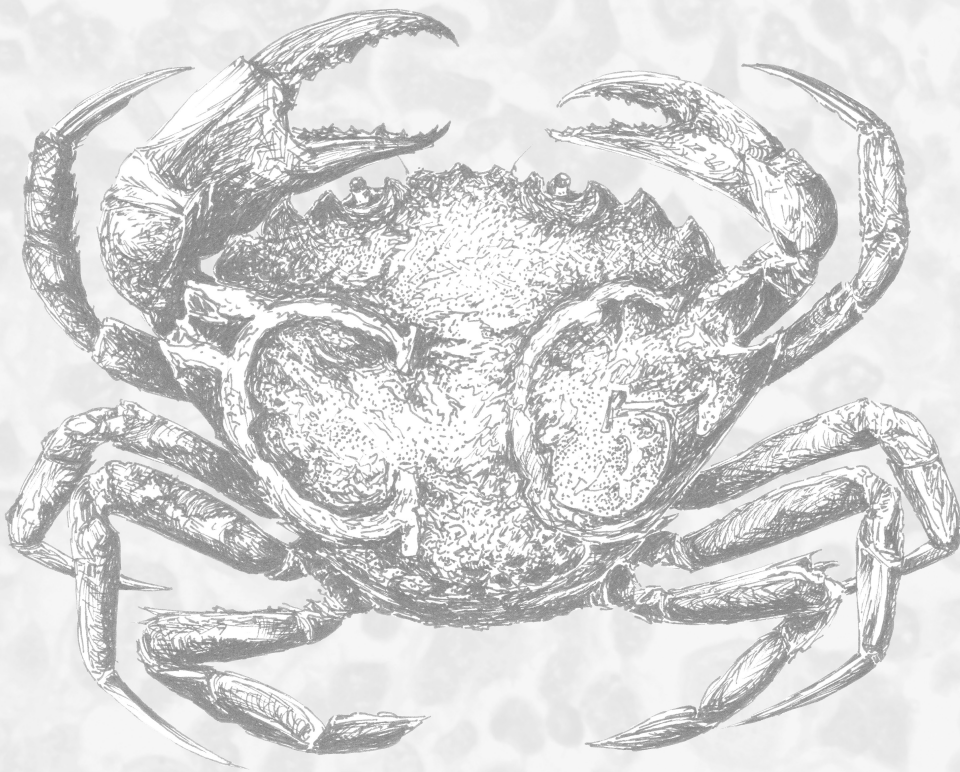
In a first discovery phase, the MBD sequencing workflow was optimized on 8 NB cell lines, whereby through data integration with mRNA expression, candidate biomarkers were selected. To evaluate these candidates, a high-throughput semi-automated methylation-specific PCR (MSP) pipeline was developed and used to test 43 biomarker MSP assays on an independent cohort of 89 primary tumors selected for risk classification and survival. This successfully led to the identification of putative prognostic methylation biomarkers (§paper 1; [229]).

Subsequently, I set up a more extensive discovery by mapping the DNA methylome of 87 primary NB tumors using MBD sequencing. Differential methylation analyses between prognostic patient groups were applied to prioritize new candidate biomarkers. In total, 78 MSP assays were designed for top-ranking differentially methylated regions and tested on two independent cohorts of 132 and 177 primary tumors, respectively. Further, a new statistical framework was developed to identify a robust set of MSP assays of which the percentage of methylated assays allows accurate outcome prediction. As such, multiple prognostic single-gene methylation biomarkers were identified and validated, as well as a prognostic 58-marker methylation signature, predicting overall and event-free survival. This study encompasses the largest DNA methylation biomarker study in NB so far (§paper 2; [230] and §paper 3; [110]).

Along with the 87 primary tumors (stage 1, 2, 3 and 4) of our biomarker validation study, also 15 stage 4S tumors were profiled by MBD sequencing, creating DNA methylation maps of 102 heterogeneous primary NB tumors. This unique resource of methylation information was shared with the NB research community through an open access data descriptor, making the MBD sequencing data easily reusable (§paper 3; [110]).

Finally, I characterized the promoter DNA methylation portrait of stage 4S NB (§1.2.1.3), by performing differential methylation analyses between *MYCN* non-amplified stage 4S, stage 4 and stage 1/2 tumors, using the MBD sequencing data of 41 primary tumors. I showed that specific chromosomal locations are enriched for stage 4S differentially methylated promoters and that stage 4S tumors show characteristic hypermethylation of specific subtelomeric promoters. Additionally, our MBD sequencing data illustrated that important oncogenic pathways, neural crest development and differentiation, and epigenetic processes are differentially regulated in stage 4S tumors. These findings open new avenues for further research, to gain more insights in the NB pathology in general and stage 4S specifically (§paper 4; [231]).

3



Results

paper 1

Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers

Anneleen Decock, Maté Ongenaert, Jasmien Hoebeeck, Katleen De Preter, Gert Van Peer, Wim Van Criekinge, Ruth Ladenstein, Johannes H Schulte, Rosa Noguera, Raymond L Stallings, An Van Damme, Geneviève Laureys, Joëlle Vermeulen, Tom Van Maerken, Frank Speleman and Jo Vandesompele

published in Genome Biology 13(10):R95, 2012

impact factor 2012: 10.288

times cited (d.d. 11/09/2016; Google Scholar): 36

PAPER 1: GENOME-WIDE PROMOTER METHYLATION ANALYSIS IN NEUROBLASTOMA IDENTIFIES PROGNOSTIC METHYLATION BIOMARKERS

Anneleen Decock^{*1}, Maté Ongenaert^{*1}, Jasmien Hoebeek^{1,2}, Katleen De Preter¹, Gert Van Peer¹, Wim Van Criekinge^{3,4,5}, Ruth Ladenstein⁶, Johannes H Schulte⁷, Rosa Noguera⁸, Raymond L Stallings^{9,10}, An Van Damme¹¹, Geneviève Laureys¹², Joëlle Vermeulen¹³, Tom Van Maerken^{1,14}, Frank Speleman¹ and Jo Vandesompele¹

**Contributed equally. Correspondence to joke.vandesompele@ugent.be.*

¹Center for Medical Genetics, Ghent University Hospital, Medical Research Building, De Pintelaan 185, Ghent, Belgium. ²Faculty of Education, Health and Social Work, University College Ghent, Ghent, Belgium. ³Department of Mathematical Modelling, Statistics and Bio-informatics, Ghent University, Coupure Links 653, Ghent, Belgium. ⁴MDxHealth, Tour 5 GIGA, Avenue de l'Hôpital 11, Liège, Belgium. ⁵NXTGNT, Ghent University, De Pintelaan 185, Ghent, Belgium. ⁶Children's Cancer Research Institute, St Anna Kinderkrebsforschung, Zimmermannplatz 10, Vienna, Austria. ⁷University Children's Hospital Essen, Hufelandstraße 55, Essen, Germany. ⁸Department of Pathology, Medical School, University of Valencia, Blasco Ibañez 17, Valencia, Spain. ⁹National Children's Research Centre, Our Lady's Children's Hospital, Crumlin, Dublin 12, Ireland. ¹⁰Department of Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, York House, York Street, Dublin 2, Ireland. ¹¹Department of Pediatrics, Brussels University Hospital, Laarbeeklaan 101, Brussels, Belgium. ¹²Department of Pediatric Hematology and Oncology, Ghent University Hospital, De Pintelaan 185, Ghent, Belgium. ¹³Pédiatrie, Hôpital de Jolimont, Rue Ferrer 159, La Louvière (Haine-Saint-Paul), Belgium. ¹⁴Department of Clinical Chemistry, Microbiology and Immunology, Ghent University Hospital, De Pintelaan 185, Ghent, Belgium.

Contribution of AD: In a joint effort, AD drafted the manuscript and took part in data generation and analysis, including collecting and preparing samples for the high-throughput methylation-specific PCR, testing the candidate biomarkers, performing MSP analyses (testing associations between biomarker methylation and NB risk factors and survival analyses), and analyzing the mRNA expression data.

Abstract

Accurate outcome prediction in neuroblastoma, which is necessary to enable the optimal choice of risk-related therapy, remains a challenge. To improve neuroblastoma patient stratification, this study aimed to identify prognostic tumor DNA methylation biomarkers. To identify genes silenced by promoter methylation, we first applied two independent genome-wide methylation screening methodologies to eight neuroblastoma cell lines. Specifically, we used re-expression profiling upon 5-aza-2'-deoxycytidine (DAC) treatment and massively parallel sequencing after capturing with a methyl-CpG-binding domain (MBD sequencing). Putative methylation markers were selected from DAC-upregulated genes through a literature search and an upfront methylation-specific PCR on 20 primary neuroblastoma tumors, as well as through MBD sequencing in combination with publicly available neuroblastoma tumor gene expression data. This yielded 43 candidate biomarkers that were subsequently tested by high-throughput methylation-specific PCR on an independent cohort of

89 primary neuroblastoma tumors that had been selected for risk classification and survival. Based on this analysis, methylation of *KRT19*, *FAS*, *PRPH*, *CNR1*, *QPCT*, *HIST1H3C*, *ACSS3* and *GRB10* was found to be associated with at least one of the classical risk factors, namely age, stage or *MYCN* status. Importantly, *HIST1H3C* and *GNAS* methylation was associated with overall and/or event-free survival. This study combines two genome-wide methylation discovery methodologies and is the most extensive validation study in neuroblastoma performed thus far. We identified several novel prognostic DNA methylation markers and provide a basis for the development of a DNA methylation-based prognostic classifier in neuroblastoma.

Keywords: DNA methylation, neuroblastoma, biomarker, prognosis, MBD sequencing, methylation-specific PCR (MSP), 5-aza-2'-deoxycytidine (DAC)

Background

Neuroblastoma (NB) is a neuroectodermal tumor that originates from precursor cells of the sympathetic nervous system and represents the most common extra-cranial solid tumor of early childhood. NB displays a highly variable clinical course, ranging from spontaneous regression to life-threatening disease [1].

Despite advances in multimodal anticancer therapies, survival rates for children with aggressive NB remain disappointingly low. Survival rates vary widely, depending on clinical features, such as age at diagnosis and tumor stage, as well as biological characteristics of the tumor. Amongst the latter, *MYCN* amplification has been used for many years as a genetic marker for therapy stratification [1]. More recently, a subset of high-risk tumors with non-amplified *MYCN* and 11q deletions was identified, while absence of segmental aberrations upon genome-wide DNA copy number analysis was found to be associated with excellent survival [2,3]. In order to facilitate the comparison of risk-based clinical trials, a new consensus approach for pretreatment risk classification has been designed including genetic parameters [1,4]. Despite this progress, additional markers for therapeutic stratification are warranted in order to avoid under- or overtreatment and to improve selection of ultra-high-risk patients for new experimental

therapies. Recently, prognostic mRNA and microRNA (miRNA) signatures were developed to accommodate this need [5–7]. Here, we propose that the use of DNA methylation markers is a new and promising method for prognostic classification.

DNA methylation is the addition of a methyl group to carbon 5 of the cytosine within the CpG dinucleotide. Dense clusters of CpG dinucleotides, termed CpG islands, are often present in gene promoters and methylation of those regions typically results in transcriptional silencing of the gene. As such, abnormal DNA methylation in cancer cells leads to aberrant expression patterns [8]. In NB, the most described epigenetic alterations are DNA methylation of *CASP8* [9] and *RASSF1A* [10], both associated with risk factors, such as *MYCN* amplification, age at diagnosis and tumor stage [11–15]. Recently, a few genome-wide methylation screening methodologies have been applied in NB, including re-expression analysis after treatment with 5-aza-2'-deoxycytidine (DAC), DNA methylation promoter arrays after capturing with methylated DNA immunoprecipitation (MeDIP) and methylation microarrays. These studies indicate that aberrant DNA methylation makes an important contribution towards NB tumor biology by downregulating specific genes and show the potential of using DNA methylation in future patient therapy stratification protocols [16–18].

Furthermore, the power of DNA methylation as a non-invasive, sensitive and specific biomarker has been demonstrated by measuring DNA methylation of *RASSF1A* in serum of primary NB patients [15] (for a detailed review see [19]). In order to improve the outcome prediction of NB patients, this study aims at establishing robust DNA methylation biomarkers that can identify patients with unfavorable prognosis.

Results

Discovery and integrated analysis: genome-wide methylation screening for selection of candidate biomarkers

The experimental setup of the study is summarized in Figure 1. In order to identify DNA methylation biomarkers in NB, we first applied two genome-wide methylation screening methodologies on eight NB cell lines: microarray after re-expression analysis and massively parallel sequencing after capturing with a methyl-CpG-binding domain (MBD sequencing). The genome-wide assessment of gene expression reactivation upon DAC treatment is an indirect method to detect DNA methylation as the influence of the demethylating effect is measured at the transcriptional level using oligonucleotide chips. Out of 54,675 probes, a total of 3,624 were upregulated after DAC treatment compared to untreated controls (RankProd false discovery rate (FDR) < 5%), of which 1,665 were upregulated at least twofold in at least one cell line. Using a cutoff of at least a twofold difference between the DAC-treated and the untreated sample, 989 probes were re-expressed in at least 2 cell lines. In order to select specific and sensitive methylation biomarkers from this high number of reactivated probes, an integrated bioinformatics approach was applied. The 1,665 upregulated probes identified by RankProd analysis were further filtered using a genome-wide promoter alignment strategy, referred to as the 'broad

approach' in Hoque et al. [20]. This strategy consists of a genome-wide multiple alignment of promoter regions, where similar sequence regions thus cluster together and where the 'distance' (the number of nodes in the hierarchical alignment model) is shown to be able to predict novel biomarkers. Such approaches using DAC re-expression data have previously been successfully applied to enrich towards truly methylated genes [20,21]. We selected 150 genes that were either in the 'neighborhood' (less than 8 nodes away) of a known methylation marker or that clustered together in the promoter sequence alignment with a high number of reactivation events (at least two genes in the cluster showed at least three reactivation events). Integration with (NB) literature, using an in-house developed text-mining-based approach (using NCBI E-Utils to query PubMed, using all known gene aliases in combination with either DNA methylation-related or NB-related search terms), and selection for genes located in genomic regions reported as recurrently affected by DNA copy number changes in NB, eventually led to the selection of 120 candidate biomarkers, comprising 30 novel candidate markers and 90 known methylation markers in other tumor types. To obtain direct evidence for DNA methylation and to further select prognostic biomarkers, the selected 120 candidate biomarkers were tested on the DAC-treated and untreated NB cell lines CLB-GA, LAN-2, N206, SH-SY5Y and SJNB-1, and primary NB samples (9 low-risk survivors (LR-SURV) and 11 high-risk deceased (HR-DOD) patients; for details see Material and methods), using high-throughput methylation-specific PCR (MSP). In the NB cell lines, the DAC-treated samples show less methylation calls in comparison to untreated samples (130 MSP assays (64%) are more frequently methylated in the untreated samples), and taking all MSP assays into account the average number of methylated samples per assay is 0.39 for the DAC-treated cell lines versus

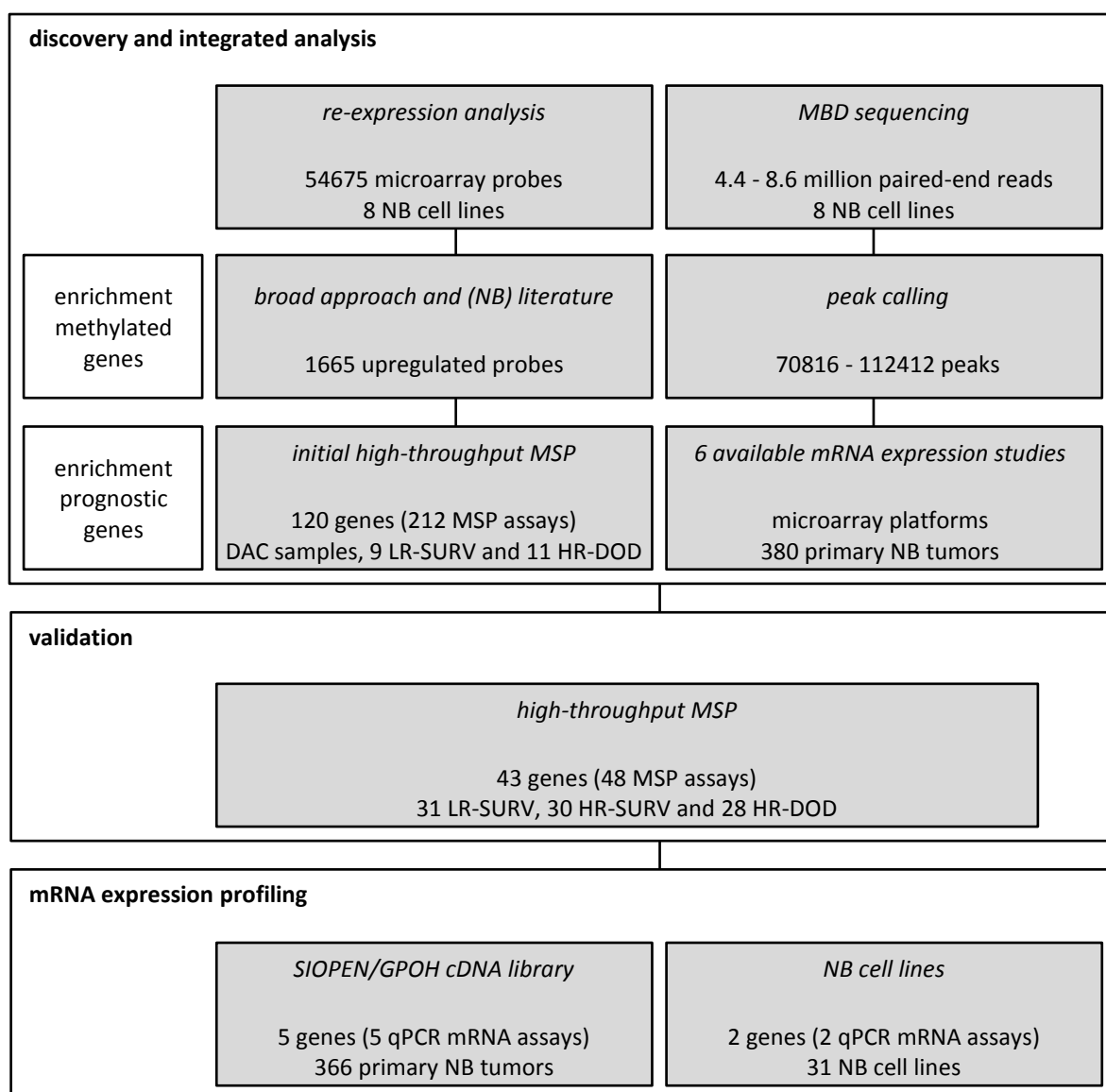


Figure 1. Combining genome-wide methylation discovery and validation, several novel prognostic DNA methylation markers were identified in neuroblastoma (NB). Starting points are a microarray based re-expression study after treatment with 5-aza-2'-deoxycytidine (DAC) and a next-generation sequencing experiment using an enrichment strategy towards methylated DNA (methyl-CpG-binding domain (MBD) capture). Both were performed on the same panel of eight NB cell lines. Applying a bioinformatics and text-mining-based approach on the re-expression data, 120 candidate genes were selected and tested using an initial high-throughput methylation-specific PCR (MSP) screen. The MBD sequencing data were combined with public mRNA expression studies to enrich for potential prognostic biomarkers. Using a rank-based scoring system, a final selection of 43 candidates was made, which were then tested using MSP on 89 primary NB samples (in the following subgroups: HR-DOD: high-risk patients that die of disease; HR-SURV: high-risk patients with long follow-up; LR-SURV: low-risk patients with long follow-up). Finally, mRNA expression levels of seven DNA methylation biomarkers were determined. qPCR: quantitative polymerase chain reaction.

1.47 for the untreated cell lines ($p = 0.0002$), revealing dense methylation in genes upregulated upon DAC treatment and efficient demethylation by DAC (data not shown). The complete results of the initial high-throughput MSP screening on the primary NB samples can

be found in Additional file 1.

The second genome-wide DNA methylation screening methodology we applied, to the same eight NB cell lines, was MBD sequencing: massively parallel sequencing of methylation-enriched DNA fragments, whereby the

enrichment is based on the capture of methylated sheared DNA using the high affinity of the methyl-CpG-binding domain (MBD) of the protein MBD2 towards methylated cytosines. Sequencing yielded 4.4 to 8.6 million paired-end reads, depending on the cell line, and after peak calling 70,816 to 112,412 peaks were detected, representing genomic regions methylated in the corresponding cell line. Between 7,612 and 11,178 of these peaks (around 10% of all identified peaks) are located in promoter regions of annotated genes (-1,500 bp to +1,000 bp around the transcription start site (TSS)). These 'methylation peaks' were visualized in the Integrative Genomic Viewer [22], showing that promoter regions that are well known to be heavily methylated in NB were confirmed - for example, the protocadherin β (*PCDHB*) family cluster (Additional file 2) [23,24]. In some regions (for example, in the promoter regions of *HIST1H3C* and *ACSS3*) it was also possible to distinguish different DNA methylation profiles between *MYCN* amplified (IMR-32, LAN-2 and N206) and *MYCN* non-amplified (SH-SY5Y, SK-N-AS, CLB-GA and SJNB-1) NB cell lines (Additional file 2). Using the R/BioC package DESeq [25], 510 regions were identified as differentially methylated between *MYCN* amplified and non-amplified cell lines, of which 95 are in close proximity to an annotated TSS (-1,500 bp to +1,000 bp). Also, some miRNAs appeared to be methylated in their promoter region.

After peak calling, we also performed gene set enrichment analysis [26], using a custom, ranked list of genes with at least one MBD peak present in a region -1,500 bp to +500 bp around its TSS, in order to explore whether promoter regions that are enriched after MBD capture are often re-expressed as well upon DAC treatment. This analysis clearly showed a high enrichment score for each cell line (enrichment scores from 0.32 to 0.36; FDR q-value < 0.01), demonstrating that a large portion of methylated regions (captured by MBD) are indeed reactivated upon DAC treatment. The overlap between the two

genome-wide datasets can be further explored by intersecting them. In total, 183 genes are both reactivated upon DAC treatment (at least 1 log2 difference after and before treatment) and have an MBD peak in their promoter regions (-1,500 bp to +1,000 bp around the TSS) in at least 2 of the 8 investigated NB cell lines. Of these 183 genes, 46 are both re-expressed and methylated in 3 cell lines, 9 in 4 cell lines and 5 in at least 5 cell lines.

As we feared that only using cell lines in the selection phase of potential prognostic DNA methylation biomarkers would lead to the identification of methylated markers not necessarily related to prognosis, six publicly available mRNA expression studies [27–34] were included in the analysis. In these studies, which comprise mRNA expression data of 380 primary NB tumors, identifying differentially expressed probes (genes) between prognostic groups would allow us to pinpoint potential prognostic methylated promoter regions in our methylome maps. Finally, a rank-based scoring system was used to prioritize genes that show methylation, re-expression after DAC treatment and differential expression (related to risk) across the prognostic groups. This score scheme uses the individual ranks of each analysis. In brief, DAC reactivation is ranked according to FDR rate (as determined by RankProd analysis), MBD sequencing data are ranked according to peak p-values and expression data are ranked according to FDR (determined by RankProd analysis). Each data source is given the same weight and a combined rank is calculated (for details, see Materials and methods). This scoring system combined all generated data and allowed us to select 43 top-ranking and thus strong prognostic methylation candidate genes without the need to use rather artificial threshold values for the different datasets.

Validation: determining the prognostic power of DNA methylation biomarkers

For these 43 genes, 48 MSP assays were designed and tested on 3 NB cell lines (IMR-32, SK-N-AS and SH-SY5Y) and the HCT-116 DKO cell lines, along with an independent cohort of 89 primary NB samples. Within the 89 primary NB sample set all three prognostic groups (LR-SURV, HR-DOD and high-risk survivors (HR-SURV); for details see Material and methods) were approximately equally represented. The complete matrix with all MSP results of all samples and a global overview of the MSP results per assay can be found in Additional file 3. Over 60% of the designed assays indeed detected methylation for the respective marker in at least 10% of the selected NB tumors. Ten MSP assays (*COL6A3*, *miR-1225*, *miR-3177*, *PCDHA6*, *PLXNC1*, *ANKRD43*, *ADRB2*, *APOE*, *miR-671* and *QPCT*) revealed methylation in at least 75% of the patient samples, and the MSP assays for *KCND2*, *PRPH*, *KRT19* (assay 83159) and *TNFRSF10D* were methylated in 50% to 75% of the patient samples. We could also detect DNA methylation in the promoter region of *miR-1225*, *miR-3177*, *miR-671* and *miR-663*, methylated in 99%, 99%, 79% and 4% of the patient samples, respectively.

Unique in this study is the use of three discrete prognostic patient groups, which allowed us to assess differential methylation across all these prognostic groups. Therefore, we performed hierarchical cluster analysis on the methylation data of all 48 MSP assays on the entire NB tumor cohort, revealing two clusters with a separation between high-risk (HR) and low-risk (LR) patients (heatmap in Additional file 3). Furthermore, the overall methylation pattern in the primary NB tumor samples was compared by calculating the number of methylation events for each sample. This indicates that HR patients show, on average, more methylation events compared to LR patients ($p < 0.001$; HR-DOD, 17.21 methylation events (95% confidence interval (CI) 15.62 to 18.81); HR-SURV, 17.13 methylation events (95% CI 15.81 to 18.46); LR-SURV, 13.00 methylation events (95% CI 11.86 to 14.14)). Also on the

individual marker level, some MSP assays are differentially methylated across the prognostic patient groups: *KRT19* and *ACSS3*. These genes are more frequently methylated in HR patients compared to LR patients (Table 1). Within the HR group, *HIST1H3C* shows a tendency to be more frequently methylated in HR-DOD compared to HR-SURV samples (21% in HR-DOD versus 7% in HR-SURV), while *KRT19* (32% versus 48%) and *ACSS3* (25% versus 47%) show the inverse pattern.

Some individual MSP assays were also associated with one or more NB risk factors (stage, *MYCN* status and age at diagnosis), and are thus potential prognostic biomarkers in NB (Table 1). In this analysis, the age at diagnosis was tested using two different age cutoffs. The 12 months cutoff was chosen as it was used for therapy stratification and as a criterion in the sample selection. The more recently established cutoff of 18 months [1,35,36] was also taken into account. Newly discovered methylated markers are *FAS*, *PRPH*, *CNR1*, *QPCT*, *HIST1H3C*, *ACSS3* and *GRB10*, methylation of which is associated with at least one of the NB risk factors. Table 1 further indicates that the difference in the methylation status of *HIST1H3C* and *ACSS3* between *MYCN* non-amplified and amplified NB cell lines as detected by MBD sequencing is reflected in the MSP results of the primary tumors as well, as *HIST1H3C* and *ACSS3* are almost exclusively methylated in *MYCN* amplified samples.

Survival analysis using the complete MSP data set indicates that patients with less methylation events showed better survival rates than patients with a high number of methylation events ($p = 0.01$; Additional file 3), as this analysis principally discriminates HR and LR patients. In order to assess to what extent our MSP data set is able to predict overall survival (OS) in HR-SURV versus HR-DOD patients, leave-one-out decision tree analysis was performed and repeated 58 times (the number of HR patients). For this analysis, we only included the

Table 1. Several individual markers are differentially methylated between the prognostic groups and neuroblastoma risk factors.

type	subtype	<i>KRT19</i>		<i>FAS</i>		<i>PRPH</i>		<i>CNR1</i>		<i>QPCT</i>		<i>HIST1H3C</i>		<i>ACSS3</i>		<i>GRB10</i>	
		number	%	number	%	number	%	number	%	number	%	number	%	number	%	number	%
prognostic group	LR-SURV	0/31	0	1/31	3	14/31	45	2/31	6	18/31	58	0/31	0	0/31	0	6/31	19
	HR-SURV	14/30	48	8/30	27	24/30	80	10/30	33	25/30	83	2/30	7	14/30	47	13/30	43
	HR-DOD	9/28	32	6/28	21	19/28	68	10/28	36	24/28	86	6/28	21	7/28	25	11/28	39
INSS stage	stage 1	0/21	0	0/21	0	8/21	38	1/21	5	13/21	62	0/21	0	0/21	0	4/21	19
	stage 2	1/12	8	2/12	17	8/12	67	2/12	17	8/12	67	0/12	0	1/12	8	3/12	25
	stage 3	9/17	53	4/17	24	13/17	77	9/17	53	15/17	88	3/17	18	9/17	58	8/17	47
	stage 4	13/39	33	9/39	15	28/39	72	10/39	26	31/39	80	5/39	18	11/39	28	15/39	39
MYCN amplification status	MYCN non-amplified	7/50	14	2/50	4	24/50	48	5/50	10	31/50	62	0/50	0	2/50	4	14/50	28
	MYCN amplified	16/39	41	13/39	33	33/39	85	17/39	44	36/39	92	8/39	21	19/39	49	16/39	41
age at diagnosis	> 12 months	21/53	40	14/53	26	37/53	69	18/53	34	46/53	87	8/53	15	21/53	40	24/53	45
	< 12 months	2/36	6	1/36	3	20/36	56	4/36	11	21/36	58	0/36	0	0/36	0	6/36	17
	> 18 months	20/45	44	13/45	29	33/45	73	17/45	38	40/45	89	8/45	18	19/45	49	23/45	51
	< 18 months	3/44	7	2/44	5	24/44	55	5/44	11	27/44	61	0/44	0	2/44	5	7/44	16
overall total		23/89	26	15/89	17	57/89	64	22/89	25	67/89	75	8/89	9	21/89	24	30/89	34
statistics on type (Fisher's exact p-value)																	
		<i>KRT19</i>		<i>FAS</i>		<i>PRPH</i>		<i>CNR1</i>		<i>QPCT</i>		<i>HIST1H3C</i>		<i>ACSS3</i>		<i>GRB10</i>	
prognostic group		<0.001		0.151		0.112		0.068		0.165		0.0624		<0.001		0.405	
MYCN amplification status		0.0594		0.008		0.008		0.008		0.017		0.0146		<0.001		0.708	
INSS stage		0.007		0.287		0.221		0.059		0.683		0.448		0.008		0.700	
age cutoff 12 months		0.008		0.045		0.579		0.138		0.059		0.123		<0.001		0.059	
age cutoff 18 months		0.002		0.045		0.326		0.059		0.045		0.0594		0.0015		0.012	

Note. The number (percentage) of methylated samples in each stratum is given. P-values according to Fisher's exact test, corrected for multiple testing (Benjamini-Hochberg). HR-DOD: high-risk deceased patients; HR-SURV: high-risk patients alive for at least 1,000 days follow-up; INSS: International Neuroblastoma Staging System; LR-SURV: low-risk patients alive for at least 1,000 days follow-up. P-values in bold indicate significant associations.

data from MSP assays (un)methylated in at least three samples. Comparison of the 58 generate decision trees showed that 4 DNA methylation biomarkers (*CNR1*, *ACSS3*, *HIST1H3C* and *PRPH*) are included in at least 50% of the resulting classifiers. Then, leave-one-out decision tree analysis was redone, but this time using only the methylation data of *CNR1*, *ACSS3*, *HIST1H3C* and

PRPH. Afterwards, the predictions for all 58 HR samples were visualized in a Kaplan-Meier plot (Figure 2). This analysis indicates that the combined methylation status of *CNR1*, *ACSS3*, *HIST1H3C* and *PRPH* has the potential to discriminate between HR-SURV and HR-DOD patients ($p = 0.058$).

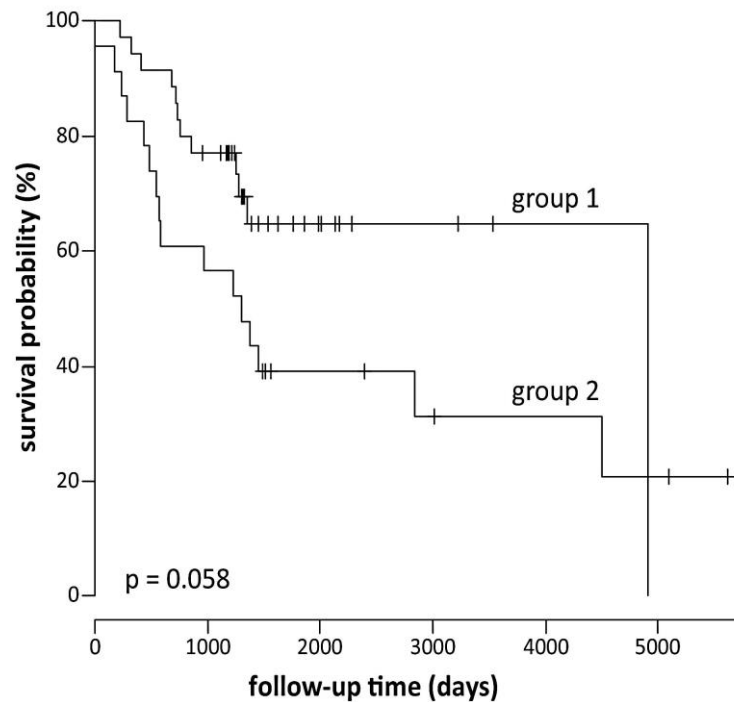


Figure 2. The combined methylation status of *CNR1*, *ACSS3*, *HIST1H3C* and *PRPH* can potentially discriminate high-risk patients. The Kaplan-Meier plot shows overall survival in the high-risk samples of the high-throughput MSP screening according to their predicted overall survival status based on leave-on-out decision tree analysis using the methylation data of *CNR1*, *ACSS3*, *HIST1H3C* and *PRPH*. Group 1 is predicted to survive, group 2 to die of disease. The p-value is determined using log-rank test (Mantel-Cox).

Survival analysis was also performed on the individual marker level. We first tested differences between the HR-DOD and LR-SURV groups using the univariate log-rank test (with multiple testing correction). This first analysis indicates that six genes (*KRT19*, *FAS*, *CNR1*, *HIST1H3C*, *ACSS3* and *GNAS*) are significantly related to survival when comparing these patient groups. As we also want to discriminate the HR patient groups (HR-DOD and HR-SURV), we then used the entire dataset (all samples) to assess which of these six genes were associated with survival (in a specific stratum only, such as

only in *MYCN* non-amplified samples). These results are shown in Table 2. According to log-rank tests, *HIST1H3C* methylation is associated with both OS and event-free survival (EFS), while *GNAS* methylation is associated with EFS. As NB is a heterogeneous disease, these biomarkers may be suited to a specific subgroup of patients for predicting survival. For example, *HIST1H3C* methylation only occurs in high-stage tumors with *MYCN* amplification (6/17 (35%) in HR-DOD patients versus 2/22 (9%) in HR-SURV patients). Figure 3 shows the Kaplan-Meier plots for *HIST1H3C* and *GNAS* methylation (OS or EFS and

OS in specific strata related to one of the risk factors).

mRNA expression profiling: determining transcriptional silencing of DNA methylation biomarkers

As it is known that promoter methylation may cause transcriptional silencing of the gene, we further measured the mRNA expression levels of five promising DNA methylation biomarkers that were methylated in a substantial fraction of HR patients (*CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT*). Quantitative RT-PCR assays were developed and tested on 366 primary NB tumor samples. Table 3 displays the results of the comparisons of the expression levels of each DNA methylation biomarker between the different NB tumor

stages, *MYCN* non-amplified and amplified tumors, the two age groups (using both the 12 and 18 months cutoff), and surviving and deceased patients. As an example, the mRNA expression levels of these genes across the NB tumor stages are depicted in Additional file 4. Out of the 366 primary NB tumors, 245 could be assigned to one of the prognostic groups defined in this study (Additional file 4), which allowed us to assess differential mRNA expression between these groups as well. For all genes mRNA expression levels were significantly higher in the LR group compared to the HR groups. As methylation of these genes was mainly detected in the HR groups, this suggests that methylation may contribute to the transcriptional silencing of these genes.

Table 2. Several individual DNA methylation markers are associated with survival.

statistics	<i>KRT19</i>	<i>FAS</i>	<i>CNR1</i>	<i>HIST1H3C</i>	<i>ACSS3</i>	<i>GNAS</i>
HR-DOD versus LR-SURV (p-value OS)	0.037	0.028	0.043	0.002	0.002	0.012
HR-DOD versus LR-SURV (p-value EFS)	0.039	0.049	0.039	0.039	0.079	0.039
HR-DOD versus LR-SURV and HR-SURV (p-value OS)	0.687	0.639	0.423	0.039	0.691	0.221
HR-DOD versus LR-SURV and HR-SURV (p-value EFS)	0.665	0.467	0.414	0.041	0.939	0.041
HR-DOD versus LR-SURV and HR-SURV (p-value OS (stratum))			age < 12 months 0.035	stage 4 0.033		MYCN0 0.033
HR-DOD versus LR-SURV and HR-SURV (p-value EFS (stratum))		age < 12 months 0.014				MYCN0 0.001

Note. Log-rank test statistics (Mantel-Cox) are given (multiple testing correction by Benjamini-Hochberg) for comparison between the ultra-high-risk group (HR-DOD) versus the low-risk group (LR-SURV), and between the ultra-high-risk group (HR-DOD) and all survivors (LR-SURV and HR-SURV). If a significant association ($p < 0.05$) was found in a particular stratum (associated with risk factors), this stratum is shown (multiple testing correction for the different comparisons by Benjamini-Hochberg). EFS: event-free survival; MYCN0: *MYCN* non-amplified; OS: overall survival. P-values in bold indicate significant associations.

Table 3. The mRNA expression level of several markers associates with neuroblastoma risk factors, prognostic groups and survival.

grouping variable	statistics	<i>CNR1</i>	<i>GRB10</i>	<i>KRT19</i>	<i>PRPH</i>	<i>QPCT</i>
stage	Kruskal-Wallis p-value	<0.001	0.008	0.118	0.010	<0.001
<i>MYCN</i> amplification status	Mann-Whitney p-value	<0.001	<0.001	<0.001	<0.001	<0.001
age cutoff 12 months	Mann-Whitney p-value	<0.001	0.609	<0.001	0.005	<0.001
age cutoff 18 months	Mann-Whitney p-value	<0.001	0.810	<0.001	0.003	0.006
overall survival status	Mann-Whitney p-value	<0.001	0.003	0.023	<0.001	<0.001
prognostic group	Kruskal-Wallis p-value	<0.001	0.002	0.005	<0.001	<0.001

The statistical test used is shown and p-values (corrected for multiple testing using Benjamini-Hochberg) are indicated. P-values in bold indicate significant associations.

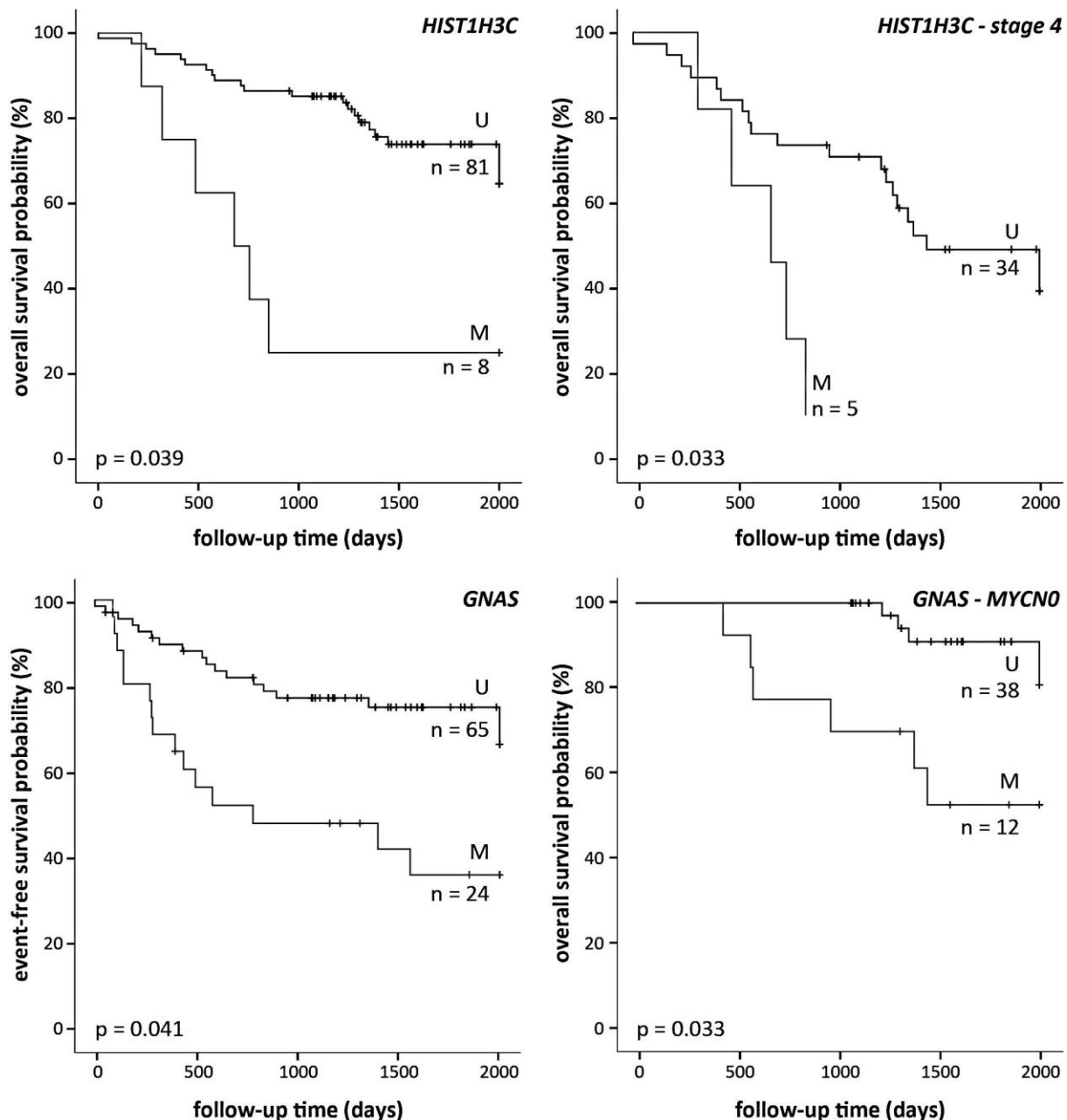


Figure 3. Methylation of *HIST1H3C* and *GNAS* is associated with worse survival outcome. Kaplan-Meier plots on the left show overall survival or event-free survival for all 89 primary neuroblastoma samples, those on the right overall survival in a specific stratum based on one of the risk factors only. Survival curves indicated with 'M' are the methylated samples, survival curves associated with the unmethylated assay are indicated with 'U'. The numbers of patients are indicated (n) and p-values are determined using log-rank test (Mantel-Cox; multiple testing correction by Benjamini-Hochberg). Time is censored to 2,000 days. MYCN0: *MYCN* non-amplified.

Survival analysis using Cox proportional hazards further shows that low mRNA expression levels of *CNR1* (hazard ratio 0.768; 95% CI 0.619 to 0.953; $p = 0.028$), *GRB10* (hazard ratio 0.613; 95% CI 0.433 to 0.866; $p = 0.015$) and *PRPH* (hazard ratio 0.714; 95% CI 0.566 to 0.922; $p = 0.015$) were significantly associated with poor

survival. After dichotomization of the mRNA expression data, using the median relative mRNA expression value as a cutoff, Kaplan-Meier survival curves were plotted (log-rank test; Additional file 4).

An interesting observation in our MBD sequencing and MSP data is the fact that

HIST1H3C and *ACSS3* are differentially methylated between *MYCN* non-amplified and amplified NB cell lines and primary tumors (Table 1; Additional file 2). To further explore this finding, the *HIST1H3C* and *ACSS3* MSP assays were tested on 31 NB cell lines, of which 10 were *MYCN* non-amplified and 21 *MYCN* amplified (Additional file 5). In addition, we also profiled *HIST1H3C* and *ACSS3* mRNA expression levels in these cell lines, in order to assess the direct relationship between promoter methylation and mRNA expression and to compare this relationship between *MYCN* non-amplified and amplified cell lines. The significant differential methylation status of *HIST1H3C* and *ACSS3* between *MYCN* non-amplified and amplified samples was confirmed in the NB cell lines (*HIST1H3C*, methylated in 15/21 (71%)

MYCN amplified cell lines and in 2/10 (20%) *MYCN* non-amplified cell lines, $p = 0.018$; *ACSS3*, methylated in 20/21 (95%) *MYCN* amplified cell lines and in 3/10 (30%) *MYCN* non-amplified cell lines, $p < 0.001$). Moreover, expression of *HIST1H3C* mRNA was significantly lower in methylated samples compared to unmethylated samples, both in *MYCN* amplified ($p = 0.005$) and *MYCN* non-amplified ($p = 0.044$) cell lines (Figure 4). These data support the idea that *HIST1H3C* promoter methylation contributes to transcriptional silencing of the gene. Figure 4 further indicates that the *MYCN* status itself is not significantly associated with *HIST1H3C* mRNA expression levels ($p = 0.204$). As *ACSS3* is expressed at very low mRNA levels, we could not correlate its mRNA expression data with the methylation data (data not shown).

MYCN amplification status - methylation call

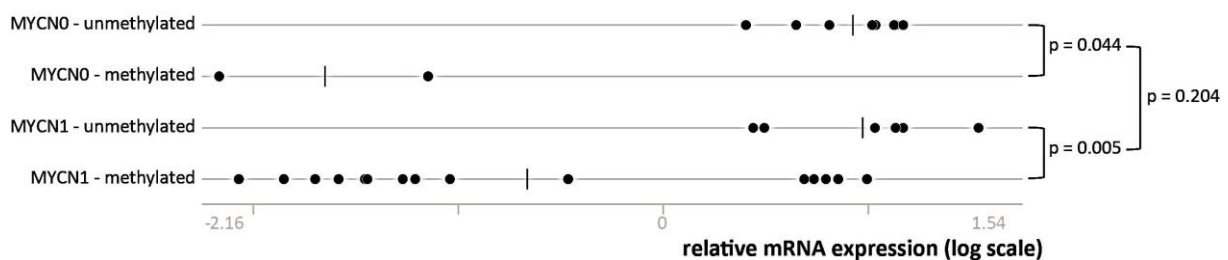


Figure 4. *HIST1H3C* has lower mRNA expression levels in neuroblastoma cell lines in which the *HIST1H3C* promoter is methylated. Thirty-one neuroblastoma cell lines were categorized according to their *MYCN* amplification and *HIST1H3C* methylation status. The relative *HIST1H3C* mRNA expression level of each of these cell lines is indicated. P-values according to Mann-Whitney test are also indicated. MYCN0: *MYCN* non-amplified; MYCN1: *MYCN* amplified.

Discussion

Thus far, most of the studies analyzing DNA methylation patterns in NB have been candidate gene-based, with the methylation status of the promoter region for only a limited number of genes being tested. These candidate genes were selected based either on prior knowledge of NB tumor biology or on the fact of being methylated in other tumor types. As a consequence, only few DNA methylation biomarkers, such as *KRT19*, *TNFRSF10D*, *CASP8*, *ZMYND10* and *RASSF1A*, were previously related to NB risk

factors or survival [11,13–15,18,37–41]. In order to identify new DNA methylation biomarkers in NB, we applied a multilevel experimental approach. In the discovery phase we established a genome-wide methylome map of eight NB cell lines. These cell lines were profiled using gene expression microarrays before and after DAC treatment, and using MBD capture followed by next-generation sequencing (NGS). The combination of both methodologies enabled the identification of regions that are both methylated and undergo re-expression upon DAC treatment. So far, only MeDIP chips were

used in whole promoter profiling studies on NB [9], making this study the first one using NGS for unbiased and more sensitive assessment of genome-wide DNA methylation patterns in NB. Our results emphasize the potential of this epigenetic sequencing technique, as it enables the investigation of the methylome or epigenome of a sample in great detail at a feasible cost.

Integration of these methylome maps with genome-wide gene expression profiles led to a selection of 43 candidate biomarkers that were tested on 89 primary NB patient samples. All samples were assigned to one of three discrete prognostic patient groups (low-risk survivors (LR-SURV), high-risk deceased (HR-DOD) and high-risk survivors (HR-SURV)). While most NB methylation studies did not discriminate between HR-SURV and HR-DOD patients, we believe this is an important clinical question, as both prognostic groups are currently considered high-risk and uniformly treated, making the present study unique in its concept. As we make use of amplified bisulfite-converted DNA, only limited amounts (100 to 200 ng) of tumor DNA are required to test over 100 MSP assays. The MBD sequencing results greatly help in designing the assays in the most informative regions, which is important as the assay location is critically important, again confirmed in this study for a number of genes for which multiple assays were designed (for example, *TGFBI* and *KRT19*). The combination of the number of samples and assays used in this study further makes it the most comprehensive methylation study in NB. Furthermore, the high-throughput validation pipeline allows fast and accurate follow-up validation of potential candidate DNA methylation biomarkers for large numbers of patients. Indeed, PCR-based detection methodologies are robust and can thus be used in a wide range of laboratory settings for a low price without the need of special equipment other than for qPCR and (microfluidic) electrophoresis, both present in most molecular

laboratories. The presented DNA methylation screening and validation methodology can thus easily be adapted by (cancer) researchers addressing similar questions in other research fields.

In this study, several novel biomarkers were established in addition to known DNA methylation biomarkers in NB, such as *KRT19*, *TGFBI*, *TNFRSF10D* and *TNFRSF10A* [14,18,37,42,43]. Interestingly, some of these novel genes were previously reported to be important in NB biology (without reference to their epigenetically altered status) or were described as epigenetic biomarkers in other tumor entities, such as *FAS*, which encodes a member of the tumor necrosis factor receptor (TNFR) superfamily [44–50]. Several other novel methylation biomarkers were also shown to be differentially methylated between HR and LR patients, and many of these were associated with NB risk factors or with survival. However, discriminating HR-DOD and HR-SURV patients is challenging. While only a few individual MSP designs (*HIST1H3C*, *KRT19* and *ACSS3*) were moderately discriminatory between these two HR subgroups, the combined methylation data analysis of *CNR1*, *ACSS3*, *HIST1H3C* and *PRPH* indicates the potential of DNA methylation biomarkers in stratifying HR NB patients. In this study, the difficulty of identifying individual biomarkers that differentiate between HR-DOD and HR-SURV patients may be explained by the fact that NB cell lines were used in the discovery phase, thus enriching for genes discriminating between HR and LR patients as NB cell lines can be considered models for aggressive HR tumors. To accommodate this, we plan to perform a large-scale discovery using MBD capture followed by NGS on primary NB tumors equally distributed over the three prognostic groups used here.

PRPH is one of the novel biomarkers identified and is differentially methylated across the prognostic patient groups. This gene encodes the cytoskeletal protein peripherin found in

neurons of the peripheral nervous system, and is probably associated with maturation of the neural phenotype and hence serves as a differentiation marker for tumors derived from the neural crest [51]. In our study, *PRPH* methylation was mainly detected in more advanced tumor stages. Since promoter methylation may cause transcriptional silencing of the gene and advanced NB tumor stages are less differentiated [52], this is in line with the idea that high levels of peripherin contribute to more differentiated tumor stages. As demonstrated in this study, this idea is further strengthened by the fact that *PRPH* mRNA expression levels gradually decreased with increasing aggressiveness of the tumor. As whole genome sequence analysis recently showed that genes involved in neuritogenesis are recurrently affected in high-stage NB [53], the identification of *PRPH* methylation opens new research perspectives regarding NB therapy.

Next to protein-coding genes, some MSP assays were designed in the promoter region of miRNAs. Aberrant miRNA expression contributes majorly to NB tumor biology and has been extensively studied during the past few years. Most of these studies used miRNA microarrays or high-throughput RT-qPCR to analyze the miRNA expression patterns in primary NB tumor samples [54–56]. Although a broad deregulation of the miRNA expression profile in NB has been described, miRNA promoter hypermethylation is relatively unexplored. Up until now, the only miRNA for which the promoter region is known to be methylated in NB is *miR-200b* [57]. Interestingly, *miR-1225*, *miR-3177* and *miR-671* were found to be methylated in their promoter region in more than 75% of the NB tumors in our study. Currently, little is known about the putative function of these miRNAs, as they are not well described or not described at all in the literature [58,59].

Another interesting finding is that *MYCN* non-amplified and amplified samples show differential promoter methylation of *HIST1H3C*

and *ACSS3*. Currently, little is known about the association between *MYCN* and DNA methylation of certain genes in NB, nor about the underlying molecular mechanisms. Previously, Teitz et al. [9,60] showed that DNA methylation of *CASP8* is almost exclusively associated with *MYCN* amplification in both NB cell lines and primary tumors. They further noticed that *CASP8* was hemimethylated (only one allele) in stage 1, 2 and 3 NB, which may indicate that complete methylation of *CASP8* may be coupled to another event, such as amplification of the *MYCN* gene. While this suggests that *MYCN* amplification is functionally linked to complete methylation of both *CASP8* alleles, it is not clear if these two events occur concurrently, or if one event leads to the other. Obviously, genes differentially methylated between *MYCN* non-amplified and amplified samples need to be further functionally characterized, as this may lead to new insights into NB biology.

Conclusions

Although international collaboration in the field of NB has invested tremendous effort in optimizing patient stratification and therapy protocols, OS rates remain low. This study shows that DNA methylation biomarkers have the potential to refine current risk assessment schemes. In contrast to most NB methylation studies that are candidate gene-based, we applied two genome-wide detection methodologies to discover hypermethylated regions in NB: re-expression analysis after demethylating DAC treatment and NGS after MBD capture. Furthermore, we present a high-throughput and semi-automated MSP pipeline, which was used to test the candidate DNA methylation markers on a large patient tumor cohort. We have identified novel aberrant promoter hypermethylation of protein coding genes and miRNAs in NB. Some of these DNA methylation biomarkers are associated with NB

risk factors and/or survival, emphasizing the prognostic value of these markers and their potential to be used in a DNA methylation-based prognostic classifier in NB. The use of such a DNA methylation signature, discriminating HR patients, is demonstrated here by the combined methylation data analysis of *CNR1*, *ACSS3*, *HIST1H3C* and *PRPH*. Furthermore, some DNA methylation biomarkers showed low levels of mRNA expression in patient groups with high methylation levels. This suggests that promoter methylation may contribute to transcriptional silencing of these genes, which may be important in the pathogenesis of NB. Encouraged by these results, we will now extensively further validate these DNA methylation biomarkers and refine the methylome map of different prognostic NB patient groups.

Materials and methods

Neuroblastoma cell lines and primary tumors

In total, 33 well-characterized NB cell lines, authenticated using array comparative genomic hybridization and short tandem repeat genotyping, were included in this study (Additional files 2 and 5). DNA was isolated using the QIAamp DNA Mini Kit (Qiagen, Venlo, The Netherlands). In addition, 109 primary tumor samples of NB patients were collected prior to therapy at the Ghent University Hospital (Ghent, Belgium), the University Children's Hospital Essen (Essen, Germany), Our Lady's Children's Hospital Dublin (Dublin, Ireland) or the Hospital Clínico Universitario (Valencia, Spain). Informed consent was obtained from each patient's guardian and the study was approved by the ethical committee of the Ghent University Hospital (approval number B67020109912). Clinical characteristics of the patients are shown in Additional files 1 and 3. All NB patient samples were assigned to one of three defined risk groups based on risk parameters (tumor stage,

MYCN status and age at diagnosis) and disease outcome. First, HR patients that died of disease (HR-DOD) as defined by stage 2/3, *MYCN* amplified, DOD; stage 4, age at diagnosis < 12 months, *MYCN* amplified, DOD; or stage 4, age at diagnosis > 12 months, DOD (n = 39). Second, HR patients alive (HR-SURV) after follow-up time > 1,000 days (n = 30). Third, LR patients alive (LR-SURV) defined by stage 1/2, *MYCN* non-amplified, follow-up time > 1,000 days; stage 3, *MYCN* non-amplified, age < 12 months, follow-up time > 1,000 days (status at last known follow-up is alive; n = 40). The clinical data of the 366 primary NB tumors (SIOPEN/GPOH cDNA library [6]), used to test the mRNA expression levels of the most promising DNA methylation biomarkers, can be found in Additional file 4.

Microarray after re-expression analysis

Eight NB cell lines (CHP-902R, CLB-GA, IMR-32, LAN-2, N206, SH-SY5Y, SK-N-AS and SJNB-1) were grown in the presence of 3 μ M DAC (Sigma, Bornem, Belgium) for 3 days, as previously described, and untreated controls were also prepared [61]. After harvesting, RNA was extracted with the RNeasy Mini kit (Qiagen), accompanied by RNase free DNase treatment on column (Qiagen). After RNA quality check on the Experion (Bio-Rad, Nazareth, Belgium), sample preparation, hybridization to Affymetrix Human Genome U133 Plus 2.0 oligonucleotide chips and scanning were carried out according to the manufacturer's protocol at the VIB MicroArray Facility. Standard quality metrics (simpleaffy BioC package [62] boxplots, visual inspection of the slides, 5'-3' degradation plot) demonstrated that the oligonucleotide chip data were of good quality. The BioC affy package was used to normalize (gc-RMA normalization) the expression levels and to obtain present/absent (expression/no expression) MAS 5.0 calls for each probe set. For all cell lines and for each probe set, the number of reactivation events was counted (absent in untreated cells and

present in treated cells). Expression data (before and after DAC treatment) have been deposited into the Gene Expression Omnibus [GEO: GSE31229], according to the MIAME guidelines.

MBD sequencing

DNA samples (1 µg DNA) of the eight NB cell lines were sheared (Covaris S2) to an average length of 200 bp. Fragment distribution was determined by the Agilent 2100 Bioanalyzer and the concentration was determined using the Quant-iT PicoGreen dsDNA HS Assay Kit (Invitrogen, Ghent, Belgium). Starting from 200 ng sheared DNA, the MethylCollector Kit (ActiveMotif, La Hulpe, Belgium) was used to enrich for methylated fragments. Library preparation for multiplex Illumina sequencing was done by combining the DNA Sample Prep Master Mix Set 1 (New England Biolabs, Frankfurt am Main, Germany) and the Multiplexing Sample Preparation Oligo Kit (Illumina). Size selection of the library was done on a 2% agarose gel. Fragments of around 300 bp (± 50 bp) were excised and purified. Illumina library amplification (21 cycles) was performed and concentration was determined. Paired-end sequencing was used for high confidence mapping of captured fragments (2×45 bp sequencing - Illumina GAIIx, NEXTGNT). Paired-end reads were mapped on the human reference genome (GRCh37) using Bowtie 0.12.7 and peaks were called using MACS 1.4beta. For differential methylation analysis, PCR duplicates were removed and sequence tags counted by using the BioC packages Short-Read and rtracklayer [63,64]. Sequence tag counts per sample were used to compose a count matrix that could be processed by the BioC package DESeq [25]. Sequencing data (raw sequence files, WIG files for visualization of the mapping results and the BED peak files as determined by MACS) have been deposited into GEO [GEO:GSE31353].

Selection of candidate biomarkers

Initial high-throughput MSP

In total, 212 MSP assays (Additional file 1) were designed in the promoter region of 120 corresponding genes re-expressed after DAC treatment, and tested on both the DAC-treated and untreated NB cell lines, 9 LR-SURV patients and 11 HR-DOD patients (Additional file 1). A total of 500 to 1,000 ng DNA of these samples was bisulfite-treated (EZ DNA Methylation Kit, Zymo Research, Irvine, CA, USA), eluted in 30 µl elution buffer and then tested on the BioTrove OpenArray (Life Technologies, Ghent, Belgium). Beta actin (ACTB) was used as a control and to normalize samples. The *in vitro* methylated HCT-116 DKO cell line (treated with SssI, Zymo Research) was used as a positive control. The methylation status for each MSP assay was determined, and called methylated if the melting temperature (T_m) of the amplicon was within a specific interval as defined by the positive control sample. These methylation calls were further analyzed by determining specificity and sensitivity of the HR-DOD samples versus LR-SURV samples.

Publicly available mRNA expression studies

Six publicly available mRNA expression studies [27–34] [GEO:GSE19274, GEO:GSE16237, GEO:GSE14880, GEO:GSE12460, GEO:GSE13136, GEO:GSE3960] were analyzed using RankProd analysis (BioC package [18]), to identify differentially expressed probes between prognostic groups (high-risk versus low-risk, high-stage versus low-stage, and *MYCN* amplified versus *MYCN* non-amplified).

Scoring system

Each analysis score of a promoter region (for example, RankProd FDR-value and p-value for differential expression between risk groups, and p-values of the peak after MBD sequencing) was ranked and given a score, ranging from tan (1) to 0 according to their rank. These individual scores were then summed and 43 top-ranking genes were selected for further analysis.

High-throughput MSP

MSP assays were designed to only amplify the bisulfite-converted target region of interest and do not anneal to genomic DNA. As each primer contains at least two CpG sites, this means that a PCR product will only be generated if the template is methylated. We choose not to design the according U primers (that would amplify the non-methylated bisulfite-converted DNA) as we do not assess methylation in a quantitative way. After *in silico* assay evaluation, 48 selected MSP primers (including the ACTB control; Additional file 3) were empirically validated on the Roche LightCycler 480 (LC480) using the *in vitro* methylated HCT-116 DKO (positive control), the HCT-116 DKO (negative control) and NB cell lines. Based on melting curve and amplicon size analysis, all assays were considered amplicon-specific. The MSP assays were tested on 89 samples, selected from the previously described patient groups (31 LR-SURV patients, 28 HR-DOD patients and 30 HR-SURV patients; Additional file 3). A no template control (NTC) sample was loaded as well. For all samples, 500 to 1,000 ng DNA was bisulfite-treated (EZ DNA Methylation Kit, Zymo Research) and eluted in 40 µl elution buffer. Prior to MSP, bisulfite-treated DNA (BT-DNA) was amplified using the EpiTect Whole Bisulfite Kit (Qiagen), starting from 100 ng BT-DNA. After amplification, the yield was determined by the Qubit 2.0 fluorometer in combination with the Quant-iT PicoGreen dsDNA BR Assay Kit (Invitrogen). The MSP was performed on the LC480 and plates were prepared using the Tecan freedom Evo robot, using a design that assures that all samples were tested for the same assay in the same run [65]. MSP amplifications were performed in 10 µl containing 5 µl LC480 SYBR Green I Master Mix (2x; Roche, Vilvoorde, Belgium), 1 mg/ml bovine serum albumin (Roche), 1 mM MgCl₂ (Roche), 125 nM forward and reverse primer (IDT, Leuven, Belgium), sample (20 ng amplified BT-

DNA) and nuclease-free water (Sigma). MSP conditions were as follows: activation for 10 minutes at 95°C, 45 amplification cycles (10 s at 95°C, 30 s at 60°C and 5 s at 72°C), followed by melting curve analysis (5 s at 95°C - melting curve from 60 to 95°C) and cool down to 45°C. Afterwards, the size of the amplicons was determined using the Caliper LabChip GX. A MSP assay was considered methylated if (1) its Cq-value < 35 (calculated by the LC480 software using the second derivative maximum method), (2) its melting temperature (T_m) differed no more than 2°C from that of the positive control sample, and (3) the amplicon length differed no more than 10 bp from the band size of the positive control sample. In addition, the band height, as determined by the LabChip GX software, was required to be higher than 20.

mRNA expression profiling

The mRNA expression levels of *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT* were profiled on the NB SIOOPEN/ GPOH cDNA library generated from 366 primary NB tumor samples (Additional file 4) [6]. For each DNA methylation marker a qPCR mRNA assay was designed and validated *in silico* and *in vitro* (Additional file 4) [66]. PCR plates were prepared as described in the previous section and RT-qPCR was performed on the LC480 as described in [6]. Relative gene expression levels were then normalized using the geometric mean of five reference sequences (*HPRT1*, *SDHA*, *UBC*, *HMBS* and *AluSq*) [67]. For *HIST1H3C* and *ACSS3*, a qPCR mRNA assay (Additional file 5) was designed and tested on 31 NB cell lines on which the corresponding MSP assay was tested as well. Here, qPCR amplifications were performed in 5 µl containing 2.5 µl SsoAdvanced SYBR Green Supermix (2x; Bio-Rad), 0.25 µl forward and reverse primer (5 µM each) and 2 µl cDNA sample (corresponding to 5 ng cDNA). Relative gene expression levels were normalized using the geometric mean of the reference sequences *SDHA*, *UBC* and *AluSq*. All RT-qPCR data analysis

was done in qbase-PLUS version 2.0 (Biogazelle, Ghent, Belgium) [65]. Logged and normalized qPCR data can be found in Additional file 4 and 5.

Statistical analysis

Statistical analyses were performed using IBM SPSS software version 19.0. All statistical tests were two-sided and p-values < 0.05 were considered statistically significant. Differential methylation across the prognostic groups was determined by the Chi square test. The relationship between the methylation status and NB risk factors was determined using Fisher's exact test. Univariate survival analysis was performed with the Kaplan-Meier method and log-rank statistics (Mantel-Cox) to determine the impact of methylation status on EFS and OS. EFS was defined as the time between initial diagnosis and relapse or death of disease, or time between diagnosis and last follow-up if no event had occurred. OS is the time to disease-related death or last follow-up. Hierarchical clustering and leave-one-out decision tree analysis were performed using R 2.13.0 (rpart package). The relationship between logged mRNA expression levels and the prognostic groups, OS status and NB risk factors was determined using the non-parametric Kruskal-Wallis test or Mann-Whitney test. Hazard ratios between logged mRNA expression data and survival were estimated using the Cox proportional hazard model. Kaplan-Meier curves were created by dichotomizing the logged mRNA expression data, using the median mRNA expression value as a cutoff. For *HIST1H3C*, the relationship between logged mRNA expression levels and the methylation status of the gene, and the *MYCN* status, was determined using the Mann-Whitney test. For all the above mentioned statistical tests, multiple hypothesis testing correction was performed (Benjamini-Hochberg method by using the R function p.adjust).

Additional materials

Additional files can be found at: <http://genomebiology.biomedcentral.com/articles/10.1186/gb-2012-13-10-r95>

Additional File 1. Clinical patient annotation, MSP assays and results on the BioTrove discovery platform. Part BioTrove clinical annotation. Patient characteristics of the samples used for the initial high-throughput MSP screening on the BioTrove OpenArray. PatientID is a unique patient number, the group indicates the risk. Clinical characteristics given are the age at diagnosis in months, International Neuroblastoma Staging System (INSS) stage, *MYCN* amplification status (0 is non-amplified and 1 is amplified), follow-up time (FU) in days after diagnosis, and overall survival (OS) and event-free survival (EFS) time in days after diagnosis. OS indicates whether the patient was alive (0) at the last known FU or died of disease (1). Similar for EFS, indicating events, such as relapse or progression. Empty cells represent missing values. **Part BioTrove assays.** MSP assays used in the initial high-throughput MSP screening on the BioTrove OpenArray. For each tested gene, the assay name(s) and corresponding forward and reverse primer(s) (5' to 3') are indicated, as well as the genomic location of the amplicon on the hg19 reference genome. **Part BioTrove results.** Results of the initial high-throughput MSP screening on the BioTrove OpenArray. For each assay and sample, the methylation call (dark blue is methylated, green unmethylated) is given. Samples are subdivided into two prognostic groups (HR-DOD and LR-SURV). The specificity and sensitivity of the assays is indicated as well. Assays selected for further testing are indicated in yellow. HR-DOD: high-risk deceased patients and LR-SURV: low-risk patients alive for at least 1000 days follow-up.

Additional File 2. Visualization of the

protocadherin beta gene cluster and the *HIST1H3C* promoter region in the Integrative Genomic Viewer. First panel. Visualization of a part of the protocadherin β (*PCDHB*) gene family cluster in IGV. For each NB cell line (SK-N-AS, CLB-GA, SH-SY5Y, SJNB-1, CHP-902R, IMR-32, LAN-2 and N206), the number of sequencing tags at each position is shown and the location of detected peaks is indicated with a red bar. The captured sequences clearly overlap with CpG islands in each individual member of this gene cluster. **Second panel.** Visualization of the promoter region of *HIST1H3C* in IGV. The four NB cell lines at the bottom (CHP-902R, IMR-32, LAN-2 and N206) are *MYCN* amplified. Three of these cell lines clearly show sequence tags in the CpG island, while the four NB cell lines on top (SK-N-AS, CLB-GA, SH-SY5Y and SJNB-1), which are *MYCN* non-amplified cell lines, do not show any signal.

Additional File 3. Clinical patient annotation, summary of clinical characteristics, MSP assays, results and summarized results (per clinical parameter) on the LC480 platform for 89 NB patient samples. Assays differentially methylated between prognostic groups and between NB risk factors are discussed in detail, as well as extended analyses on the MSP data (hierarchical clustering (heatmap) and survival analysis according to the number of methylation events (Kaplan-Meier plot)). **Part LC480 clinical annotation.** Patient characteristics of the samples used for the high-throughput MSP screening on the Roche LC480. PatientID is a unique patient number, the group indicates the risk. Clinical characteristics given are the age at diagnosis in months, International Neuroblastoma Staging System (INSS) stage, *MYCN* amplification status (0 is non-amplified and 1 is amplified), follow-up time (FU) in days after diagnosis, and overall survival (OS) and event-free survival (EFS) time in days after diagnosis. OS indicates whether the patient was alive (0) at the last known FU or died of disease

(1). Similar for EFS, indicating events such as relapse or progression. Segmental aberrations for chromosome 1p, 11q and 17q are indicated with normal, partial loss (PL), whole loss (WL), partial gain (PG) or whole gain (WG). Empty cells represent missing values. **Part LC480 overview annotation.** Summary of the clinical characteristics of the 89 primary NB samples used in the high-throughput MSP screening. **Part LC480 assays.** MSP assays used in the high-throughput MSP screening on the Roche LC480. For each tested gene, the assay name(s) and corresponding forward and reverse primer(s) (5' to 3') are indicated, as well as the genomic location of the amplicon on the hg19 reference genome. **Part LC480 results.** Results of the high-throughput MSP screening on the Roche LC480. For each assay and for each sample, the methylation call (dark blue is methylated, green unmethylated) is given. The patient samples are subdivided into three prognostic groups (LR-SURV, HR-DOD and HR-SURV). CL: cell line; Neg: negative control (HCT-116 DKO cell line); NTC: no template control. **Part LC480 overview results.** Summary of the MSP results of the high-throughput study. The number of methylated samples for a particular MSP assay within each prognostic group and for the entire sample set is given (percentage of methylated samples between brackets). Assays are ranked descending on the number of overall methylated samples. Assays discussed in more detail in the results section are indicated in grey. **Part Group associations.** Detailed information on a selection of assays, differentially methylated between prognostic groups (per prognostic group and per combination of risk factors). The number (percentage) of methylated samples in each stratum is given. **Part Risk factor associations.** Detailed information on a selection of assays, differentially methylated between NB risk factors (per (combination of) risk factors). The number (percentage) of methylated samples in each stratum is given. **Part Heatmap.** Hierarchical clustering based on the MSP data of

the 89 primary NB tumor samples. This analysis reveals two clusters, indicated with group 1 and group 2, which predominantly distinguish high-risk and low-risk NB patient samples. Each sample is assigned to a prognostic group (LR-SURV, HR-DOD and HR-SURV). **Part KM number.** Kaplan-Meier plot: overall survival in the samples of the high-throughput MSP screening according to the number of methylation events. The purple line indicates patients with 0 to 16 methylation events, the red line patients with 16 to 27 methylation events. This 16 methylation events cutoff was used, as the average number of methylation events in one sample is 16. The p-value is determined using a log-rank test (Mantel-Cox). Time is indicated in days, starting from diagnosis. HR-DOD: high-risk deceased patients; HR-SURV: high-risk patients alive for at least 1000 days follow-up; INSS: International Neuroblastoma Staging System; LR-SURV: low-risk patients alive for at least 1000 days follow-up.

Additional File 4. Clinical annotation, summary of clinical characteristics, qPCR assays and results of qPCR experiments on 366 NB patient samples (SIOPEN/GPOH cDNA library). Boxplots of the expression levels for *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT* in each of the five different NB stages (stages 1, 2, 3, 4 and 4S). A Kaplan-Meier plot shows overall survival according to the relative mRNA expression levels of *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT*. **Part SIOPEN clinical annotation.** Patient characteristics of the samples of the NB SIOPEN/GPOH cDNA library. PatientID is a unique patient number, the group indicates the risk. Clinical characteristics given are the age at diagnosis in months, International Neuroblastoma Staging System (INSS) stage, *MYCN* amplification status (0 is non-amplified and 1 is amplified), overall survival (OS) and event-free survival (EFS) time in days and months after diagnosis, respectively. OS indicates whether the patient was alive (0) at the last known follow-up or died of disease (1).

Similar for EFS, indicating events, such as relapse or progression. Empty cells represent missing values. **Part SIOPEN overview annotation.** Summary of the clinical characteristics of the 366 primary NB samples of the SIOPEN/GPOH cDNA library. **Part SIOPEN assays.** qPCR mRNA assays used in the mRNA expression profiling of the DNA-methylation biomarkers *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT*. For each tested gene, the assay name and corresponding forward and reverse primer (5' to 3') are indicated, as well as the genomic location of the amplicon on the hg19 reference genome. **Part SIOPEN results.** Results of the mRNA expression measurement of the DNA-methylation biomarkers *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT*. The Cq values were converted to relative quantities and log2 values. Relative gene expression levels were then normalized using the geometric mean of five reference sequences (*HPRT1*, *SDHA*, *UBC*, *HMBS* and *AluSq*). These logged and normalized qPCR data are given. **Part Expression.** Relative mRNA expression distribution of *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT* in each of the five different NB tumor stages (stage 1, 2, 3, 4 and 4S). In the box plots, the right and left hinge of the boxes represents the 75th percentile and 25th percentile, respectively. Whiskers, lines that extend from the box to the highest and lowest values, indicate the data range. Significant p-values according to the Kruskal-Wallis test (corrected for multiple testing using Benjamini-Hochberg) are indicated. **Part Survival.** Kaplan-Meier plots: overall survival in the NB SIOPEN/GPOH samples according to the relative mRNA expression levels of *CNR1*, *GRB10*, *KRT19*, *PRPH* and *QPCT*. Survival curves indicated with 'High' and 'Low' are the samples assigned to the high and low mRNA expression group, respectively, using the median relative mRNA expression value as a cutoff to create the groups. Significant p-values according to the log-rank test (Mantel-Cox; corrected for multiple testing using Benjamini-Hochberg) are indicated. Time is indicated in days, starting from diagnosis

and censored to 4000 days (censored samples are indicated with vertical lines crossing the overall survival curves). HR-DOD: high-risk deceased patients; HR-SURV: high-risk patients alive for at least 1000 days follow-up; INSS: International Neuroblastoma Staging System; LR-SURV: low-risk patients alive for at least 1000 days follow-up.

Additional File 5. Quantitative PCR and MSP assays for *HIST1H3C* and *ACSS3* and matched results (expression levels - methylation call) for a panel of 31 NB cell lines. Part *HIST1H3C* and *ACSS3* qPCR mRNA assay. qPCR mRNA assays used in the mRNA expression profiling of the DNA-methylation biomarkers *HIST1H3C* and *ACSS3*. For each gene, the assay name and corresponding forward and reverse primer (5' to 3') are indicated. **Part *HIST1H3C* - *ACSS3* MSP cell lines.** Results of the *HIST1H3C* and *ACSS3* MSP screen on 31 NB cell lines. The methylation call (dark blue is methylated, green unmethylated) of each cell line is indicated, as well as the *MYCN* amplification status. **Part *HIST1H3C* qPCR cell lines.** Results of the mRNA expression measurement of the DNA-methylation biomarker *HIST1H3C*. The Cq values were converted to relative quantities and converted to log2 values. Relative gene expression levels were then normalized using the geometric mean of the reference sequences *SDHA*, *UBC* and *AluSq*. These logged and normalized qPCR data are given.

Acknowledgments

The authors thank Els De Smet, Nurten Yigit, Justine Nuytens, Johan Vandersmissen and Jean-Pierre Renard (MDxHealth) for their kind help and NXTGNT for the excellent sequencing services. AD, KDP and TVM are supported by a grant of the Research Foundation Flanders (FWO), and GVP by a grant of the Ghent University Special Research Fund (BOF). RLS was a recipient of grants from Science Foundation

Ireland (07/IN.1/B1776), Children's Medical and Research Foundation, and National Institutes of Health (NIH; 5R01CA127496). RN was a recipient of grants (RD06/0020/0102) from Thematic Network of Cooperative Research on Cancer (RTICC), Institute of Health Carlos III (ISCIII) and European Regional Development Fund (ERDF), and Spanish Health Research Fund (FIS; 2010/PI015). This study was further supported by the Emmanuel Van der Schueren Foundation (scientific partner of the Flemish League Against Cancer (VLK)), Fournier-Majoie Foundation (FFM), UGent GOA (grant number 01G01910), Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT) and FWO (grant number G0530.12N).

References

- [1] Øra I et al. (2011). Progress in treatment and risk stratification of neuroblastoma: impact on future clinical and basic research. *Seminars in Cancer Biology*; 21(4):217-228.
- [2] Schleiermacher G et al. (2011). Segmental chromosomal alterations lead to a higher risk of relapse in infants with MYCN-non-amplified localized unresectable/disseminated neuroblastoma (a SIOPEN collaborative study). *British Journal of Cancer*; 105(12):1940-1948.
- [3] Vandesompele J et al. (2005). Unequivocal delineation of clinicogenetic subgroups and development of a new model for improved outcome prediction in neuroblastoma. *Journal of Clinical Oncology*; 23(10):2280-2299.
- [4] Cohn SL et al. (2009). The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. *Journal of Clinical Oncology*; 27(2):289-297.
- [5] Schulte JH et al. (2010). Accurate prediction of neuroblastoma outcome based on miRNA expression profiles. *International Journal of Cancer*; 127(10):2374-2385.
- [6] Vermeulen J et al. (2009). Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/COG/ GPOH study. *Lancet Oncology*; 10(7):663-671.
- [7] De Preter K et al. (2011). miRNA expression profiling enables risk stratification in archived and fresh neuroblastoma tumor samples. *Clinical Cancer Research*; 17(24):7684-7692.

- [8] Portela A et al. (2010). Epigenetic modifications and human disease. *Nature Biotechnology*; 28(10):1057-1068.
- [9] Teitz T et al. (2000). Caspase 8 is deleted or silenced preferentially in childhood neuroblastomas with amplification of MYCN. *Nature Medicine*; 6(5):529-535.
- [10] Astuti D et al. (2001). RASSF1A promoter region CpG island hypermethylation in haemochromocytomas and neuroblastoma tumours. *Oncogene*; 20(51):7573-7577.
- [11] Michalowski MB et al. (2008). Methylation of tumor-suppressor genes in neuroblastoma: the RASSF1A gene is almost always methylated in primary tumors. *Pediatric Blood and Cancer*; 50(1):29-32.
- [12] Yang Q et al. (2004). Methylation-associated silencing of the heat shock protein 47 gene in human neuroblastoma. *Cancer Research*; 64(13):4531-4538.
- [13] Grau E et al. (2010). Hypermethylation of apoptotic genes as independent prognostic factor in neuroblastoma disease. *Molecular Carcinogenesis*; 50(3):153-162.
- [14] Banelli B et al. (2005). Distinct CpG methylation profiles characterize different clinical groups of neuroblastic tumors. *Oncogene*; 24(36):5619-5628.
- [15] Misawa A et al. (2009). RASSF1A hypermethylation in pretreatment serum DNA of neuroblastoma patients: a prognostic marker. *British Journal of Cancer*; 100(2):399-404.
- [16] Murphy DM et al. (2009). Global MYCN transcription factor binding analysis in neuroblastoma reveals association with distinct E-box motifs and regions of DNA hypermethylation. *PLoS One*; 4(12):e8154.
- [17] Buckley PG et al. (2010). Genome-wide DNA methylation analysis of neuroblastic tumors reveals clinically relevant epigenetic events and large-scale epigenomic alterations localized to telomeric regions. *International Journal of Cancer*; 128(10):2296-2305.
- [18] Caren H et al. (2011). Identification of epigenetically regulated genes that predict patient outcome in neuroblastoma. *BMC Cancer*; 11:66.
- [19] Decock A et al. (2011). Neuroblastoma epigenetics: from candidate gene approaches to genome-wide screenings. *Epigenetics*; 6(8):962-970.
- [20] Hoque MO et al. (2008). Genome-wide promoter analysis uncovers portions of the cancer methylome. *Cancer Research*; 68(8):2661-2670.
- [21] Ongenaert M et al. (2008). Discovery of DNA methylation markers in cervical cancer using relaxation ranking. *BMC Medical Genomics*; 1:57.
- [22] Huss M (2010). Introduction into the analysis of high-throughput-sequencing based epigenome data. *Briefings in Bioinformatics*; 11:512-523.
- [23] Abe M et al. (2005). CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas. *Cancer Research*; 65(3):828-834.
- [24] Banelli B et al. (2011). A pyrosequencing assay for the quantitative methylation analysis of the PCDHB gene cluster, the major factor in neuroblastoma methylator phenotype. *Laboratory Investigation*; 92(3):458-465.
- [25] Anders S et al. (2010). Differential expression analysis for sequence count data. *Genome Biology*; 11(10):R106.
- [26] Subramanian A et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*; 102(43):15545-15550.
- [27] Cole KA et al. (2011). RNAi screen of the protein kinome identifies checkpoint kinase 1 (CHK1) as a therapeutic target in neuroblastoma. *PNAS*; 108(8):3336-3341.
- [28] Ohtaki M et al. (2010). A robust method for estimating gene expression states using Affymetrix microarray probe level data. *BMC Bioinformatics*; 11:183.
- [29] Bourdeaut F et al. (2009). Cholinergic switch associated with morphological differentiation in neuroblastoma. *Journal of Pathology*; 219(4):463-472.
- [30] Janoueix-Lerosey I et al. (2008). Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature*; 455(7215):967-970.
- [31] Łastowska M et al. (2007). Identification of candidate genes involved in neuroblastoma progression by combining genomic and expression microarrays with survival data. *Oncogene*; 26(53):7432-7444.
- [32] Wang Q et al. (2006). Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. *Cancer Research*; 66(12):6050-6062.
- [33] Balamuth NJ et al. (2010). Serial transcriptome analysis and cross-species integration identifies centromere-associated protein E as a novel neuroblastoma target. *Cancer Research*; 70(7):2749-2758.
- [34] Wang K et al. (2011). Integrative genomics

- identifies LMO1 as a neuroblastoma oncogene. *Nature*; 469(7329):216-220.
- [35] Schmidt ML et al. (2005). Favorable prognosis for patients 12 to 18 months of age with stage 4 nonamplified MYCN neuroblastoma: a Children's Cancer Group Study. *Journal of Clinical Oncology*; 23(27):6474-6480.
- [36] London WB et al. (2005). Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's Oncology Group. *Journal of Clinical Oncology*; 23(27):6459-6465.
- [37] Yagyu S et al. (2008). Circulating methylated-DCR2 gene in serum as an indicator of prognosis and therapeutic efficacy in patients with MYCN nonamplified neuroblastoma. *Clinical Cancer Research*; 14(21):7011-7019.
- [38] Yang Q et al. (2007). Methylation of CASP8, DCR2, and HIN-1 in neuroblastoma is associated with poor outcome. *Clinical Cancer Research*; 13(11):3191-3197.
- [39] Banelli B et al. (2010). Outcome prediction and risk assessment by quantitative pyrosequencing methylation analysis of the SFN gene in advanced stage, high-risk, neuroblastic tumor patients. *International Journal of Cancer*; 126(3):656-668.
- [40] Kamimatsuse A et al. (2009). Detection of CpG island hypermethylation of caspase-8 in neuroblastoma using an oligonucleotide array. *Pediatric Blood and Cancer*; 52(7):777-783.
- [41] Yang Q et al. (2004). Association of epigenetic inactivation of RASSF1A with poor outcome in human neuroblastoma. *Clinical Cancer Research*; 10(24):8493-8500.
- [42] Alaminos M et al. (2004). Clustering of gene hypermethylation associated with clinical risk groups in neuroblastoma. *Journal of the National Cancer Institute*; 96(16):1208-1219.
- [43] Grau E et al. (2010). Epigenetic alterations in disseminated neuroblastoma tumour cells: influence of TMS1 gene hypermethylation in relapse risk in NB patients. *Journal of Cancer Research and Clinical Oncology*; 136(9):1415-1421.
- [44] Petak I et al. (2003). Hypermethylation of the gene promoter and enhancer region can regulate Fas expression and sensitivity in colon carcinoma. *Cell Death and Differentiation*; 10(2):211-217.
- [45] Chaopatchayakul P et al. (2010). Aberrant DNA methylation of apoptotic signaling genes in patients responsive and nonresponsive to therapy for cervical carcinoma. *American Journal of Obstetrics and Gynecology*; 202(3):281.e1-9.
- [46] Schaefer JT et al. (2000). Ceramide induces apoptosis in neuroblastoma cell cultures resistant to CD95 (Fas/APO-1)-mediated apoptosis. *Journal of Pediatric Surgery*; 35(3):473-479.
- [47] Kisenge RR et al. (2003). Expression of short-form caspase 8 correlates with decreased sensitivity to Fas-mediated apoptosis in neuroblastoma cells. *Cancer Science*; 94(7):598-605.
- [48] Takamizawa S et al. (2000). Overexpression of Fas-ligand by neuroblastoma: a novel mechanism of tumor-cell killing. *Journal of Pediatric Surgery*; 35(2):375-379.
- [49] Bernassola F et al. (1999). Induction of apoptosis by IFNgamma in human neuroblastoma cell lines through the CD95/CD95L autocrine circuit. *Cell Death and Differentiation*; 6(7):652-660.
- [50] Gross N et al. (2001). Expression of Fas (APO-1/CD95) and Fas ligand (FasL) in human neuroblastoma. *Medical and Pediatric Oncology*; 36(1):111-114.
- [51] Pedersen WA et al. (1993). Expression and distribution of peripherin protein in human neuroblastoma cell lines. *International Journal of Cancer*; 53(3):463-470.
- [52] Ambros IM et al. (2002). Morphologic features of neuroblastoma (Schwannian stroma-poor tumors) in clinically favorable and unfavorable groups. *Cancer*; 94(5):1574-1583.
- [53] Molenaar JJ et al. (2012). Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature*; 483(7391):589-593.
- [54] Chen Y et al. (2007). Differential patterns of microRNA expression in neuroblastoma are correlated with prognosis, differentiation, and apoptosis. *Cancer Research*; 67(3):976-983.
- [55] Mestdagh P et al. (2009). A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biology*; 10(6):R64.
- [56] Schulte JH et al. (2008). MYCN regulates oncogenic microRNAs in neuroblastoma. *International Journal of Cancer*; 122(3):699-704.
- [57] Ragusa M et al. (2010). MIR152, MIR200B, and MIR338, human positional and functional neuroblastoma candidates, are involved in neuroblast differentiation and apoptosis. *Journal of Molecular Medicine*; 88(10):1041-1053.
- [58] Tan YC et al. (2011). Autosomal dominant polycystic kidney disease: genetics, mutations and microRNAs. *Biochimica et Biophysica Acta*; 1812(10):1202-1212.
- [59] Hansen TB et al. (2011). miRNA-dependent gene silencing involving Ago2-mediated

- cleavage of a circular antisense RNA. *EMBO Journal*; 30(21):4414-4422.
- [60] Teitz T et al. (2001). Aggressive childhood neuroblastomas do not express caspase-8: an important component of programmed cell death. *Journal of Molecular Medicine*; 79(8):428-436.
- [61] Hoebeeck J et al. (2009). Aberrant methylation of candidate tumor suppressor genes in neuroblastoma. *Cancer Letters*; 273(2):336-346.
- [62] Gentleman RC et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*; 5(10):R80.
- [63] Morgan M et al. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics*; 25(19):2607-2608.
- [64] Lawrence M et al. (2009). rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics*; 25(14):1841-1842.
- [65] Hellemans J et al. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*; 8(2):R19.
- [66] Lefever S et al. (2009). RTPrimerDB: the portal for real-time PCR primers and probes. *Nucleic Acids Research*; 37:D942-945.
- [67] Vandesompele J et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*; 3(7):RESEARCH0034.

paper 2

Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma

Anneleen Decock, Maté Ongenaert, Robrecht Cannoodt, Kimberly Verniers, Bram De Wilde, Geneviève Laureys, Nadine Van Roy, Ana P. Berbegall, Julie Bienertova-Vasku, Nick Bown, Nathalie Clément, Valérie Combaret, Michelle Haber, Claire Hoyoux, Jayne Murray, Rosa Noguera, Gaelle Pierron, Gudrun Schleiermacher, Johannes H. Schulte, Ray L. Stallings, Deborah A. Tweddle for the Children's Cancer and Leukaemia Group (CCLG), Katleen De Preter, Frank Speleman and Jo Vandesompele

published in Oncotarget 7(2):1960-1972, 2015

impact factor 2015: 5.008

times cited (d.d. 11/09/2016; Google Scholar): 2

PAPER 2: METHYL-CpG-BINDING DOMAIN SEQUENCING REVEALS A PROGNOSTIC METHYLATION SIGNATURE IN NEUROBLASTOMA

Anneleen Decock^{1,2}, Maté Ongenaert¹, Robrecht Cannoodt^{1,2,3,4,5}, Kimberly Verniers^{1,2}, Bram De Wilde^{1,2,6}, Geneviève Laureys⁶, Nadine Van Roy^{1,2}, Ana P. Berbegall⁷, Julie Bienertova-Vasku⁸, Nick Bown⁹, Nathalie Clément¹⁰, Valérie Combaret¹¹, Michelle Haber¹², Claire Hoyoux¹³, Jayne Murray¹², Rosa Noguera⁷, Gaëlle Pierron¹⁴, Gudrun Schleiermacher¹⁵, Johannes H. Schulte¹⁶, Ray L. Stallings^{17,18}, Deborah A. Tweddle¹⁹ for the Children's Cancer and Leukaemia Group (CCLG), Katleen De Preter^{1,2,3}, Frank Speleman^{1,2} and Jo Vandesompele^{1,2,3}

Correspondence to joke.vandesompele@ugent.be.

¹Center for Medical Genetics, Ghent University, De Pintelaan 185, Ghent, Belgium. ²Cancer Research Institute Ghent (CRIG), De Pintelaan 185, Ghent, Belgium. ³Bioinformatics Institute Ghent - From Nucleotides to Networks (BIG N2N), De Pintelaan 185, Ghent, Belgium. ⁴DAMBI, VIB Inflammation Research Center, Technologiepark 927, Ghent, Belgium. ⁵Department of Respiratory Medicine, Ghent University, De Pintelaan 185, Ghent, Belgium. ⁶Department of Pediatric Hematology and Oncology, Ghent University Hospital, De Pintelaan 185, Ghent, Belgium. ⁷Department of Pathology, Medical School, University of Valencia, and Health Research Institute INCLIVA, Blasco Ibañez 17, Valencia, Spain. ⁸Department of Pathological Physiology, Department of Pediatric Oncology, Masaryk University, Černoplní 212/9, Brno, Czech Republic. ⁹Northern Genetics Service, Institute of Genetic Medicine, Central Parkway, Newcastle upon Tyne, United Kingdom. ¹⁰Department of Pediatric Oncology, Institut Curie, 26 rue d'Ulm, Paris, France. ¹¹Centre Léon Bérard, Laboratoire de Recherche Translationnelle, 28 rue Laennec, Lyon, France. ¹²Children's Cancer Institute, Lowy Cancer Research Centre, UNSW, PO Box 81, Randwick NSW 2031, Australia. ¹³Pediatric Hemato-oncology, CHR Citadelle, Boulevard du 12ème de Ligne 1, Liège, Belgium. ¹⁴Unité de Génétique Somatique, Institut Curie, 26 rue d'Ulm, Paris, France. ¹⁵U830 INSERM, Recherche Translationnelle en Oncologie Pédiatrique (RTOP) and Department of Pediatric Oncology, Institut Curie, 26 rue d'Ulm, Paris, France. ¹⁶Department of Pediatric Oncology and Hematology, University Children's Hospital Essen, Hufelandstraße 55, Essen, Germany. ¹⁷National Children's Research Centre, Our Lady's Children's Hospital, Crumlin, Dublin 12, Ireland. ¹⁸Department of Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, York House, York Street, Dublin 2, Ireland. ¹⁹Newcastle Cancer Centre, Northern Institute for Cancer Research, Newcastle University, Framlington Place, Newcastle upon Tyne, United Kingdom.

Contribution of AD: AD drafted the article, collected and prepared samples, executed DNA shearing for MBD sequencing, performed differential methylation analyses on the MBD sequencing data, prioritized biomarker candidates, developed MSP assays, tested the candidate biomarkers, analyzed the MSP data, performed survival analyses and helped in the generation of the methylation signature.

Abstract

Accurate assessment of neuroblastoma outcome prediction remains challenging. Therefore, this study aims at establishing novel prognostic tumor DNA methylation biomarkers. In total, 396 low- and high-risk primary tumors were analyzed, of which 87 were profiled using methyl-CpG-binding

domain (MBD) sequencing for differential methylation analysis between prognostic patient groups. Subsequently, methylation-specific PCR (MSP) assays were developed for 78 top-ranking differentially methylated regions and tested on two independent cohorts of 132 and 177 samples, respectively. Further, a new statistical framework was used to identify a robust set of MSP assays of which the methylation score (i.e. the percentage of methylated assays) allows accurate outcome prediction. Survival analyses were performed on the individual target level, as well as on the combined multimarker signature. As a result of the differential DNA methylation assessment by MBD sequencing, 58 of the 78 MSP assays were designed in regions previously unexplored in neuroblastoma, and 36 are located in non-promoter or non-coding regions. In total, 5 individual MSP assays (located in *CCDC177*, *NXPH1*, *Inc-MRPL3-2*, *Inc-TREX1-1* and one on a region from chromosome 8 with no further annotation) predict event-free survival and 4 additional assays (located in *SPRED3*, *TNFAIP2*, *NPM2* and *CYYR1*) also predict overall survival. Furthermore, a robust 58-marker methylation signature predicting overall and event-free survival was established. In conclusion, this study encompasses the largest DNA methylation biomarker study in neuroblastoma so far. We identified and independently validated several novel prognostic biomarkers, as well as a prognostic 58-marker methylation signature.

Keywords: neuroblastoma, prognosis, DNA methylation, biomarker, biomarker signature

Introduction

Neuroblastoma (NB), a childhood tumor that originates from precursor cells of the sympathetic nervous system, is a heterogeneous disease with prognosis ranging from excellent long-term survival to high-risk with fatal outcome. In order to determine the most appropriate treatment modalities for each patient, patients are stratified into risk groups at the time of diagnosis, based on combinations of clinical (age of the patient, stage of the tumor) and biological (*MYCN* amplification status, DNA index, histopathology) parameters [1]. Use of this risk classification system has shown that patients characterized by the same clinicobiological parameters can have different disease outcomes, indicating that accurate assessment of prognosis of NB patients still remains difficult [2–4]. Therefore, additional prognostic markers are warranted, allowing a more accurate risk estimation and more rapid identification of those patients who will not benefit from current treatments.

Molecular alterations of the epigenome,

especially DNA methylation, have emerged as alternative targets of biomarker research. DNA methylation biomarkers potentially have great clinical value due to the stable nature of DNA. For this reason, there are many relevant applications of DNA methylation biomarkers in cancer. For example, they could be used for early tumor detection, tumor classification, stratification of treatment, tumor recurrence and patient prognosis, as well as predicting and monitoring a patient's response to treatment (detailed review in reference [5]). In NB, several prognostic single-gene methylation biomarkers have been reported, e.g. promoter methylation of *TNFRSF10D*, *CASP8*, *ZMYND10*, *RASSF1A*, *KRT19*, *GNAS*, *HIST1H3C*, *RB1* and *TDGF1* [6–11]. Furthermore, a CpG island methylator phenotype (CIMP), described as the aberrant and concordant methylation of multiple promoter CpG islands, has been shown to be of prognostic significance [12–16].

In this study, we aim to assess the primary NB tumor methylome in a genome-wide manner to identify differentially methylated regions (DMRs) between the prognostic patient groups, and to

use these DMRs to establish and validate new and valuable biomarkers.

Results

Methyl-CpG-binding domain (MBD) sequencing of primary tumors prioritizes differentially methylated regions (DMRs) between patient subgroups

The study design is schematically represented in Figure 1. In the discovery phase, two independent cohorts of 42 (MBD cohort I) and 45 (MBD cohort II) primary NB tumors, selected for risk classification and survival (low-risk survivors (LR-SURV), high-risk survivors (HR-SURV) and high-risk deceased (HR-DOD)), were analyzed by methyl-CpG-binding domain (MBD) sequencing (Supplemental Table 1A and B). Sheared input DNA was enriched towards methylated fragments using the high affinity of the MBD of the MeCP2 protein towards methylated cytosines. These methylation-enriched fractions, as well as the input (non-MBD-enriched) DNA of MBD cohort II were then further studied by next-generation sequencing. After raw data analyses, differentially methylated regions (DMRs) between patient subgroups were detected using DESeq, which uses count data as input. The following patient subgroups were compared: HR-SURV versus HR-DOD (on the entire cohorts, as well as on the high-risk *MYCN* amplified (HR-MYCN1) and non-amplified (HR-MYCN0) cohorts only), LR-SURV versus HR-DOD, and HR-MYCN0 versus HR-MYCN1 (Supplemental Table 2). The same analyses were performed on the input sample data in order to estimate the background signal and exclude falsely identified DMRs. The DESeq analyses yield for each region of interest the mean normalized counts per patient group, as well as the log₂ fold change and p-value for the statistical significance of the difference. By calculating the π -value ($\pi = -\ln p\text{-val} * \log_2 \text{fold change}$ [17]) for each of these regions, a new

significance score was defined, which was then used to rank the candidate prognostic DMRs. Hierarchical cluster analysis using normalized counts of the top-ranking DMRs yielded two sample clusters which mainly correspond to the patient groups used in the differential methylation analysis, highlighting the capability of our MBD sequencing analysis strategy in identifying biomarker candidates (examples shown in Supplemental Figure 1).

Methylation-specific PCR (MSP) assays are designed and tested on two independent cohorts

MBD sequencing data of the top-ranking DMRs (promoter regions and 5 kb windows) from the different prognostic comparisons were visualized in the Integrative Genomics Viewer (IGV; [18]) in order to locate the most informative (discriminative) region for MSP primer design (Figure 1). The importance of this step is illustrated by the promoter region of *HNRNPH1*, which was identified as differentially methylated between HR-SURV and HR-DOD patients, and LR-SURV and HR-DOD patients (Supplemental Figure 2). MBD regions for which no clear discriminative region could be identified were excluded from further analyses and only DMRs hypermethylated in HR-DOD or HR-MYCN1 samples were considered for further evaluation. In total, 78 MSP assays (Supplemental Table 2) were designed, analytically validated and tested on 19 NB cell lines (Supplemental Table 3), positive and negative controls (the (*in vitro* methylated) HCT-116 DKO cell line), along with two independent cohorts of 148 (MSP cohort I) and 202 (MSP cohort II) primary NB samples assigned to one of the three defined prognostic patient groups (Supplemental Table 4). Also the *ACTB* primer pair, a control assay that does not contain CpG sites and thus should always generate a PCR product, was tested on these samples to confirm successful DNA preparation (bisulfite treatment and amplification). In total,

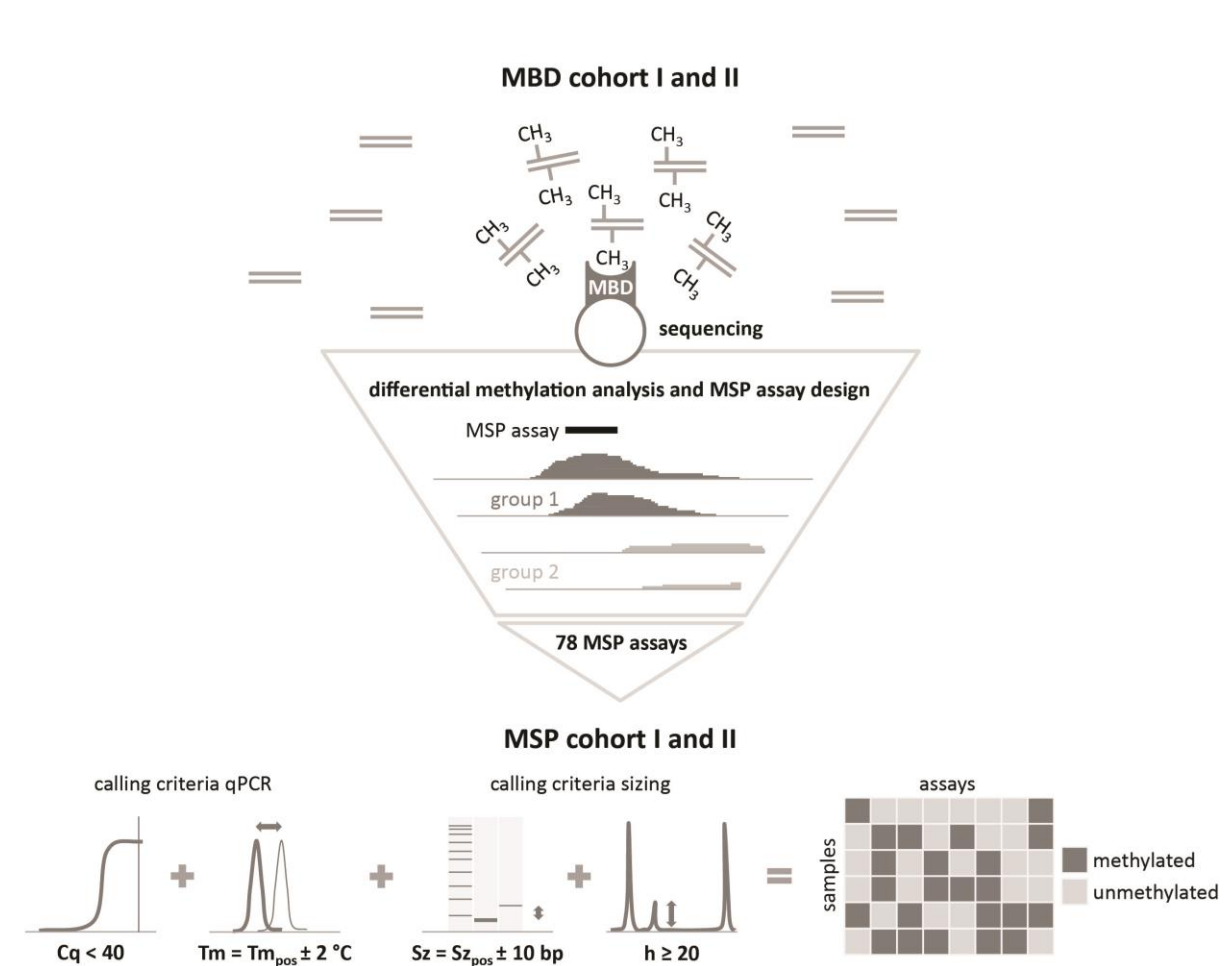


Figure 1. Schematic representation of the study design. Differentially methylated regions (DMRs) between the prognostic patient groups are identified by methyl-CpG-binding domain (MBD) sequencing on MBD cohort I and II. For the top candidate prognostic DMRs, the MBD sequencing data were visualized in order to locate the most informative region for methylation-specific PCR (MSP) assay design. These assays were subsequently tested on MSP cohort I and II. By applying specific methylation calling criteria [10], a binary dataset for each of these cohorts was constructed, which was subsequently used for survival analyses. Cq: quantification cycle; h: height; Sz: size; Tm: melting temperature. The subscript pos refers to the data of the positive control sample.

16 samples of MSP cohort I and 25 samples of MSP cohort II failed for this assay, probably due to low DNA quality, and were therefore excluded from the study.

MSP confirms the validity of MBD sequencing in identifying candidate methylation biomarkers

In both MSP cohort I and II, primary tumor samples of HR-DOD and HR-MYC*N1* patients show more methylation events compared to either survivors ($p = 0.001$ for both cohorts; Supplemental Figure 3A and 3B) and HR-MYC*N0* patients ($p < 0.001$ for both cohorts; Supplemental Figure 3C and 3D), respectively. This again confirms the validity of MBD

sequencing data in identifying candidate markers, as all MSP assays were designed in regions identified in the MBD sequencing data as being hypermethylated in HR-DOD or HR-MYCIN1 patients. To further strengthen MBD sequencing as a powerful technology for identification of genome-wide differential methylation, the genomic locations of the in-house designed MSP assays were compared to the genomic locations of the cytosines interrogated on the Infinium HumanMethylation450 BeadChip Kit (HM450 array; Illumina). Of note, 58 MSP assays (74.36%) do not overlap with an interrogated cytosine on the HM450 array, and would thus not have been identified using this array technology (e.g. promoter region of *UHRF2* in Supplemental

Table 1. Survival analyses on the individual MSP assay level identify new biomarkers for overall and event-free survival.

overall survival variable	log-rank p	MSP cohort I			log-rank p	MSP cohort II		
		univariable logistic regression				univariable logistic regression		
	p	p	OR	95% CI	p	p	OR	95% CI
assay006 (SPRED3)	0.005	0.043	2.26	1.03 - 4.96	0.014	0.030	2.08	1.08 - 4.04
assay008 (TNFAIP2)	0.008	0.009	3.13	1.33 - 7.40	0.025	0.020	2.28	1.14 - 4.57
assay011	0.002	0.001	3.56	1.64 - 7.75	< 0.001	< 0.001	3.82	1.81 - 8.08
assay062 (NPM2)	0.021	0.010	4.19	1.41 - 12.46	0.038	0.036	2.38	1.06 - 5.33
assay087 (NXPH1)	0.014	0.043	2.26	1.03 - 4.96	0.003	0.004	2.60	1.37 - 4.95
assay108 (CYYR1)	0.024	0.004	3.46	1.48 - 8.13	0.046	0.040	1.93	1.03 - 3.61
assay111 (CCDC177)	0.002	< 0.001	3.68	1.71 - 7.89	0.020	0.022	2.11	1.11 - 3.99
assay113 (Inc-MRPL3-2)	0.002	0.004	3.03	1.43 - 6.43	0.034	0.056	1.84	0.98 - 3.45
assay116 (Inc-TREX1-1)	0.008	0.056	3.27	0.97 - 10.98	0.004	0.021	2.87	1.18 - 7.01
event-free survival variable	log-rank p	univariable logistic regression			log-rank p	univariable logistic regression		
	p	p	OR	95% CI	p	p	OR	95% CI
assay011	0.007	0.006	2.77	1.35 - 5.70	< 0.001	< 0.001	4.18	2.02 - 8.67
assay087 (NXPH1)	0.018	0.017	2.58	1.19 - 5.62	0.003	0.004	2.56	1.35 - 4.83
assay111 (CCDC177)	< 0.001	< 0.001	4.94	2.33 - 10.48	0.007	0.004	2.52	1.34 - 4.74
assay113 (Inc-MRPL3-2)	0.035	0.040	2.12	1.04 - 4.32	0.038	0.030	1.98	1.07 - 3.67
assay116 (Inc-TREX1-1)	0.022	0.060	3.33	0.95 - 11.70	0.019	0.055	2.38	0.98 - 5.80

Note. For each individual MSP assay, the log-rank p-values, and the p-value, odds ratio (OR) and 95% confidence interval (CI) of the univariable logistic regression analyses are shown. Methylation of the individual markers is associated with worse overall and event-free survival.

Figure 4). Also, 36 MSP assays (46.15%) are located in non-promoter or non-coding regions.

Survival analyses on the individual MSP assay level identify new prognostic biomarkers

Overall, the percentage of methylated samples per MSP assay ranges from 96.97% to 2.27% in MSP cohort I, and from 97.18% to 1.70% in MSP cohort II, and variable percentages between the prognostic patient groups are detected (Supplemental Table 4). The results of the survival analyses (log-rank test) on each

individual MSP assay and the different patient (sub)cohorts are indicated in Supplemental Table 2. Although the survival analyses on the high-risk subgroups did not yield significant results, analyses on the entire cohorts identified 9 individual prognostic MSP assays for overall survival (OS) and 5 assays for event-free survival (EFS) that were significantly detected in both MSP cohort I and II (Table 1). For EFS, these assays are located in the promoter region or gene body of *CCDC177* and *NXPH1*, and the long non-coding RNAs *Inc-MRPL3-2* and *Inc-TREX1-1*. The additional prognostic assays for OS are

Table 2. The nine individual prognostic MSP assays are differentially methylated between patient groups with distinct neuroblastoma risk factors (MSP cohort I).

		MSP cohort I								
factor - number (percentage)		assay006 (<i>SPRED3</i>)	assay008 (<i>TNFAIP2</i>)	assay011	assay062 (<i>NPM2</i>)	assay087 (<i>NXPH1</i>)	assay108 (<i>CYYR1</i>)	assay111 (<i>CCDC177</i>)	assay113 (<i>Inc-MRPL3-2</i>)	assay116 (<i>Inc-TREX1-1</i>)
INSS stage	stage 1 (n = 27)	3 (11.11)	0 (0.00)	6 (22.22)	0 (0.00)	2 (7.41)	1 (3.70)	2 (7.41)	8 (29.63)	1 (3.70)
	stage 2 (n = 18)	1 (5.56)	1 (5.56)	7 (38.89)	1 (5.56)	1 (5.56)	0 (0.00)	5 (27.78)	4 (22.22)	0 (0.00)
	stage 3 (n = 33)	15 (45.45)	10 (30.30)	17 (51.52)	5 (15.15)	16 (48.48)	11 (33.33)	16 (48.48)	13 (39.39)	4 (12.12)
	stage 4 (n = 54)	18 (33.33)	17 (31.48)	35 (64.81)	10 (18.52)	18 (33.33)	17 (31.48)	32 (59.26)	28 (51.85)	7 (12.96)
<i>MYCN</i> amplification status	<i>MYCN</i> 0 (n = 96)	11 (11.46)	11 (11.46)	39 (40.63)	5 (5.21)	12 (12.50)	13 (13.54)	29 (30.21)	28 (29.17)	2 (2.08)
	<i>MYCN</i> 1 (n = 36)	26 (72.22)	17 (47.22)	26 (72.22)	11 (30.56)	25 (69.44)	16 (44.44)	26 (72.22)	25 (69.44)	10 (27.78)
age at diagnosis	≤ 12 months (n = 54)	6 (11.11)	1 (1.85)	17 (31.48)	1 (1.85)	9 (16.67)	1 (1.85)	11 (20.37)	11 (20.37)	1 (1.85)
	> 12 months (n = 78)	31 (39.74)	27 (34.62)	48 (61.54)	15 (19.23)	28 (35.90)	28 (35.90)	44 (56.41)	42 (53.85)	11 (14.10)
	≤ 18 months (n = 63)	10 (15.87)	5 (7.94)	22 (34.92)	1 (1.59)	10 (15.87)	2 (3.17)	14 (22.22)	16 (25.40)	1 (1.59)
	> 18 months (n = 39)	27 (39.13)	23 (33.33)	43 (62.32)	15 (21.74)	27 (39.13)	27 (39.13)	41 (59.42)	37 (53.62)	11 (15.94)
factor - statistics (p)		assay006 (<i>SPRED3</i>)	assay008 (<i>TNFAIP2</i>)	assay011	assay062 (<i>NPM2</i>)	assay087 (<i>NXPH1</i>)	assay108 (<i>CYYR1</i>)	assay111 (<i>CCDC177</i>)	assay113 (<i>Inc-MRPL3-2</i>)	assay116 (<i>Inc-TREX1-1</i>)
INSS stage		0.002	< 0.001	0.003	0.056	< 0.001	< 0.001	< 0.001	0.085	0.302
<i>MYCN</i> amplification status		< 0.001	< 0.001	0.002	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
age at diagnosis (cutoff 12 months)		< 0.001	< 0.001	0.001	0.002	0.018	< 0.001	< 0.001	< 0.001	0.027
age at diagnosis (cutoff 18 months)		0.004	< 0.001	0.002	< 0.001	0.004	< 0.001	< 0.001	0.001	0.005

(continues)

Note. For each of the nine individual prognostic MSP assays the number (percentage) of methylated samples in each stratum of MSP cohort I is given. P-values are according to the Fisher's exact test. INSS: International Neuroblastoma Staging System; *MYCN*0: *MYCN* non-amplified; *MYCN*1: *MYCN* amplified.

Table 2. The nine individual prognostic MSP assays are differentially methylated between patient groups with distinct neuroblastoma risk factors (MSP cohort II).

(continued)

factor - number (percentage)		MSP cohort II								
		assay006 (<i>SPRED3</i>)	assay008 (<i>TNFAIP2</i>)	assay011	assay062 (<i>NPM2</i>)	assay087 (<i>NXPH1</i>)	assay108 (<i>CYYR1</i>)	assay111 (<i>CCDC177</i>)	assay113 (<i>Inc-MRPL3-2</i>)	assay116 (<i>Inc-TREX1-1</i>)
INSS stage	stage 1 (n = 27)	2 (7.41)	1 (3.70)	7 (25.93)	0 (0.00)	3 (11.11)	10 (37.04)	11 (40.74)	10 (37.04)	1 (3.70)
	stage 2 (n = 17)	4 (23.53)	3 (17.65)	8 (47.06)	0 (0.00)	4 (23.53)	5 (29.41)	6 (35.29)	4 (23.53)	0 (0.00)
	stage 3 (n = 27)	7 (25.93)	6 (22.22)	14 (51.85)	5 (18.52)	9 (33.33)	12 (44.44)	9 (33.33)	14 (51.85)	6 (22.22)
	stage 4 (n = 103)	41 (39.81)	35 (33.98)	83 (80.58)	24 (23.30)	48 (46.60)	50 (48.54)	66 (64.08)	56 (54.37)	16 (15.53)
MYCN amplification status	MYCN0 (n = 115)	11 (9.57)	14 (12.17)	63 (54.78)	11 (9.57)	22 (19.13)	39 (33.91)	47 (40.87)	43 (37.39)	7 (6.09)
	MYCN1 (n = 60)	43 (71.67)	30 (50.00)	48 (80.00)	18 (30.00)	40 (66.67)	37 (61.67)	43 (71.67)	40 (66.67)	16 (26.67)
age at diagnosis	≤ 12 months (n = 53)	1 (1.89)	1 (1.89)	14 (26.42)	0 (0.00)	8 (15.09)	12 (22.64)	10 (18.87)	11 (20.75)	3 (5.66)
	> 12 months (n = 124)	53 (42.74)	44 (35.48)	99 (79.84)	29 (23.39)	56 (45.16)	65 (52.42)	82 (66.13)	73 (58.87)	20 (16.13)
	≤ 18 months (n = 74)	10 (13.51)	4 (5.41)	27 (36.49)	2 (2.70)	14 (18.92)	23 (31.08)	22 (29.73)	22 (29.73)	8 (10.81)
	> 18 months (n = 103)	44 (42.72)	41 (39.81)	86 (83.50)	27 (26.21)	50 (48.54)	54 (52.43)	70 (67.96)	62 (60.19)	15 (14.56)
factor - statistics (p)		assay006 (<i>SPRED3</i>)	assay008 (<i>TNFAIP2</i>)	assay011	assay062 (<i>NPM2</i>)	assay087 (<i>NXPH1</i>)	assay108 (<i>CYYR1</i>)	assay111 (<i>CCDC177</i>)	assay113 (<i>Inc-MRPL3-2</i>)	assay116 (<i>Inc-TREX1-1</i>)
INSS stage		0.006	0.005	< 0.001	0.002	0.003	0.439	0.004	0.066	0.066
MYCN amplification status		< 0.001	< 0.001	0.001	0.001	< 0.001	0.001	< 0.001	< 0.001	< 0.001
age at diagnosis (cutoff 12 months)		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.085
age at diagnosis (cutoff 18 months)		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.006	< 0.001	< 0.001	0.505

Note. For each of the nine individual prognostic MSP assays the number (percentage) of methylated samples in each stratum of MSP cohort II is given. P-values are according to the Fisher's exact test. INSS: International Neuroblastoma Staging System; MYCN0: *MYCN* non-amplified; MYCN1: *MYCN* amplified.

located in the promoter region of *SPRED3*, *TNFAIP2*, *NPM2* and *CYYR1*. For MSP assay011, which has prognostic value for both OS and EFS, the amplicon is located on chr8:143 498 349 - 143 498 469 (flanking genes are *TSNARE1* (downstream on \pm 14 kb) and *BAI1* (upstream on \pm 47 kb on opposite strand)). The corresponding results of the univariable logistic regression analyses are also shown in Table 1, and associations between the prognostic DNA methylation biomarkers and established prognostic NB risk factors (*MYCN* amplification, age at diagnosis (both 12 and 18 month cutoff) and International Neuroblastoma Staging System (INSS) stage [19]) are shown in Table 2.

A 58-marker methylation signature with accompanying methylation score cutoff of 25% predicts overall and event-free survival

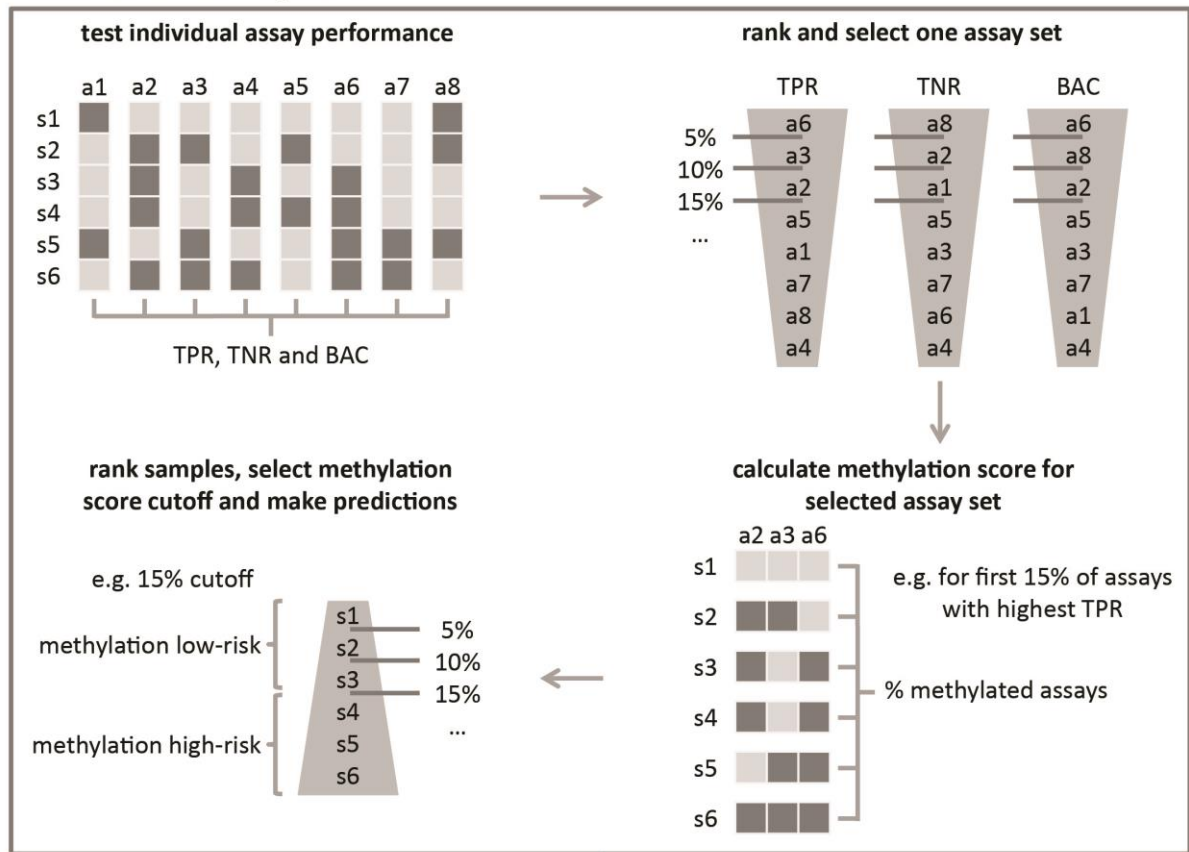
As all MSP assays were designed in regions identified as hypermethylated in HR-DOD or HR-MYCN1 samples and as the MSP data show association with outcome (Supplemental Figure 3A-3D), the possibility of establishing a robust and accurate multimarker signature for OS and EFS based on the number of methylation events was explored. To this purpose, a new statistical framework was developed, which allows identification of a robust set of MSP assays of which the methylation scores (i.e. the percentage of methylated assays in each sample) allow accurate outcome prediction (details in Materials and Methods and Figure 2). The signature was trained on MSP cohort I and tested on MSP cohort II. For the high-risk subgroups, the resulting signature was not prognostic, but using the entire sample cohorts, a set of 58 MSP assays (Supplemental Table 4) with a methylation score cutoff of 25% was put forward and shown to significantly predict OS ($p < 0.001$ for both cohorts, log-rank test) and EFS ($p = 0.001$ for MSP cohort I and $p < 0.001$ for MSP cohort II). For MSP cohort I, OS at 5 years of

follow-up is 80.14% (95% confidence interval (CI) 72.06 - 89.11) for the group of patients at methylation low-risk, compared to 47.74% (34.43 - 66.18) for the group of patients at methylation high-risk. The 5-year EFS is 80.54% (72.40 - 89.61) and 55.22% (40.92 - 74.51) in the methylation low- and high-risk groups, respectively. For MSP cohort II, OS at 5 years of follow-up is 86.67% (79.92 - 93.98) for the methylation low-risk group, compared to 44.20% (34.14 - 57.23) for the methylation high-risk group. Here, the 5-year EFS is 86.86% (79.86 - 94.47) and 53.34% (42.06 - 67.65) in the methylation low- and high-risk groups, respectively. The corresponding Kaplan-Meier curves are depicted in Figure 3. Power analyses using these survival rates illustrate that the MSP cohorts contain sufficient numbers of samples to obtain 90% power at 5% significance level. The signature has a balanced accuracy (BAC) of 70.12% for OS and 65.71% for EFS on MSP cohort I. On MSP cohort II, these values are 71.28% and 67.97%, respectively. Univariable logistic regression analyses also illustrate that the signature predicts OS and EFS, and multivariable logistic regression analyses show that the signature is a significantly independent predictor of OS in MSP cohort II after controlling for known risk factors (Supplemental Table 5). Associations between the signature predictions and established NB risk factors are shown in Table 3.

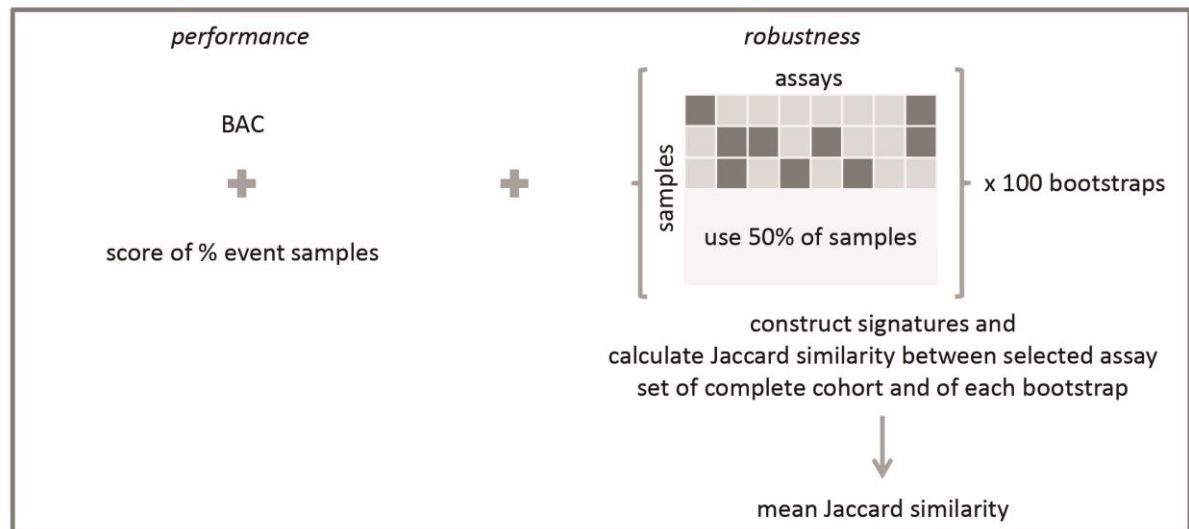
Discussion

MYCN amplification was identified as first genetic prognostic marker, in addition to age at diagnosis and tumor stage, which is still used today in therapeutic stratification [1]. Further studies have attempted to explore additional parameters to improve prognostic classification. Most notably, these include large chromosomal imbalances as well as transcriptome-based gene signatures. Given the low mutation burden,

1. construction of 1323 signatures



2. evaluation of performance and robustness of each constructed signature



3. signature selection

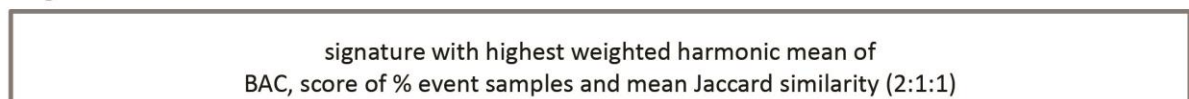


Figure 2. A new statistical framework was developed to identify a robust multimarker signature for accurate outcome prediction. The framework consists of three major steps: (1) signatures construction, (2) evaluation of the performance and robustness of the constructed signatures and (3) the selection of the final signature. Details of every step are described in the materials and methods section. a: assay; BAC: balanced accuracy; s: sample; TNR: true negative rate (specificity); TPR: true positive rate (sensitivity).

results

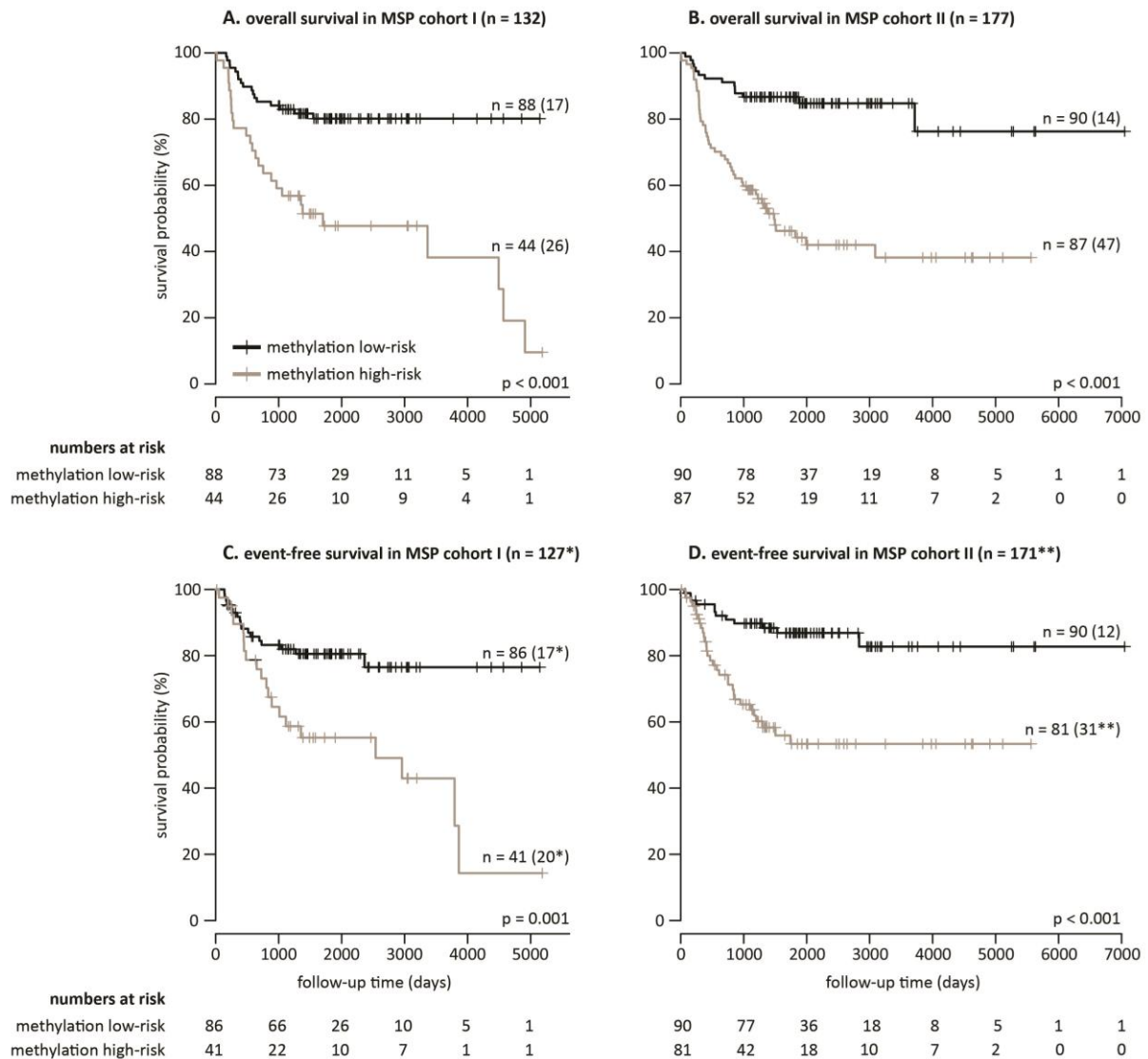


Figure 3. A robust 58-marker methylation signature and methylation score of 25% predicts overall and event-free survival. Kaplan-Meier curves and log-rank p-values for overall survival (MSP cohort I in A. and MSP cohort II in B.) and event-free survival (MSP cohort I in C. and MSP cohort II in D.) are shown. The numbers of patients at methylation low- and high-risk as predicted by the 58-marker signature are indicated. The numbers in parentheses in the plots refer to the number of patients that experienced an event (deceased of disease for overall survival, and relapse, progression or deceased of disease for event-free survival). *Missing follow-up time for two methylation low-risk patients and three methylation high-risk patients. **Missing follow-up time for five methylation high-risk patients, and event status and follow-up time for one patient.

more recent sequencing efforts did not deliver significant novel tools for prognostic stratification [20], although *ALK* mutation status is of importance for including patients for targeted therapy with *ALK* inhibitors. Recent studies have shown that NB biology is also strongly determined by the epigenetic profile of the tumor, which has paved the way for prognostic DNA methylation biomarker research. During the past years, multiple prognostic single-

gene methylation biomarkers have been described in NB; also a so-called CpG island methylator phenotype (CIMP) was found to be of prognostic value [6–16]. Here, we studied the NB methylome in a genome-wide manner to establish and validate novel prognostic biomarkers for OS and EFS.

Several features contribute to the novel and comprehensive aspect of our study. A first important feature is the number of analyzed

tumor samples. In total, 396 primary tumors were included, which is the largest series studied to date. Most reported studies only rely on NB cell lines or on a relatively limited number of tumors in the discovery phase and thus fall short in covering the NB heterogeneity, or lack independent validation on large sample cohorts. Of note, previous studies on mRNAs and microRNAs in NB have emphasized that biomarkers are of little or no utility if they are

not validated on an independent patient cohort [21,22]. Here, MBD sequencing was applied to 87 primary tumors, carefully selected for risk classification, allowing optimal biomarker discovery, and two independent cohorts of 132 and 177 primary tumors were used to test the selected candidate biomarkers. Power analyses further emphasize that these large sample collections result in adequate power of the study.

Table 3. The 58-marker signature predictions are associated with established neuroblastoma risk factors.

MSP cohort I			
factor - number (percentage)		OS signature prediction	EFS signature prediction
INSS stage	stage 1 (n = 27)	0 (0.00)	0 (0.00)
	stage 2 (n = 18)	1 (5.56)	1 (5.56)
	stage 3 (n = 33)	15 (45.45)	15 (45.45)
	stage 4 (n = 54)	28 (51.85)	28 (51.85)
MYCN amplification status	MYCN0 (n = 96)	17 (17.71)	17 (17.71)
	MYCN1 (n = 36)	27 (75.00)	27 (75.00)
age at diagnosis	≤ 12 months (n = 54)	5 (9.26)	5 (9.26)
	> 12 months (n = 78)	39 (50.00)	39 (50.00)
	≤ 18 months (n = 63)	8 (12.70)	8 (12.70)
	> 18 months (n = 69)	36 (52.17)	36 (52.17)
MSP cohort II			
factor - number (percentage)		OS signature prediction	EFS signature prediction
INSS stage	stage 1 (n = 27)	4 (14.81)	4 (14.81)
	stage 2 (n = 17)	4 (23.53)	4 (23.53)
	stage 3 (n = 27)	7 (25.93)	7 (25.93)
	stage 4 (n = 103*)	72 (69.90)	71 (69.61)
MYCN amplification status	MYCN0 (n = 115)	38 (33.04)	38 (33.04)
	MYCN1 (n = 60)*	47 (78.33)	46 (77.97)
age at diagnosis	≤ 12 months (n = 53)	1 (1.89)	1 (1.89)
	> 12 months (n = 124*)	86 (69.35)	85 (69.11)
	≤ 18 months (n = 74)	9 (12.16)	9 (12.16)
	> 18 months (n = 103*)	78 (75.73)	77 (75.49)

Note. For both OS and EFS, the number (percentage) of methylation high-risk samples in each stratum of MSP cohort I and II is given. All associations are statistically significant ($p < 0.001$; Fisher's exact test). *Missing EFS status for one patient. EFS: event-free survival; INSS: International Neuroblastoma Staging System; MYCN0: MYCN non-amplified; MYCN1: MYCN amplified; OS: overall survival.

Another important feature is that we made use of MBD sequencing of primary NB tumors in the discovery phase to identify novel biomarker candidates. Compared to the Illumina methylation arrays, which were previously applied to NB tumors, MBD sequencing interrogates more CpGs (approximately 18% of

all CpGs versus < 2% for the arrays [23]) and thus allows extension of the biomarker discovery phase to previously unexplored regions. MBD sequencing also has higher genomic coverage than methodologies based on antibodies (methylated DNA immunoprecipitation (MeDIP)) [24]. This genome-wide assessment of the DNA

methylation pattern is reflected in the final selection of MSP assays, as we have shown that most of the assays would not have been identified using the HM450 array and that a substantial part of the assays is located in non-promoter or non-coding regions. These findings support previous studies in other cancer types that show that it is important to extend the search for potentially clinical applicable DNA methylation biomarker to the entire methylome rather than focusing on promoter CpG islands of which methylation is in most cases inversely correlated to their transcriptional activity [5]. The prognostic relevance of the selected candidate biomarkers was further analyzed in two large independent cohorts using our previously established high-throughput and semi-automated MSP pipeline [10]. As these cohorts include a considerable number of both high-risk survivors and non-survivors, the candidates could not only be tested on the entire sample cohorts, but also on the high-risk cohorts only. This analysis is very valuable, for the reason that the need for prognostic biomarkers is the highest within this group of patients. However, although differential methylation analyses and hierarchical clustering on the MBD sequencing data illustrate that high-risk survivors and high-risk non-survivors show different methylation patterns, the MSP screens did not identify markers that were significantly prognostic in both MSP high-risk cohort I and II. Importantly, this does not mean that high-risk DNA methylation biomarkers cannot be found. It only indicates that the methylation differences in the DMRs (of 2 kb or 5 kb) in the MBD sequencing data between these high-risk groups are too subtle to be easily translated in an MSP assay which only interrogates a few CpGs. Therefore, the possibility of establishing high-risk methylation biomarkers based on genome-wide bisulfite sequencing, which allows analysis of the methylome at the single CpG level, should be addressed in the future. These future studies might also benefit from focusing on more

homogeneous high-risk patient groups, for example by only studying *MYCN* amplified or non-amplified samples, as the heterogeneity within our high-risk cohort might also have counteracted the possibility of establishing high-risk DNA methylation biomarkers.

Nevertheless, our validation efforts allowed robust identification of prognostic assays on the entire patient cohorts. Newly discovered individual prognostic methylation biomarkers for event-free survival (EFS) are *CCDC177* and *NXPH1*, and *SPRED3*, *TNFAIP2*, *NPM2* and *CYYR1* for overall survival (OS). Interestingly, some of these biomarkers are linked with neural processes and/or have already been described in other tumor types. For example, *NXPH1* encodes the neurexophilin 1 protein that forms a very tight complex with alpha neurexins, a group of proteins that promote adhesion between dendrites and axons, and methylation of this gene was previously described as potential diagnostic biomarker for breast cancer [26]. *TNFAIP2* (tumor necrosis factor, alpha-induced protein 2) was also found to be hypermethylated in colorectal cancer [27] and *NPM2* (nucleophosmin/ nucleoplasmin 2) in melanoma [28] and acute myeloid leukemia [29]. Alterations of sequence and expression of *CYYR1* (cysteine/tyrosine-rich 1) were previously observed in neuroendocrine tumors [30]. Remarkably, also three non-coding methylation biomarkers for OS and EFS were identified (*Inc-MRPL3-2*, *Inc-TREX1-1* and assay011). Assay011 is located on chr8:143 498 349 - 143 498 469, but further annotation is not available for this region. These findings again underscore the importance of screening the entire methylome for biomarker discovery. Of note, the role of methylation of these non-promoter CpGs in NB is currently unclear and should also be topic of further investigation, as it has been shown that DNA methylation outside promoters may also be crucial for gene regulation [25]. Clearly, this might reveal new aspects of NB tumorigenesis.

Finally, a new statistical framework was applied

to identify a robust set of MSP assays of which the methylation scores of the samples allow accurate outcome prediction. Both for OS and EFS, a 58-marker signature with a methylation score cutoff of 25% was selected based on the data of MSP cohort I. Survival analyses on both MSP cohort I and II indicate that the signature displays prognostic value for OS and EFS, and is a significant independent predictor of OS in MSP cohort II after controlling for established NB risk factors. All newly discovered individual prognostic methylation biomarkers are part of the signature and further inspection of the other assays included in the signature shows biomarkers previously described in other tumor types, as well as genes previously linked to NB, such as *NAV2*, which functions in axonal elongation and is required for all-trans retinoic acid to induce neurite outgrowth in human NB cells [31]. Also in this regard, the present study is unique, since combining multiple individual methylation assays into a single biomarker signature is not previously reported in NB, with the exception of testing the CpG island methylator phenotype (CIMP), but this assay panel was simply adopted from the colorectal cancer research field. Yet, it should be tested whether these established DNA methylation biomarkers can further improve the performance of our 58-marker signature.

In conclusion, the applications of DNA methylation biomarkers in cancer management are versatile and these should definitely be further explored in the context of NB. During the past decades, many efforts have been made to identify prognostic DNA methylation biomarkers for NB, but currently no such biomarkers have made it to the clinic, as they lack comprehensive validation. In our study, we performed genome-wide methylation profiling of primary NB tumors using MBD sequencing to discover novel prognostic methylation biomarkers and subsequently tested top candidates in two independent cohorts using MSP. As such, we comprised 396 patients in total, which greatly

increases the validity of the study and makes it the largest DNA methylation biomarker study in NB to date. We robustly identified several novel individual biomarkers for OS and EFS, and could develop a prognostic 58-marker signature of which a methylation score cutoff of 25% allows accurate outcome prediction in the total patient cohorts. Furthermore, on the validation cohort, this signature was an independent predictor of OS after controlling for known NB risk factors, clearly indicating its clinical relevance. As such, this study forms a solid basis for further investigation of our biomarkers and signature in NB subgroups which could not be robustly examined in our cohorts (low-risk non-survivors and more homogeneous high-risk subgroups). Ideally, also the integration with other DNA methylation biomarkers and -omic data should be further explored to fully optimize the assessment of NB prognosis and appropriate stratification of patient treatment.

Materials and methods

Neuroblastoma cell lines and primary tumors

In total, 437 primary NB tumor samples were used to establish four independent sample cohorts: MBD cohort I (n = 42), MBD cohort II (n = 45), MSP cohort I (n = 148) and MSP cohort II (n = 202). Also 19 NB cell lines (Supplemental Table 3) were included in the study. All primary tumor samples were assigned to one of three previously defined [10] risk groups based on NB risk parameters (INSS stage, *MYCN* amplification status and age of the patient at diagnosis) and disease outcome: (1) high-risk patients that died of disease (HR-DOD), (2) high-risk survivors (HR-SURV), or (3) low-risk survivors (LR-SURV). Samples were collected at the Centre Léon Bérard (n = 125, Lyon, France), the Hospital Clínico Universitario (n = 86; Valencia, Spain), the Ghent University Hospital (n = 80; Ghent, Belgium), the Sydney Children's Hospital (n = 48; Sydney, Australia), the Institut Curie (n = 37,

Paris, France), the Children's Cancer and Leukemia Group (n = 29, Leicester, UK), the Our Lady's Children's Hospital Dublin (n = 13; Dublin, Ireland), the University Hospital Brno (n = 11, Brno, Czech Republic) and the University Children's Hospital Essen (n = 8; Essen, Germany). Detailed clinical characteristics of the patients and a summary across the different subcohorts are given in Supplemental Table 1. The study was approved by the ethical committee of the Ghent University Hospital (approval number: B67020109912).

Methyl-CpG-binding domain sequencing

DNA fragmentation and MBD-based capturing of 42 (MBD cohort I) and 45 (MBD cohort II) samples were performed as described in [32] and Decock et al., in preparation. Briefly, 200-500 ng sheared DNA was used to enrich for methylated fragments using the MethylCap kit (MBD from MeCP2; Diagenode). For each captured fraction of the samples of MBD cohort I, DNA library preparation was performed using the NEBNext DNA Library Prep Master Mix Set for Illumina (New England Biolabs) in combination with the Multiplexing Sample Preparation Oligonucleotide Kit (Illumina) for paired-end adapter ligation. For the input and enriched fractions of the samples of MBD cohort II, library preparation was automated on an Apollo 324 Next Generation Sequencing Library Preparation System (IntegenX), making use of the PrepX ILM DNA Library Kit (IntegenX) in combination with the Multiplexing Sample Preparation Oligonucleotide Kit. Paired-end sequencing was performed on an Illumina GAIIx (MBD cohort I; PE 2 x 45 bp) and HiSeq2000 (MBD cohort II; PE 2 x 51 bp).

Methylation-specific PCR

Experimental MSP conditions and methylation calling were done as previously described [10] and are shown in Figure 1. Here, 78 technically validated MSP primer pairs (and the

methylation-independent *ACTB* control assay; Supplemental Table 2) were tested on amplified bisulfite-treated DNA from 19 NB cell lines and two independent cohorts of 148 (MSP cohort I) and 202 (MSP cohort II) patients, selected from the previously defined prognostic patient groups.

Bioinformatics and statistical analyses

Methyl-CpG-binding domain sequencing

Raw MBD sequencing data were demultiplexed and converted to FASTQ files. Quality control was performed by FastQC, followed by paired-end read mapping to the human reference genome (hg19) using Bowtie2 [33] and SAMtools [34]. PCR duplicates were marked by Picard and mapping quality control was done by SAMStat [35] and bamUtil. Peaks were called using MACS [36]. Data have been deposited into the Gene Expression Omnibus (GEO; GSE69224 and GSE69243). Count matrices for differential methylation analyses between the prognostic patient subgroups in DESeq [37] were constructed using the R ShortRead [38] and rtracklayer [39] packages. Here, for both MBD cohorts, two count datasets were constructed. The first one represents a table that reports for each MBD-enriched sample the number of mapped reads that are assigned to the promoter region (-1,500 bp to +500 bp around transcription start site (TSS)) of the different Ensembl Transcripts (release 68), and the second one to 5 kb genomic windows (2.5 kb overlapping moving windows). Hierarchical clustering was performed using the R gplots and RColorBrewer packages.

Methylation-specific PCR

For survival analyses on the MSP data, the Kaplan- Meier method was used to estimate overall and event-free survival (OS and EFS) probabilities, and survival functions were compared with the log-rank test (R survival package). OS time was defined as the time

between diagnosis and disease-related death or last follow-up. EFS time is the time between diagnosis and first occurrence of relapse, progression or death of disease, or last follow-up. P-values < 0.05 were considered statistically significant. All individual assays were tested, as well as a multimarker signature that was established on MSP cohort I using a new statistical framework (Figure 2 and Cannoodt et al., in preparation). This framework involves three major steps: (1) the construction of signatures, (2) the evaluation of the performance and robustness of each constructed signature, and (3) the signature selection. The construction of signatures (step 1) is based on the performance of the individual assays, which is evaluated by determining the following statistical metrics: sensitivity (true positive rate (TPR)), specificity (true negative rate (TNR)) and balanced accuracy (BAC). Each of these metrics was subsequently used to rank the assays (from highest to lowest value) and a cutoff, defined by percentiles of the ranked list (from 0% to 100% with 5% increment; 21 possible cutoffs), was applied to select a specific assay set. Then, the methylation score (i.e. the percentage of methylated assays) for each of the samples is calculated and used to rank the samples (from lowest to highest value). Again, a percentage cutoff is applied on the ranked list, which allows making risk predictions for each sample. Samples with a methylation score above the cutoff have a high risk. Samples with a methylation score below the cutoff have a low risk. Given the number of tested metrics to evaluate the individual assay performance (3 possibilities), the number of possible cutoffs to select a specific assay set (21 possibilities), and the number of possible methylation score cutoffs (21 possibilities), 1,323 signatures were constructed and further evaluated on their performance and robustness (step 2). The performance of the constructed signatures was examined by determining the BAC, as well as a score that reflects how well the percentage of

predicted samples with an event equals the true percentage of samples with an event (score of % event samples in Figure 2). The robustness of the constructed signatures was tested by performing 100 bootstraps, creating a subcohort containing half of the samples. For each of these 100 subcohorts, signatures were constructed as described above and for each combination of parameters the Jaccard similarity index [40] between the selected assay set on the entire cohort and the bootstrap cohort was computed. The robustness of the signature is then reflected in the mean Jaccard similarity of the 100 bootstraps (R caret package). In order to select a final signature (step 3), the performance and robustness metrics are combined in a weighted harmonic mean, and the signature with the highest value is retained. Also power analyses (SAS Power and Sample Size), and univariable and multivariable logistic regression analyses (R survival package) were performed. Included factors in the multivariable analyses are: the *MYCN* amplification status (*MYCN* amplified versus non-amplified as reference), age of the patient at diagnosis (> 18 months versus ≤ 18 months as reference [41]), INSS stage (stage 4 versus not stage 4 as reference) and the signature prediction (methylation high-risk versus methylation low-risk as reference). Associations between the prognostic DNA methylation biomarkers and established NB risk factors (*MYCN* amplification, age at diagnosis (cutoff of 12 months and 18 months) and INSS stage) were determined using Fisher's exact test.

Fundings

This study is supported by the Fournier-Majoie Foundation (FFM), the Belgian National Lottery, the Flemish League Against Cancer (VLK) and UGent GOA (grant number 01G01910). AD was a recipient of a grant of the Research Foundation Flanders (FWO) and an Emmanuel Van der Schueren research grant (VLK). RN was a recipient of grants (RD12/0036/0020 and

2014/01008) from Thematic Network of Cooperative Research on Cancer (RTICC), Institute of Health Carlos III (ISCIII) and European Regional Development Fund (ERDF), and Spanish Health Research Fund (FIS). MH is a recipient of grants from the National Health and Medical Research Council (Australia), and Cancer Institute NSW. In France, this study was supported by the Annenberg Foundation and the Ligue Nationale Contre le Cancer. Funding was also obtained from SiRIC/INCa (Grant INCa-DGOS-4654). The CCLG Tissue Bank is funded by Cancer Research UK and CCLG, and supported by contributing CCLG Centres, including members of the ECMC Paediatric Network. RLS was a recipient of grants from Science Foundation Ireland (07/IN.1/B1776), Children's Medical and Research Foundation, and National Institutes of Health (NIH; 5R01CA127496).

Acknowledgments

The authors thank NXTGNT for the sequencing services, Lise Van den Haute for the kind help in the MSP screens and MDxHealth for facilitating the Caliper LabChip GX experiments. The authors also thank Olivier Delattre, head of Unité de Génétique Somatique, Eve Lapouble, Institut Curie, Paris, and the UK CCLG Tissue Bank for access to samples.

Conflicts of interest

No potential conflicts of interest were disclosed.

Supplemental materials

Supplemental materials can be found at: [http://www.impactjournals.com/oncotarget/index.php?journal=oncotarget&page=article&op=view&path\[\]=6477&path\[\]=17887](http://www.impactjournals.com/oncotarget/index.php?journal=oncotarget&page=article&op=view&path[]=6477&path[]=17887)

Supplemental Figure 1. Hierarchical cluster analyses highlight the capability of the methyl-CpG-binding domain (MBD) sequencing analysis

strategy in identifying candidate biomarkers. A. Clustering using the top 500 hyper- and hypomethylated promoter regions in high-risk non-survivors (HR-DOD) compared to low-risk survivors (LR-SURV) in MBD cohort II. **B.** Clustering using the top 500 hyper- and hypomethylated promoter regions in high-risk non-survivors compared to high-risk survivors (HR-SURV) in MBD cohort II. **C.** Clustering using the top 500 hyper- and hypomethylated promoter regions in high-risk *MYCN* amplified (HR-MYCN1) samples compared to high-risk *MYCN* non-amplified samples (HR-MYCN0).

Supplemental Figure 2. Visualization of the MBD sequencing data of the *HNRNP1* promoter region allows identification of the most informative (discriminative) region for MSP assay design. The location of three different MSP assays in the *HNRNP1* promoter region is shown (blue bars in the upper panel), as well as the number of sequencing tags at each position for each primary tumor of MBD cohort II (lower panel). Assay 1 is located in the region that is most discriminative between high-risk non-survivors (red) and high-risk survivors (orange)/low-risk survivors (green), while assay 2 and 3 are located in less informative regions (with fuzzy methylation patterns).

Supplemental Figure 3. Methylation-specific PCR confirms the validity of methyl-CpG-binding domain sequencing in identifying candidate methylation markers. Number of methylation events of 68 MSP assays (designed in regions identified in the MBD sequencing data as being hypermethylated in non-survivors) in survivors (LR-SURV and HR-SURV) and non-survivors of MSP cohort I (**A.**; Mann-Whitney, $p = 0.001$) and MSP cohort II (**B.**; Mann-Whitney, $p = 0.001$). Number of methylation events of 23 MSP assays (designed in regions identified in the MBD sequencing data as being hypermethylated in high-risk *MYCN* amplified samples) in high-risk *MYCN* non-amplified and amplified samples of

MSP cohort I (C.; Mann-Whitney, $p < 0.001$) and MSP cohort II (D.; Mann-Whitney, $p < 0.001$).

Supplemental Figure 4. Visualization of the MBD sequencing data of the promoter region of *UHRF2* shows that the corresponding MSP assay would not have been identified using the HM450 array. The location of the in-house designed MSP assay (large blue bar in upper panel) and probes of the HM450 array (small blue bars in the upper panel) are shown, as well as the number of sequencing tags at each position for each primary tumor of MBD cohort II (lower panel). High-risk non-survivors are indicated in red, high-risk survivors in orange and low-risk survivors in green.

Supplemental Table 1. In total, 437 annotated primary neuroblastoma DNA samples were collected and assigned to a specific study subcohort. A. Detailed characteristics. Each sample is characterized by a unique patientID and is assigned to a prognostic risk group (LR-SURV, HR-SURV or HR-DOD) and subcohort (MBD cohort I, MBD cohort II, MSP cohort I or MSP cohort II). Clinical characteristics given are the age at diagnosis in months, International Neuroblastoma Staging System (INSS) stage, *MYCN* amplification status (0 is non-amplified and 1 is amplified), and overall survival (OS) and event-free survival (EFS) status and time after diagnosis in days. The OS status indicates whether the patient was alive (0) at the last known follow-up or died of disease (1). Similarly, the EFS status indicates events such as relapse, progression or death. NAs represent missing values. LR-SURV: low-risk survivors, HR-DOD: high-risk deceased patients, HR-SURV: high-risk survivors. **B.** Summary. Per subcohort an overview of the clinical characteristics is given. *Only samples with a positive *ACTB* call were included, as only these were used in the analyses.

Supplemental Table 2. Seventy-eight MSP

assays were designed, technically validated and tested for overall and event-free survival prediction. For each assay, if available, the gene annotation is shown, as well as the forward and reverse primer (5' to 3'), and the genomic location of the amplicon on the hg19 reference genome. The corresponding region of interest in the methyl-CpG-binding domain sequencing data and the comparisons in which the region was identified as differentially methylated (indicated by yes, followed by the group that is hypermethylated) are indicated, as well as log-rank p-values for overall and event-free survival in the corresponding test cohort. Furthermore, it is shown whether the assay is part of the 58-marker signature or not.

Supplemental Table 3. Nineteen neuroblastoma cell lines are included in the study. For each cell line, the *MYCN* amplification status is shown, as well as its corresponding number in the MSP screen.

Supplemental Table 4. The LabChip GX size and height, and LC480 Cq and Tm value were combined to construct a dichotomous calling matrix. A. Results on MSP cohort I. For each assay and sample, the methylation call (dark blue (1) is methylated, yellow (0) unmethylated) is given. The patient samples are subdivided into three prognostic groups (LR-SURV, HR-DOD and HR-SURV). LR-SURV: low-risk survivors, HR-DOD: high-risk deceased patients, HR-SURV: high-risk survivors, CL: cell line, U-HCT: negative control (HCT-116 DKO cell line), M-HCT: positive control (*in vitro* methylated HCT-116 DKO cell line) and NTC: no template control. **B.** Results on MSP cohort II. **C.** Summary on MSP cohort I and MSP cohort II (only *ACTB* positive samples are taken into account). The number and percentage of methylated samples for a particular MSP assay within each prognostic group and for the entire sample cohorts is given.

Supplemental Table 5. The 58-marker

methylation signature is an independent prognostic predictor of overall survival in MSP cohort II. For each variable, the p-value, odds ratio (OR) and 95% confidence interval (CI) of the univariable and multivariable logistic regression analyses are shown. The age at diagnosis cutoff is 18 months.

Editorial note

This paper has been accepted based in part on peer-review conducted by another journal and the authors' response and revisions as well as expedited peer-review in Oncotarget.

References

- [1] Cohn SL et al. (2009). The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. *Journal of Clinical Oncology*; 27(2):289-297.
- [2] Park JR et al. (2013). Children's Oncology Group's 2013 blueprint for research: neuroblastoma. *Pediatric Blood and Cancer*; 60(6):985-993.
- [3] Berthold F et al. (2003). Long-term results and risk profiles of patients in five consecutive trials (1979-1997) with stage 4 neuroblastoma over 1 year of age. *Cancer Letters*; 197(1-2):11-17.
- [4] Pearson ADJ et al. (2008). High-dose rapid and standard induction chemotherapy for patients aged over 1 year with stage 4 neuroblastoma: a randomised trial. *Lancet Oncology*; 9(3):247-256.
- [5] How Kit A et al. (2012). DNA methylation based biomarkers: practical considerations and applications. *Biochimie*; 94(11):2314-2337.
- [6] Misawa A et al. (2009). RASSF1A hypermethylation in pretreatment serum DNA of neuroblastoma patients: a prognostic marker. *British Journal of Cancer*; 100(2):399-404.
- [7] Carén H et al. (2011). Identification of epigenetically regulated genes that predict patient outcome in neuroblastoma. *BMC Cancer*; 11:66.
- [8] Grau E et al. (2011). Hypermethylation of apoptotic genes as independent prognostic factor in neuroblastoma disease. *Molecular Carcinogenesis*; 50(3):153-162.
- [9] Decock A et al. (2011). Neuroblastoma epigenetics: from candidate gene approaches to genome-wide screenings. *Epigenetics*; 6(8):962-970.
- [10] Decock A et al. (2012). Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. *Genome Biology*; 13(10):R95.
- [11] Yáñez Y et al. (2015). Two independent epigenetic biomarkers predict survival in neuroblastoma. *Clinical Epigenetics*; 7(1):16.
- [12] Abe M et al. (2005). CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas. *Cancer Research*; 65(3):828-834.
- [13] Abe M et al. (2007). Marked and independent prognostic significance of the CpG island methylator phenotype in neuroblastomas. *Cancer Letters*; 247(2):253-258.
- [14] Abe M et al. (2008). Identification of genes targeted by CpG island methylator phenotype in neuroblastomas, and their possible integrative involvement in poor prognosis. *Oncology*; 74(1-2):50-60.
- [15] Banelli B et al. (2012). A pyrosequencing assay for the quantitative methylation analysis of the PCDHB gene cluster, the major factor in neuroblastoma methylator phenotype. *Laboratory Investigation*; 92(3):458-465.
- [16] Banelli B et al. (2013). Clinical potentials of methylator phenotype in stage 4 high-risk neuroblastoma: an open challenge. *PLoS One*; 8(5):e63253.
- [17] Xiao Y et al. (2014). A novel significance score for gene selection and ranking. *Bioinformatics*; 30(6):801-807.
- [18] Huss M (2010). Introduction into the analysis of high-throughput-sequencing based epigenome data. *Briefings in Bioinformatics*; 11(5):512-523.
- [19] Brodeur GM et al. (1993). Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *Journal of Clinical Oncology*; 11(8):1466-1477.
- [20] Molenaar JJ et al. (2012). Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature*; 483(7391):589-593.
- [21] De Preter K et al. (2011). miRNA expression profiling enables risk stratification in archived and fresh neuroblastoma tumor samples. *Clinical Cancer Research*; 17(24):7684-7692.

- [22] Vermeulen J et al. (2009). Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/ COG/GPOH study. *Lancet Oncology*; 10(7):663-671.
- [23] Stirzaker C et al. (2014). Mining cancer methylomes: prospects and challenges. *Trends in Genetics*; 30(2):75-84.
- [24] Bock C et al. (2010). Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nature Biotechnology*; 28(10):1106-1114.
- [25] Lou S et al. (2014). Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome Biology*; 15(7):408.
- [26] Faryna M et al. (2012). Genome-wide methylation screen in low-grade breast cancer identifies novel epigenetically altered genes as potential biomarkers for tumor diagnosis. *FASEB Journal*; 26(12):4937-4950.
- [27] Ashktorab H et al. (2014). DNA methylome profiling identifies novel methylated genes in African American patients with colorectal neoplasia. *Epigenetics*; 9(4):503-512.
- [28] Koga Y et al. (2009). Genome-wide screen of promoter methylation identifies novel markers in melanoma. *Genome Research*; 19(8):1462-1470.
- [29] Kroeger H et al. (2008). Aberrant CpG island methylation in acute myeloid leukemia is accentuated at relapse. *Blood*; 112(4):1366-1373.
- [30] Vitale L et al. (2007). Sequence, "subtle" alternative splicing and expression of the CYR1 (cysteine/tyrosine-rich 1) mRNA in human neuroendocrine tumors. *BMC Cancer*; 7:66.
- [31] Muley PD et al. (2008). The atRA-responsive gene neuron navigator 2 functions in neurite outgrowth and axonal elongation. *Developmental Neurobiology*; 68(13):1441-1453.
- [32] De Meyer T et al. (2013). Quality evaluation of methyl binding domain based kits for enrichment DNA-methylation sequencing. *PloS One*; 8(3):e59068.
- [33] Langmead B et al. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*; 9(4):357-359.
- [34] Li H et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*; 25(16):2078-2079.
- [35] Lassmann T et al. (2011). SAMStat: monitoring biases in next generation sequencing data. *Bioinformatics*; 27(1):130-131.
- [36] Zhang Y et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*; 9(9):R137.
- [37] Anders S et al. (2010). Differential expression analysis for sequence count data. *Genome Biology*; 11(10):R106.
- [38] Morgan M et al. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics*; 25(19):2607-2608.
- [39] Lawrence M et al. (2009). rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics*; 25(14):1841-1842.
- [40] Liben-Nowell D et al. (2007). The link prediction problem for social networks. *Journal of the American Society for Information Science and Technology*; 58(7):1019-1031.
- [41] London WB et al. (2005). Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's Oncology Group. *Journal of Clinical Oncology*; 23(27):6459-6465.

paper 3

DNA methylation profiling of primary neuroblastoma tumors using methyl-CpG-binding domain sequencing

Anneleen Decock, Maté Ongenaert, Wim Van Criekinge, Frank Speleman and Jo Vandesompele

published in Scientific Data 3:160004, 2016

impact factor: not yet available

times cited (d.d. 11/09/2016; Google Scholar): 2

PAPER 3: DNA METHYLATION PROFILING OF PRIMARY NEUROBLASTOMA TUMORS USING METHYL-CpG-BINDING DOMAIN SEQUENCING

Anneleen Decock^{1,2}, Maté Ongenaert¹, Wim Van Criekinge^{3,4,5}, Frank Speleman^{1,2} and Jo Vandesompele^{1,2,6}

Correspondence to joke.vandesompele@ugent.be.

¹Center for Medical Genetics, Ghent University Hospital, De Pintelaan 185, Ghent, Belgium. ²Cancer Research Institute Ghent (CRIG), De Pintelaan 185, Ghent, Belgium. ³Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure Links 653, Ghent, Belgium. ⁴MDxHealth, 15279 Alton Parkway, Suite 100, Irvine, CA, USA. ⁵NXTGNT, Ghent University, Ottergemsesteenweg 460, Ghent, Belgium. ⁶Bioinformatics Institute Ghent - From Nucleotides to Networks (BIG N2N), De Pintelaan 185, Ghent, Belgium.

Contribution of AD: AD drafted the article, collected samples, executed DNA shearing for MBD sequencing, generated the data files for MBD sequencing visualization, and performed the differential methylation and technical validation analyses.

Abstract

Comprehensive genome-wide DNA methylation studies in neuroblastoma (NB), a childhood tumor that originates from precursor cells of the sympathetic nervous system, are scarce. Recently, we profiled the DNA methylome of 102 well-annotated primary NB tumors by methyl-CpG-binding domain (MBD) sequencing, in order to identify prognostic biomarker candidates. In this data descriptor, we give details on how this data set was generated and which bioinformatics analyses were applied during data processing. Through a series of technical validations, we illustrate that the data are of high quality and that the sequenced fragments represent methylated genomic regions. Furthermore, genes previously described to be methylated in NB are confirmed. As such, these MBD sequencing data are a valuable resource to further study the association of NB risk factors with the NB methylome, and offer the opportunity to integrate methylome data with other -omic data sets on the same tumor samples such as gene copy number and gene expression, also publically available.

Subject categories: DNA methylation, pediatric cancer, next-generation sequencing

Background and summary

Neuroblastoma (NB), a neuro-ectodermal tumor that originates from precursor cells of the sympathetic nervous system, represents the most common extra-cranial solid tumor of early childhood and is considered a heterogeneous disease driven by genetic aberrations, as during the past decades mainly genetic factors have

been described to influence the pathogenesis and disease course (including *MYCN* amplification, *ALK* amplification and mutation, hyperdiploidy, and gains and losses of specific chromosome arms (1p, 3p, 11q and 17q)) [1]. Also, recent comprehensive whole-genome sequencing studies of primary NB tumors pinpointed chromothripsis and defects in neuritogenesis genes as important tumor-driving

events in a subset of NB [2], and indicated that *MYCN*, *TERT* and *ATRX* alterations define major subgroups of high-risk NB [3,4]. However, also epigenetic mechanisms, such as DNA methylation alterations, seem to contribute to the NB biology and clinical behavior.

As reviewed in Decock et al. [5], multiple DNA methylation alterations have been described in NB, but given the rare occurrence of the disease, the number of comprehensive genome-wide DNA methylation studies analyzing primary tumor samples is limited. Hence, most studies initially make use of NB cell lines and only validate the most obvious methylation alterations in primary NB tumors. For example, a frequently applied methodology to NB cell lines is assessment of gene expression reactivation upon 5'-aza-2'-deoxycytidine (DAC) treatment, a cytosine analogue that cannot be methylated, leading to progressive DNA demethylation upon cell division. However, major drawbacks of these studies are that their discovery phases fall short in covering the NB heterogeneity, as NB cell lines are considered models for aggressive high-risk tumors, and that DNA methylation detection is indirectly assessed, as the influence of the demethylating effect is measured at the transcriptional level [6–8]. To accommodate this, the Illumina 27 and 450 K methylation arrays, directly interrogating the status of approximately 27,000 and 485,000 methylation sites, respectively, recently were applied to primary NB tumors [6,9–12]. Yet, also this technology has important limitations: the design of the arrays is heavily biased to interrogation of CpG sites previously described in literature and covers less than 2% of all CpG sites in the human genome [13].

Therefore, we generated a data set comprising of 102 primary NB tumors in which DNA methylation is assessed by massively parallel sequencing of methylation enriched DNA fragments. The applied method is based on the use of MeCP2, a member of the methyl-CpG-binding domain (MBD) protein family which

specifically binds to methylated cytosines and enables precipitation of methylated DNA fragments. This data set is unique in the NB research field, as it is the first sample cohort in which the full tumor heterogeneity is being assessed by genome-wide methylation analysis using next-generation sequencing (NGS); it was originally collected for the identification of prognostic biomarker candidates. Selected candidates were validated in independent cohorts using methylation-specific PCR and we showed that MBD sequencing allowed selection of valuable markers which would not have been identified using the Illumina methylation arrays [14].

Here, we provide a detailed description of the methodological approach and bioinformatics analyses, as well as easy access to the (analyzed) MBD sequencing data and analysis tools, allowing other researchers (inexperienced with MBD sequencing) to reuse it. Importantly, the analyzed samples are well annotated; besides overall and event-free survival data, also following NB characteristics are available: age of the patient at diagnosis, tumor stage according to the International Neuroblastoma Staging System (INSS) [15] and *MYCN* amplification status. As such, these data offer the opportunity to further explore the association of these risk factors with the NB methylome. Furthermore, integration of methylome data with other -omic data sets should be examined in order to fully map the NB biology on a genome-wide level. The present MBD sequencing data greatly facilitate these integration analyses, considering that for part of the profiled samples matching expression and array comparative genomic hybridization (aCGH) data are available [16–18] (see Methods for details).

In summary, this data descriptor outlines details on the generation and analysis of MBD sequencing data of 102 primary NB tumors (Figure 1). As NB is a rare disease and comprehensive DNA methylation studies scarce, these MBD sequencing data are very valuable

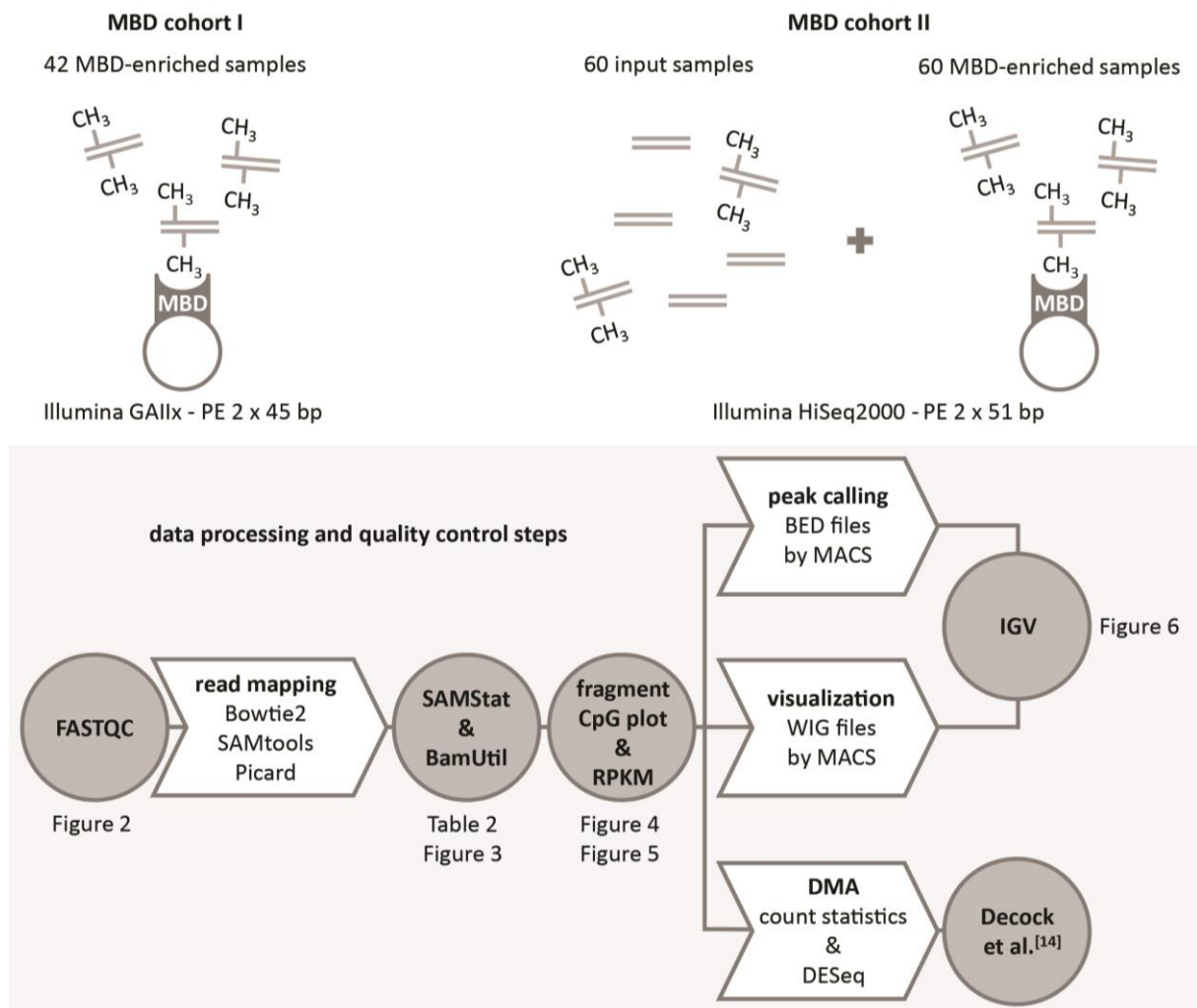


Figure 1. The MBD sequencing data of 102 primary neuroblastoma tumors are processed using different analysis tools. Depicted are the available methyl-CpG-binding domain (MBD) sequencing data sets and downstream data processing and technical validation steps. These steps are represented as arrows and circles, respectively. For each step, the applied tool or analysis is indicated. For the technical validation steps, also the corresponding data descriptor figures and tables are indicated. DMA: differential methylation analysis; IGV: Integrative Genomics Viewer; PE: paired-end; RPKM: reads per kilobase CpG island per million.

and permit further unravelling the role of DNA methylation in the NB biology.

Methods

DNA sample collection

Two independent cohorts of 42 and 60 primary tumor DNA samples, respectively annotated as MBD cohort I and II, were sequenced. Samples of fresh frozen tumors were collected at the Ghent University Hospital (n =49; Ghent, Belgium), the Hospital Clínico Universitario (n=42; Valencia, Spain), the University Children's

Hospital Essen (n =8; Essen, Germany) and the Our Lady's Children's Hospital Dublin (n=3; Dublin, Ireland), according to previously published criteria [7,14], and stage 4S tumors were also included. Detailed clinical characteristics of the patients are given in Table 1 (available online only). For samples 809 and 912, DNA was extracted from different parts of the same primary tumor. Informed consent was obtained from each patient's guardian and the study was approved by the ethical committee of the Ghent University Hospital (approval number B67020109912). Matching expression data [16,17] of 38 tumors are available through the

NCBI Gene Expression Omnibus (GEO) database (GSE21713 and GSE32664; sample IDs in Table 1 (available online only)). Matching aCGH data [18] of 38 tumors are available through ViVar [19] (<https://www.cmgg.be/vivar/>; login: review, password: review, project: Kumps et al. 2013; sample IDs in Table 1 (available online only)).

Methyl-CpG-binding domain (MBD) sequencing

DNA fragmentation

For each sample, between 400 to 1,000 ng DNA was sheared to obtain DNA fragments with an average length of 200 bp. The DNA was loaded in 120 µl TE buffer (1:5), transferred to a Snap Cap microTUBE (Covaris) and exposed to Covaris S2 Adaptive Focused Acoustics. Fragment distribution and concentration was determined on a High Sensitivity DNA chip (Agilent Technologies).

Methylated DNA capturing

Subsequently, capturing of methylated DNA fragments was done according to the MethylCap kit protocol of Diagenode using 200-500 ng DNA. Elution of the captured fraction was performed in 150 µl High Elution Buffer and DNA was purified using the MinElute PCR purification kit (Qiagen). For MBD cohort II, also input samples (10%) were prepared.

Library preparation

As MBD cohort I and II were profiled in a different time frame and NGS methodologies evolve at rapid pace, a different library preparation protocol and sequencing technology was applied to each of them. For MBD cohort I, DNA library preparation was performed using the NEBNext DNA Library Prep Master Mix Set for Illumina (New England Biolabs) in combination with the Multiplexing Sample Preparation Oligonucleotide Kit (Illumina) for paired-end adapter ligation. Size selection of the library is done on a 2% agarose gel (Bio-Rad). Fragments between 250 and 350 bp were

excised and purified using a Qiagen Gel Extraction Kit. For MBD cohort II, library preparation was automated on an Apollo 324 Next Generation Sequencing Library Preparation System (IntegenX), making use of the PrepX ILM DNA Library Kit (IntegenX). For paired-end adapter ligation the Multiplexing Sample Preparation Oligonucleotide Kit was used. Size selection was done with 1X AMPure XP beads (Agencourt) and PEG-Bead Solution.

Library amplification

PCR library amplification with appropriate Index Primers for each sample was performed using the Multiplexing Sample Preparation Oligonucleotide Kit and following PCR conditions: 30 s at 98 °C, 21 amplification cycles (10 s at 98 °C, 30 s at 65 °C and 30 s at 72 °C), 5 min at 72 °C, and held at 4 °C. PCR product purification was done using the High Pure PCR Purification Kit (Roche). QC was performed on a DNA 1000 chip (Agilent) and concentration was determined by qPCR according to the qPCR Quantification Protocol Guide of Illumina. Samples were pooled and profiled on an Illumina GAIIx (PE 2 × 45 bp) for MBD cohort I and on an Illumina HiSeq2000 (PE 2 × 51 bp) for MBD cohort II.

Data processing and analysis

Sequencing data

All crucial steps in the processing and analysis of the MBD sequencing data are summarized in Figure 1. Raw sequencing data were demultiplexed and converted to FASTQ files (with sequencing reads and quality scores). Quality control on the raw data was performed by FASTQC (version 0.9.2; <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Read mapping

Next, the sequencing reads were mapped/aligned to the human reference

genome (hg19), using the Bowtie2 [20] mapper (version 2.0.0 beta7) and FASTQ files as input. For each sample, two paired FASTQ files are available (as we performed paired-end sequencing), in which the data lines correspond to each other. To improve the mapping quality, reads were only taken into account if the sequences in both files could be mapped to the reference genome (maximum 500 bp between both paired ends). Also sequencing quality scores were used in the mapping process. The BAM format was used as output file type. PCR duplicates were marked with Picard (version 1.79; <http://broadinstitute.github.io/picard/>) and the BAM files were sorted and indexed using SAMtools [21] (version 0.1.18) and index commands. These files have been deposited as raw data files in the NCBI Gene Expression Omnibus (GEO) database (Data Citation 1 for MBD cohort I; Data Citation 2 and Data Citation 3 for MBD cohort II). FASTQ records can be extracted from the sequence alignments in the BAM files using the BEDTools bamtofastq conversion utility [22]. Starting from the SRA files, the NCBI SRA Toolkit (fastq-dump) can be used to generate the FASTQ files. Mapping quality was evaluated using SAMStat [23] (version 1.08) and BamUtil (version 1.0.2; <http://genome.sph.umich.edu/wiki/BamUtil>). Technical validation of MBD enrichment is performed by fragment CpG plot analysis [24] and by plotting the densities of the median numbers of mapped reads per kilobase per million (RPKM [25]) in all CpG islands (n=28,691) across the different subcohorts.

Peak calling

The process of converting mapped sequencing reads to coverage vectors and the detection of enriched regions (peaks) is referred to as peak detection or peak calling. Here, peak calling was done using the MACS [26] software tool (version 1.4.0 beta) and BAM files as input. BED files were generated (Data Citation 1 for MBD cohort I; Data Citation 2 and Data Citation 3 for MBD

cohort II), indicating the location and score (linked to the p-value) of the identified peaks.

Visualization

MACS is also used to output WIG files (Data Citation 1 for MBD cohort I; Data Citation 2 and Data Citation 3 for MBD cohort II), which are transformed to a binary format (TDF file; Data Citation 1 for MBD cohort I; Data Citation 2 and Data Citation 3 for MBD cohort II) by igvtools (<https://www.broadinstitute.org/igv/igvtools>) for visualization in the Integrative Genomics Viewer (IGV) [27]. An example IGV XML-session file for MBD cohort II and instructions on how to make use of this file are included in the GitHub repository (see Code availability).

Differential methylation analyses

Differential methylation analyses between sample groups are described in detail in Decock et al. [14]. Briefly, for each subcohort, two count data sets were constructed, in which for each sample the numbers of mapped reads in the promoter region of the different Ensembl Transcripts or 5 kb genomic windows are indicated. Here, we provide access to these count data sets (Supplementary Tables 3-8), which can directly be used for differential methylation analyses in DESeq [14,28].

Code availability

All tools and code that are necessary to generate the described file types are provided in a Docker container (Docker Hub; <https://hub.docker.com/r/mateongenaert/mbdttoolbox/>). More advanced analysis scripts can be found in the GitHub repository (<https://github.com/mateongenaert/MBDToolBox>).

Data records

An overview of the sample annotation and data outputs is given in Table 1 (available online only). The outputs of each step in the data

processing (read mapping: BAM files, peak calling: BED files, and visualization: WIG and TDF files) have been deposited in the GEO database. For MBD cohort I, the accession number is GSE69224 (Data Citation 1), for MBD cohort II, GSE69243 (Data Citation 2) and GSE69268 (Data Citation 3). In GEO, these data sets were submitted as SubSeries of the SuperSeries

GSE69279 (Data Citation 4). We also provide a Docker container, made available through Docker Hub, that embeds all necessary tools to generate the data files and illustrates the analysis pipeline. More advanced analysis scripts are given in the GitHub repository (see Code availability).

Table 2. Using BamUtil, basic sequencing statistics of MBD cohort I and II are computed.

statistic	MBD cohort I enriched samples			MBD cohort II enriched samples			MBD cohort II input samples		
	range	mean	median	range	mean	median	range	mean	median
total read number (e6)	4.65 - 18.20	13.38	14.17	29.74 - 66.59	45.09	44.41	20.86 - 59.51	36.00	33.19
duplicate reads (%)	0.70 - 72.00	6.46	3.39	2.55 - 79.69	31.04	19.89	2.24 - 10.47	4.17	3.68
properly paired reads (%)	48.29 - 94.51	85.64	89.29	86.86 - 97.57	95.33	95.72	94.78 - 97.55	96.50	96.59

Note. Total read number: the total number of reads in the two paired FASTQ files of a sample; duplicate reads as a percentage of the total read number; properly paired reads as a percentage of the total read number.

Technical validation

Validation of raw and mapped sequencing data

The total read number and percentage of duplicate and properly paired reads in each sample are given in Supplementary Table 1, and a summary of these sequencing statistics across the different sample cohorts can be found in Table 2.

To ensure raw data quality, FASTQC analyses were performed to determine the per base sequence quality which reflects the probability that a base has been called incorrectly [29]. Quality scores between 41 and 28, 28 and 20, and below 20 are considered base calls of very good quality, calls of reasonable quality and calls of poor quality, respectively. In order to obtain a general overview of the range of quality values across all bases at each position, the median quality score for each position in each FASTQ file was determined. Figure 2 shows the distribution of these median per base quality scores across the different sample cohorts. In general, the

quality scores of both MBD cohort I and II are of reasonable to very good quality. Given the different sequencing technologies that were used for MBD cohort I (Illumina GAIIX) and II (Illumina HiSeq2000), it is expected that the read quality of MBD cohort II is higher than that of MBD cohort I. The steadily increase and subsequent decrease in quality along the read is also expected for Illumina-based experiments [29,30].

Mapping quality is ensured by analyzing the mapping quality scores of the alignments in each sample (Supplementary Table 2). In Figure 3, the distributions of the percentages of mapped reads across the different mapping quality ranges are shown. For all subcohorts, the reads are clearly mapped with high accuracy, as almost for every sample, more than half of the mapped reads has a MAPQ ≥ 30 [23].

Validation of MBD-based enrichment

Over the past years several companies developed commercial kits for MBD-based

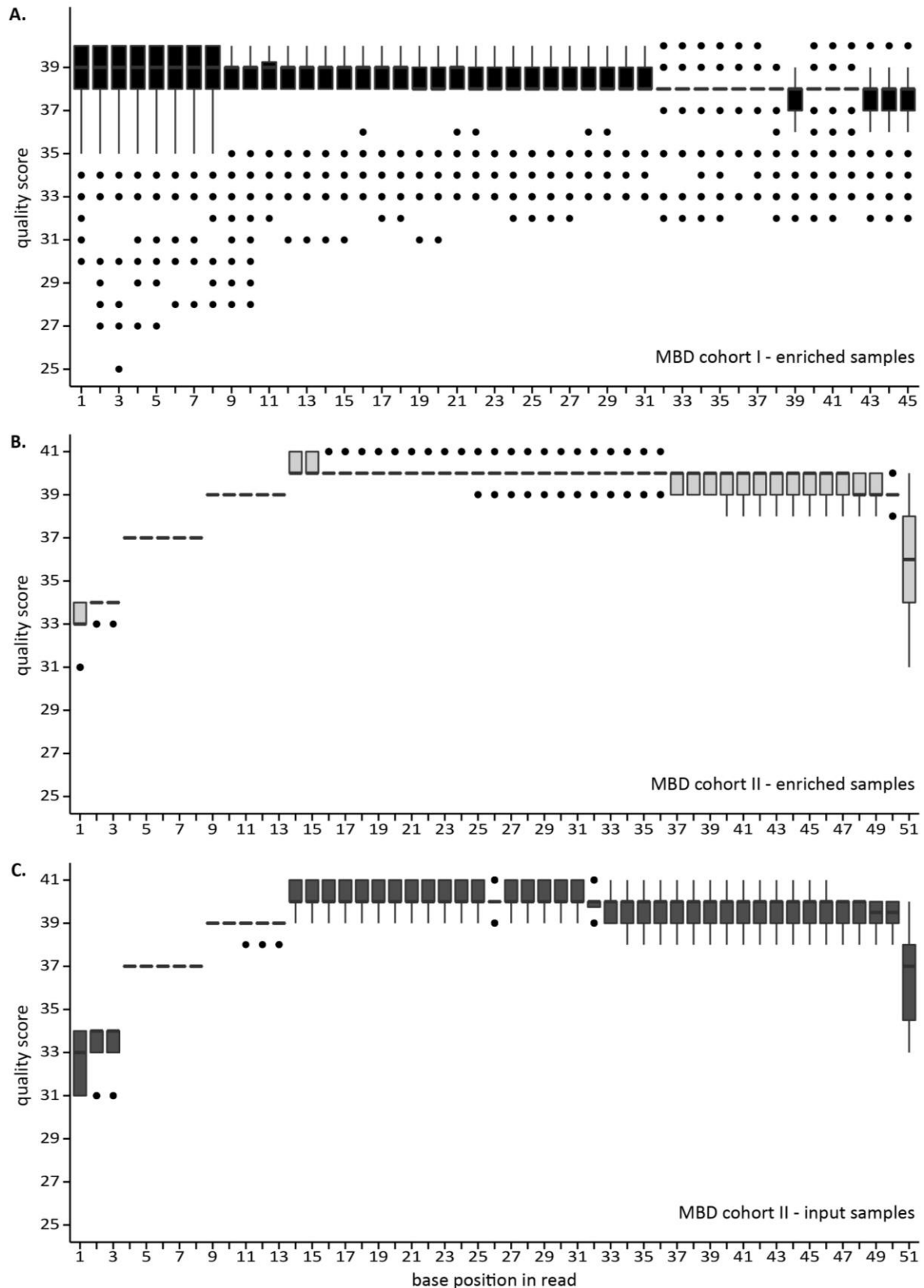


Figure 2. The per base sequence quality scores indicate that the raw sequencing data are of good quality. Shown are the distributions of the median per base quality score (determined by FASTQC) of the enriched samples of MBD cohort I (A.), and of the enriched (B.) and input (C.) samples of MBD cohort II. In the boxplots, the lower and upper hinge of the boxes represents the 25th and 75th percentile, respectively. The whiskers extend to the lowest and highest value that is within 1.5 times the interquartile range. Data beyond the end of the whiskers are outliers and plotted as dots.

results

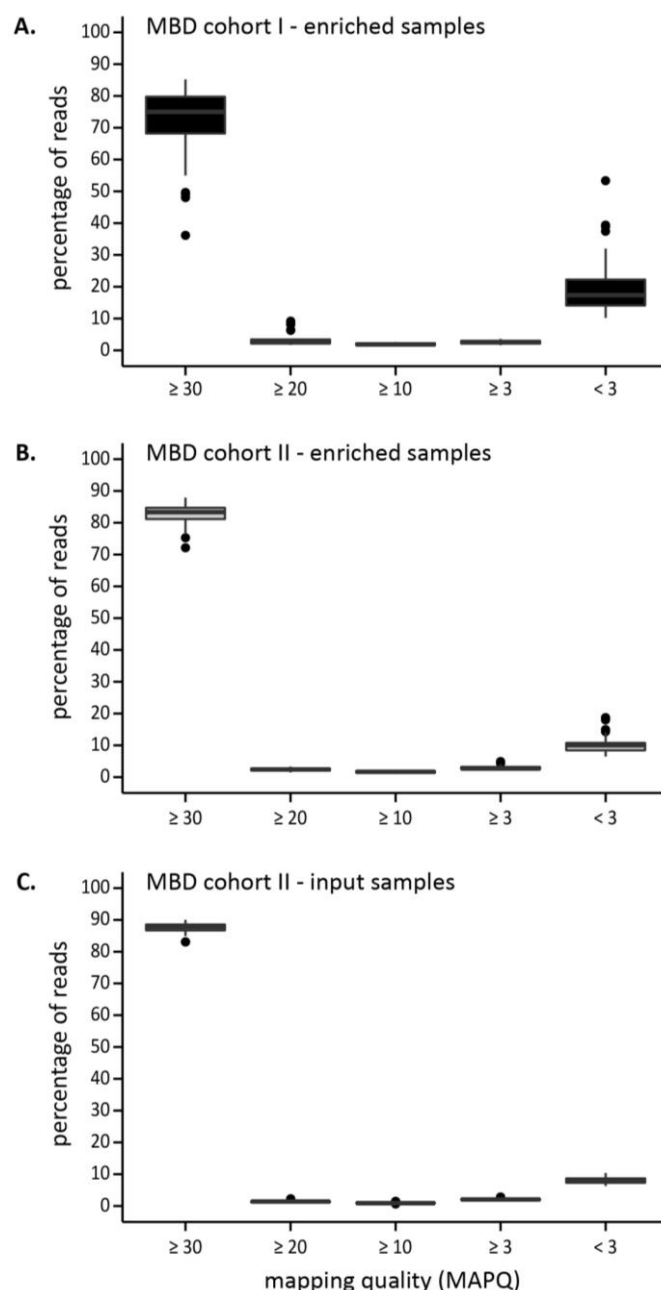


Figure 3. The mapping quality scores illustrate high mapping accuracy. Shown are the distributions of the percentages of mapped reads across the different mapping quality ranges, as determined by SAMStat (**A.** enriched samples of MBD cohort I, **B.** enriched samples of MBD cohort II and **C.** input samples of MBD cohort II). In the boxplots, the lower and upper hinge of the boxes represents the 25th and 75th percentile, respectively. The whiskers extend to the lowest and highest value that is within 1.5 times the interquartile range. Data beyond the end of the whiskers are outliers and plotted as dots.

capturing of methylated fragments. Although all of them claim to be of high quality, differences in performance exist. Careful kit selection is thus of utmost importance [24]. Here, sheared tumor DNA was enriched towards methylated fragments using the MethylCap kit of Diagenode, that makes use of the methylCap protein, consisting of the MBD of human MeCP2 fused

with glutathione-S-transferase (GST) containing an N-terminal His6-tag. A previous evaluation assessed the quality of this kit for combination with NGS by comparison with four other commercially available kits [24]. This study also compared the MBD sequencing data with reduced representation bisulfite sequencing (RRBS) and Illumina 27 K methylation array data

of the same samples. Together, these analyses showed that the MethylCap kit outperforms the others, due to a consistent combination of high yield, sensitivity and specificity [24]. In order to demonstrate that the samples of MBD cohort I and II were enriched for methylated DNA fragments after MBD-based capturing, we made use of the fragment CpG plot [24]. As this plot depicts the CpG content of the mapped fragments and the MethylCap kit theoretically only captures methylated cytosines in a CpG dinucleotide context, the fragment CpG plot can be used to evaluate the MBD-based enrichment. An overview of the CpG content of the mapped fragments per sample cohort is depicted in Figure 4. This fragment CpG plot clearly

illustrates that the MBD-enriched samples of MBD cohort I and II have a high fraction of CpG-dense fragments, while the input (non-MBD-enriched) samples of MBD cohort II are not enriched in CpG content. Additionally, using the number of mapped reads per kilobase CpG island per million (RPKM) values [25], the methylation level of each CpG island across the different subcohorts was determined. The density plot in Figure 5 indicates that the MBD-enriched samples have a higher fraction of CpG islands with an RPKM > 1 compared to the input samples of MBD cohort II. Based on these analyses, it can be concluded that the MBD-based capture successfully led to the enrichment of methylated DNA fragments.

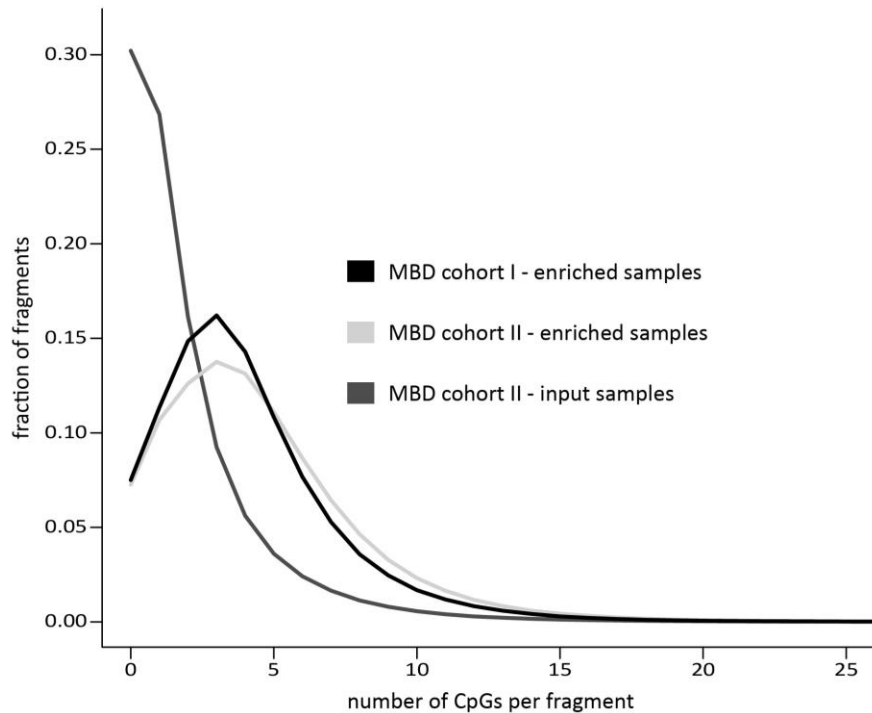


Figure 4. Fragment CpG plots demonstrate that the MBD-enriched samples have a high fraction of CpG-dense sequencing fragments. Shown are the fractions of mapped MBD sequencing fragments with different CpG counts. Per cohort, 100,000 randomly selected fragments of each sample were used to construct the plots.

Validation of methylated genes in neuroblastoma

Finally, TDF and BED files, containing sequence coverage and peak locations respectively, were loaded into IGV to visually inspect genes previously described to be methylated in NB. As

an example, the MBD sequencing data of the *PCDHB* gene cluster is shown in Figure 6. This gene cluster is frequently methylated in NB [5,31], which is confirmed by the MBD sequencing data of both MBD cohort I and II. Additionally, 78 regions identified in the MBD sequencing data as being methylated, were

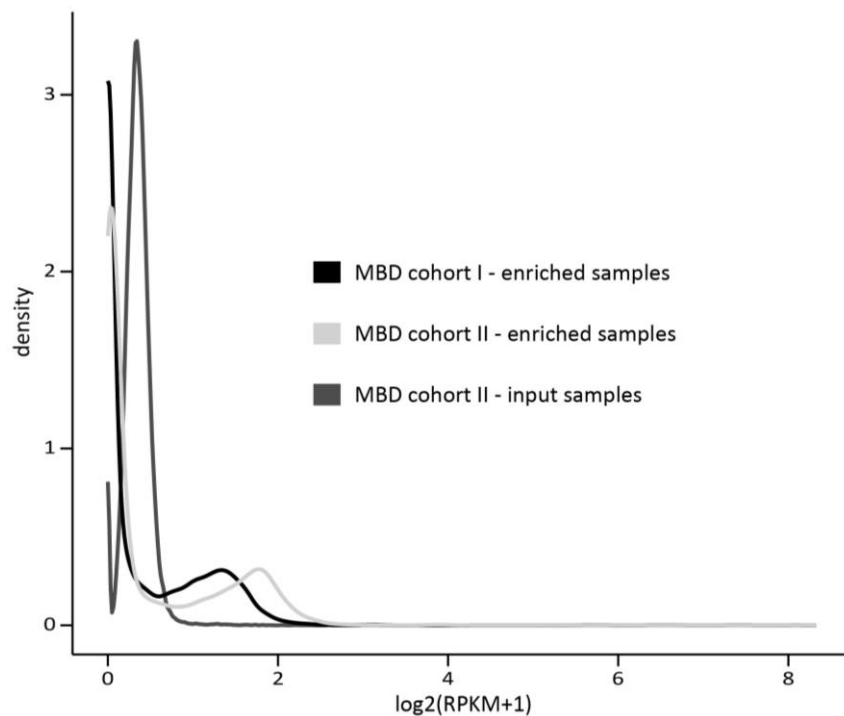


Figure 5. CpG island RPKM-values confirm enrichment towards methylated DNA fragments upon MBD capture. Shown are the densities of the median RPKM-values per subcohort. RPKM: reads per kilobase CpG island per million.

validated in two independent patient cohorts using methylation-specific PCR (MSP) [14]. These data confirm the validity of MBD sequencing in identifying methylated regions in NB.

Usage notes

The MBD sequencing data can be downloaded from the GEO database via accession numbers GSE69224 (for MBD cohort I; Data Citation 1), GSE69243 and GSE69268 (for MBD cohort II; Data Citation 2 and Data Citation 3; SuperSeries GSE69279 (Data Citation 4)). The unique GEO sample accession IDs and clinical annotation can be found in Table 1 (available online only). This table also contains the accession IDs of the matching expression and aCGH data, which allows easy data access and facilitates integration analyses.

All output files from the different steps in the MBD sequencing data processing are provided through GEO. Analysis tools and scripts have been embedded in a Docker container, to deliver an environment that runs on any supported host

platform (Windows, MAC, Linux). This Docker container, and all instructions on how it is made and how analyses can be run on the data, are made available through Docker Hub and GitHub (see Code availability). This allows researchers to try out the analysis pipeline that was used to generate the publically available data, without the need of additional infrastructure or software versions. The Docker container guarantees that the provided commands work and allows researchers to start exploring the data at the level they are experienced with.

Alternative processing tools can be tested for read mapping (e.g., BWA [32]) or identification of enriched regions (e.g., PeakRanger [33] or BALM [34]), or absolute methylation scores can be calculated (MEDIPS [35]; see Code availability). Researchers inexperienced with MBD sequencing can easily visualize their genes of interest by downloading the BED and TDF files (see Code availability). Downstream differential methylation analyses can be done with DESeq [28] (as described in Decock et al. [14]) using count data sets provided in Supplementary

results

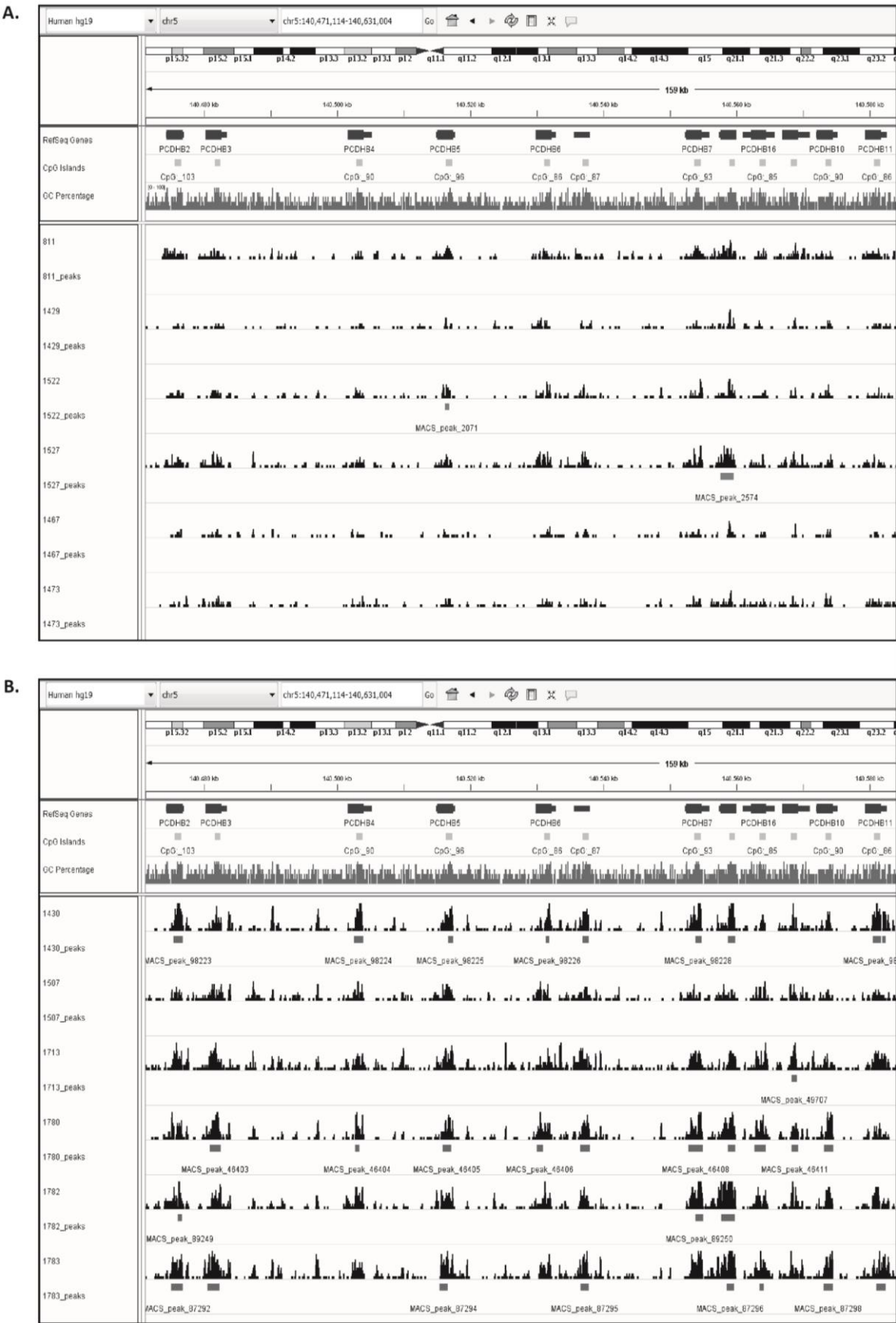


Figure 6. Visualization of the MBD sequencing data in IGV confirms methylation of the *PCDH* gene cluster. In **A.** the data of MBD cohort I is shown, in **B.** the data of MBD cohort II. The upper panels show the genes in the cluster, the location of CpG islands and the GC percentage. In the lower panels, sequence coverage of 6 high-risk patient samples is shown (peak pattern), as well as the location of identified peaks (horizontal bars).

Tables 3-8, or other software can be used, such as DiffBind [36] and edgeR [37]. Differences in absolute methylation scores can be used for RankProd [38] analyses.

Acknowledgements

The authors thank Rosa Noguera, Johannes H. Schulte and Raymond L. Stallings for sharing primary tumor samples, and NXTGNT for the sequencing services. The work is further supported by the Emmanuel Van der Schueren Foundation (scientific partner of the Flemish League Against Cancer (VLK)), the Fournier-Majoie Foundation (FFM) and the Belgian National Lottery. A.D. was a recipient of a grant of the Research Foundation Flanders (FWO) and an Emmanuel Van der Schueren research grant (VLK).

Author contributions

Drafting the article: A.D. Data generation: W.V. Data processing: M.O. Technical validation analyses: A.D. and M.O. Overall supervision of study: F.S. and J.V. All authors contributed to preparation of the manuscript and approved the final version.

Additional information

Table 1 is only available in the online version of this paper.

Supplementary Information accompanies this paper at:

<http://www.nature.com/articles/sdata20164>

Supplemental Table 1. Using BamUtil, basic sequencing statistics of each sample of MBD cohort I and II are computed. Given are the total read numbers, and the number and percentage of properly paired and duplicate reads of each sample of MBD cohort I (a) and II (enriched samples in (b); input samples in (c)).

Supplemental Table 2. Using SAMStat, the mapping quality scores of each sample of MBD cohort I and II are analyzed. Given are the

numbers and percentages of mapped reads across the different mapping quality ranges, as determined by SAMStat ((a) enriched samples of MBD cohort I, (b) enriched samples of MBD cohort II and (c) input samples of MBD cohort II).

Supplemental Table 3. Promoter count data for the MBD-enriched samples of MBD cohort I. For each MBD-enriched sample of MBD cohort I, the number of mapped reads in each Ensembl Transcript promoter region (-1500 bp to +500 bp around TSS) is given.

Supplemental Table 4. Promoter count data for the MBD-enriched samples of MBD cohort II. For each MBD-enriched sample of MBD cohort II, the number of mapped reads in each Ensembl Transcript promoter region (-1500 bp to +500 bp around TSS) is given.

Supplemental Table 5. Promoter count data for the input samples of MBD cohort II. For each input sample of MBD cohort II, the number of mapped reads in each Ensembl Transcript promoter region (-1500 bp to +500 bp around TSS) is given.

Supplemental Table 6. Window count data for the MBD-enriched samples of MBD cohort I. For each MBD-enriched sample of MBD cohort I, the number of mapped reads in each 5 kb genomic window (2.5 kb overlapping moving windows) is given.

Supplemental Table 7. Window count data for the MBD-enriched samples of MBD cohort II. For each MBD-enriched sample of MBD cohort II, the number of mapped reads in each 5 kb genomic window (2.5 kb overlapping moving windows) is given.

Supplemental Table 8. Window count data for the input samples of MBD cohort II. For each input sample of MBD cohort II, the number of mapped reads in each 5 kb genomic window (2.5

kb overlapping moving windows) is given.

Competing financial interests

The authors declare no competing financial interests.

How to cite this article

Decock, A. et al. DNA methylation profiling of primary neuroblastoma tumors using methyl-CpG-binding domain sequencing. *Sci. Data* 3:160004 doi: 10.1038/sdata.2016.4 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit

<http://creativecommons.org/licenses/by/4.0>

Metadata associated with this Data Descriptor is available at <http://www.nature.com/sdata/> and is released under the CC0 waiver to maximize reuse.

References

- [1] Brodeur GM (2003). Neuroblastoma: biological insights into a clinical enigma. *Nature Reviews. Cancer*; 3(3):203-216.
- [2] Molenaar JJ et al. (2012). Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* 483(7391):589-593.
- [3] Valentijn LJ et al. (2015). TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nature Genetics*; 47(12):1411-1414.
- [4] Peifer M et al. (2015). Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature*; 526(7575):700-704.
- [5] Decock A et al. (2011). Neuroblastoma epigenetics: from candidate gene approaches to genome-wide screenings. *Epigenetics*; 6(8):962-970.
- [6] Carén H et al. (2011). Identification of epigenetically regulated genes that predict patient outcome in neuroblastoma. *BMC Cancer*; 11:66.
- [7] Decock A et al. (2012). Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. *Genome Biology*; 13(10):R95.
- [8] Duijkers F et al. (2013). Epigenetic drug combination induces genome-wide demethylation and altered gene expression in neuro-ectodermal tumor-derived cell lines. *Cellular Oncology*; 36(5):351-362.
- [9] Mayol G et al. (2012). DNA hypomethylation affects cancer-related biological functions and genes relevant in neuroblastoma pathogenesis. *PLoS One*; 7(11): e48401.
- [10] Yáñez Y et al. (2015). Two independent epigenetic biomarkers predict survival in neuroblastoma. *Clinical Epigenetics*; 7(1):16.
- [11] Gómez S et al. (2015). DNA methylation fingerprint of neuroblastoma reveals new biological and clinical insights. *Epigenomics*; 7(7):1137-1153.
- [12] Schramm A et al. (2015). Mutational dynamics between primary and relapse neuroblastomas. *Nature Genetics*; 47(8): 872-877.
- [13] Stirzaker C et al. (2014). Mining cancer methylomes: prospects and challenges. *Trends in Genetics*; 30(2):75-84.
- [14] Decock A et al. (2015). Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncotarget*; 7(2):1960-1972.
- [15] Brodeur GM et al. (1993). Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *Journal of Clinical Oncology*; 11(8):1466-1477.
- [16] Mestdagh P et al. (2010). The miR-17-92 microRNA cluster regulates multiple components of the TGF- β pathway in neuroblastoma. *Molecular Cell*; 40(5):762-773.
- [17] Eschenburg G et al. (2012). Smac mimetic LBW242 sensitizes XIAP-overexpressing neuroblastoma cells for TNF- α -independent apoptosis. *Cancer Research*; 72(10):2645-2656.
- [18] Kumps C et al. (2013). Focal DNA copy number changes in neuroblastoma target MYCN regulated genes. *PLoS One*; 8(1):e52321.
- [19] Sante T et al. (2014). ViVar: a comprehensive platform for the analysis and visualization of structural genomic variation. *PLoS One*; 9(12):e113800.
- [20] Langmead B et al. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*; 9(4):357-359.
- [21] Li H et al. (2009). The Sequence

- Alignment/Map format and SAMtools. *Bioinformatics*; 25(16):2078-2079.
- [22] Quinlan AR et al. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*; 26(6):841-842.
- [23] Lassmann T et al. (2011). SAMStat: monitoring biases in next generation sequencing data. *Bioinformatics*; 27(1):130-131.
- [24] De Meyer T et al. (2013). Quality evaluation of methyl binding domain based kits for enrichment DNA-methylation sequencing. *PLoS One*; 8(3):e59068.
- [25] Mortazavi A et al. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*; 5(7):621-628.
- [26] Zhang Y et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*; 9(9):R137.
- [27] Thorvaldsdóttir H et al. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*; 14(2):178-192.
- [28] Anders S et al. (2010). Differential expression analysis for sequence count data. *Genome Biology*; 11(10):R106.
- [29] Minoche AE et al. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biology*; 12(11):R112.
- [30] Yu Y et al. (2014). Comprehensive RNA-Seq transcriptomic profiling across 11 organs, 4 ages, and 2 sexes of Fischer 344 rats. *Scientific Data*; 1:140013.
- [31] Banelli B et al. (2012). A pyrosequencing assay for the quantitative methylation analysis of the PCDHB gene cluster, the major factor in neuroblastoma methylator phenotype. *Laboratory Investigation*; 92(3):458-465.
- [32] Li H et al. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*; 25(14):1754-1760.
- [33] Feng X et al. (2011). PeakRanger: a cloud-enabled peak caller for ChIP-seq data. *BMC Bioinformatics*; 12:139.
- [34] Lan X et al. (2011). High resolution detection and analysis of CpG dinucleotides methylation using MBD-seq technology. *PLoS One*; 6(7):e22226.
- [35] Lienhard M et al. (2014). MEDIPS: Genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics*; 30(2):284-286.
- [36] Stark R et al. (2011). DiffBind: differential binding analysis of ChIP-seq peak data. <http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>.
- [37] Robinson M et al. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*; 26(1):139-140.
- [38] Hong F et al. (2006). RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics*; 22(22):2825-2827.

Data Citations

1. Decock A et al. Gene Expression Omnibus GSE69224 (2015).
2. Decock A et al. Gene Expression Omnibus GSE69243 (2015).
3. Decock A et al. Gene Expression Omnibus GSE69268 (2015).
4. Decock A et al. Gene Expression Omnibus GSE69279 (2015).

paper 4

Stage 4S neuroblastoma tumors show a characteristic DNA methylation portrait

Anneleen Decock, Maté Ongenaert, Bram De Wilde, Bénédicte Brichard, Rosa Noguera, Frank Speleman and Jo Vandesompele

accepted for publication in Epigenetics
impact factor 2015: 4.774
times cited (d.d. 11/09/2016; Google Scholar): 0

PAPER 4: STAGE 4S NEUROBLASTOMA TUMORS SHOW A CHARACTERISTIC DNA METHYLATION PORTRAIT

Anneleen Decock^{1,2}, Maté Ongenaert¹, Bram De Wilde^{1,2,3}, Bénédicte Brichard⁴, Rosa Noguera⁵, Frank Speleman^{1,2}, Jo Vandesompele^{1,2,6}

Correspondence to joke.vandesompele@ugent.be.

¹Center for Medical Genetics, Ghent University, De Pintelaan 185, Ghent, Belgium. ²Cancer Research Institute Ghent (CRIG), De Pintelaan 185, Ghent, Belgium. ³Department of Pediatric Hematology and Oncology, Ghent University Hospital, De Pintelaan 185, Ghent, Belgium. ⁴Cliniques Universitaires Saint-Luc, Université Catholique de Louvain, 10 avenue Hippocrate, Brussels, Belgium. ⁵Department of Pathology, Medical School, University of Valencia, and Health Research Institute INCLIVA, Blasco Ibañez 17, Valencia, Spain. ⁶Bioinformatics Institute Ghent - From Nucleotides to Networks (BIG N2N), De Pintelaan 185, Ghent, Belgium.

Contribution of AD: AD drafted the article and performed all bioinformatics and statistical analyses (differential methylation and expression analyses, cytogenetic band enrichment analyses, gene set enrichment analyses and transcription factor target enrichment analyses).

Abstract

Stage 4S neuroblastoma (NB) is a special type of NB found in infants with metastases at diagnosis and is associated with an excellent outcome due to its remarkable capacity to undergo spontaneous regression. As genomics have not been able to explain this intriguing clinical presentation, we here aimed at profiling the DNA methylome of stage 4S NB to better understand this phenomenon. To this purpose, differential methylation analyses between INSS stage 4S, stage 4 and stage 1/2 were performed, using methyl-CpG-binding domain (MBD) sequencing data of 14 stage 4S, 14 stage 4, and 13 stage 1/2 primary NB tumors (all *MYCN* non-amplified in order not to confound results). Stage 4S-specific hyper- and hypomethylated promoters were determined and further characterized for genomic localization and function by cytogenetic band enrichment, gene set enrichment, transcription factor target enrichment and differential RNA expression analyses. We show that specific chromosomal locations are enriched for stage 4S differentially methylated promoters and that stage 4S tumors show characteristic hypermethylation of specific subtelomeric promoters. Furthermore, genes involved in important oncogenic pathways, in neural crest development and differentiation, and in epigenetic processes are differentially methylated and expressed in stage 4S tumors. Based on these findings, we describe new biological mechanisms possibly contributing to the stage 4S-specific tumor biology and spontaneous regression. In conclusion, this study is the first to describe the highly characteristic stage 4S DNA methylome. These findings will open new avenues to further unravel the NB pathology in general and stage 4S disease specifically.

Keywords: neuroblastoma, stage 4S (MS), DNA methylation, methyl-CpG-binding domain (MBD) sequencing, spontaneous regression

Introduction

Neuroblastoma (NB) is a childhood tumor that originates from precursor cells of the sympathetic nervous system and is a heterogeneous disease with prognosis ranging from long-term survival to fatal outcome. NB tumor staging systems have been established, based on clinical and biological parameters (age of the patient, tumor stage, *MYCN* amplification, DNA index and histopathology), and are being used for therapeutic stratification. The special stage IV (IV-S) was first described in the staging system of Evans et al., introduced in 1971, and defined patients who would otherwise have been stage I or II (with localized primary tumors), but who had disseminated disease confined to liver, skin or bone marrow, or any combination of these, without involvement of the skeleton. Remarkably, these patients did not have the dismal prognosis of other metastatic NB, as most of these tumors are characterized by spontaneous regression without any therapeutic intervention [1,2]. Approximately twenty years later, the International Neuroblastoma Staging System (INSS) refined the criteria of this tumor stage, renamed as stage 4S, and limited it to children younger than 12 months of age at diagnosis [3]. In 2005, this age limit was extended to 18 months in the International Neuroblastoma Risk Group Staging System (INRGSS), as it was demonstrated that an age cutoff of 18 months resulted in a stratification that optimized the prognostic contribution of the age at diagnosis [4]. In the INRGSS, this special tumor stage is indicated as MS [5].

Several genomic alterations have been identified in NB, but so far no specific changes, apart from hyperdiploidy and near-triploidy with numerical chromosome alterations common to localized tumors without *MYCN* amplification, have been associated with stage 4S NB. This indicates that other molecular mechanisms, such as epigenetic alterations, may play a role in stage 4S NB [6]. In this study, we therefore specifically focused on

identifying patterns of DNA methylation that are characteristic of this particular NB tumor stage.

Results

Stage 4S tumors demonstrate a unique promoter methylation portrait

In order to profile the stage 4S NB methylome, we re-analyzed methyl-CpG-binding domain (MBD) sequencing data of 41 *MYCN* non-amplified primary NBs [7,8], including INSS stage 1 (n = 7), stage 2 (n = 6), stage 4 (n = 14) and stage 4S (n = 14) tumors. As patients with stage 4S NB, in sharp contrast to stage 4 patients, show excellent outcome despite widespread metastases at diagnosis, we first performed a differential methylation analysis between these tumor stages. In total, 5,914 and 10,974 promoters were identified as being hyper- or hypomethylated in stage 4S tumors, respectively (Supplemental table 1). Next, stage 4S tumors were also compared to stage 1/2 tumors, which are also associated with a favorable patient outcome, but patients with stage 1/2 tumors do not present with metastatic disease. Here, 1,288 and 3,854 promoters were hyper- and hypomethylated in stage 4S tumors, respectively (Supplemental table 1). Subsequently, we determined the number of stage 4S-specific hyper- and hypomethylated promoters by determining the overlap between these two differential methylation analyses and by excluding promoters differentially methylated in the stage 1/2 versus stage 4 comparison. As such, we identified 393 hypermethylated and 1,150 hypomethylated promoters that typify the stage 4S methylome (Figure 1A; Supplemental table 1). Similarly, also for stage 4 and stage 1/2 NB, specific promoter methylation patterns were determined (Figure 1A; Supplemental table 1). Principal component analysis using the normalized read counts of the promoters included in Figure 1A visualizes the NB stage-specific promoter methylation patterns (Figure

1B). A dense clustering is observed for stage 4S tumors and this tumor stage seems most similar to stage 1/2.

Specific chromosomal locations are enriched in stage 4S differentially methylated promoters

Next, we determined the chromosomal distribution of the promoters contributing to the specific stage 4S methylation portrait. The percentages of hyper- and hypomethylated promoters located on the different chromosomes is indicated in Figure 2A and Figure 2B, respectively. Importantly, as chromosomes differ in length and gene density, these chromosome percentages were calculated based on the number of identified promoters corrected for the total number of promoters located on the corresponding chromosome. Stage 4S hypermethylated promoters are

overrepresented ($p < 0.05$) on chr17 and chr19, and underrepresented ($p < 0.05$) on chr1, chr3, chr4, chr5, chr7, chr10, chr11, chr14, chr15, chr18, chr20 and chr21 (Figure 2A). Stage 4S hypomethylated promoters are overrepresented ($p < 0.05$) on chr3, chr7, chr8, chr13 and chr18, and underrepresented ($p < 0.05$) on chr2, chr6, chr12, chr16, chr17, chr19 and chr22 (Figure 2B). To further pinpoint specific chromosomal regions enriched in stage 4S differentially methylated promoters, cytogenetic band enrichment analyses in WebGestalt [9,10] were performed (Figure 2C and Supplemental table 2). These analyses indicate that stage 4S hyper- and hypomethylated promoters are not randomly distributed across the genome, but that specific chromosomal cytogenetic bands are significantly enriched in stage 4S differentially methylated promoters (Figure 2).

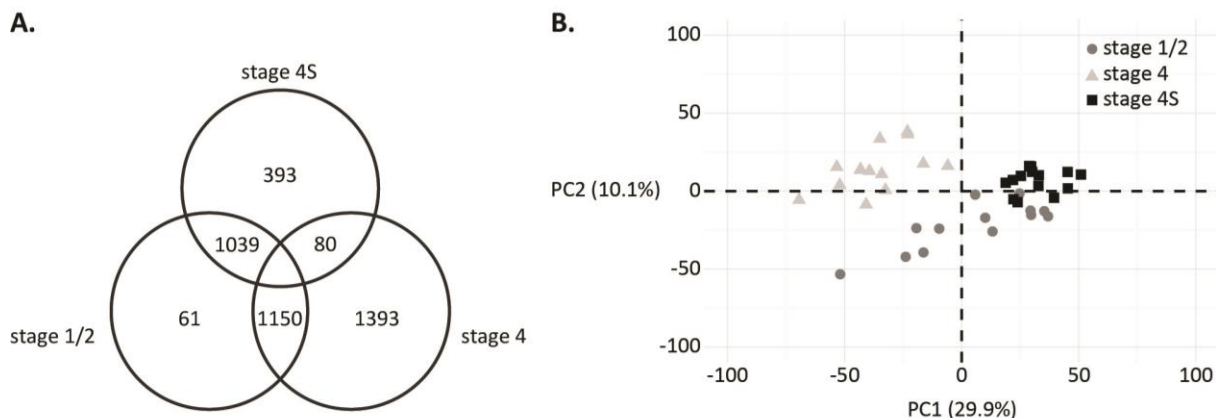


Figure 1. MBD sequencing analysis of *MYCN* non-amplified stage 4S, stage 4 and stage 1/2 tumors identifies stage-specific promoter methylation portraits. A. Diagram of the promoters identified as hypermethylated in the respective sample groups. For example, 393 and 1,150 promoters are specifically hyper- and hypomethylated in stage 4S, respectively. **B.** Principal component (PC) analysis using the normalized read counts of the promoters included in the diagram. Shown are the data for the first two PCs. PC1 explains 29.9% of the variability of the data and PC2 10.1%.

Specific subtelomeres are enriched in stage 4S hypermethylated promoters

Telomerase activity upregulation and alternative lengthening of telomeres are key mechanisms in counteracting telomere shortening in NB and contribute to tumor cell immortalization [11–14]. As DNA methylation of subtelomeres

represents an additional way to suppress telomere elongation [15,16] and as tumor regression (and thus loss of immortality) is a hallmark of stage 4S NB, we specifically analyzed the subtelomeric methylation patterns of the different NB tumor stages. To this purpose, we first determined the proportions of promoters of the different stage-specific methylation portraits

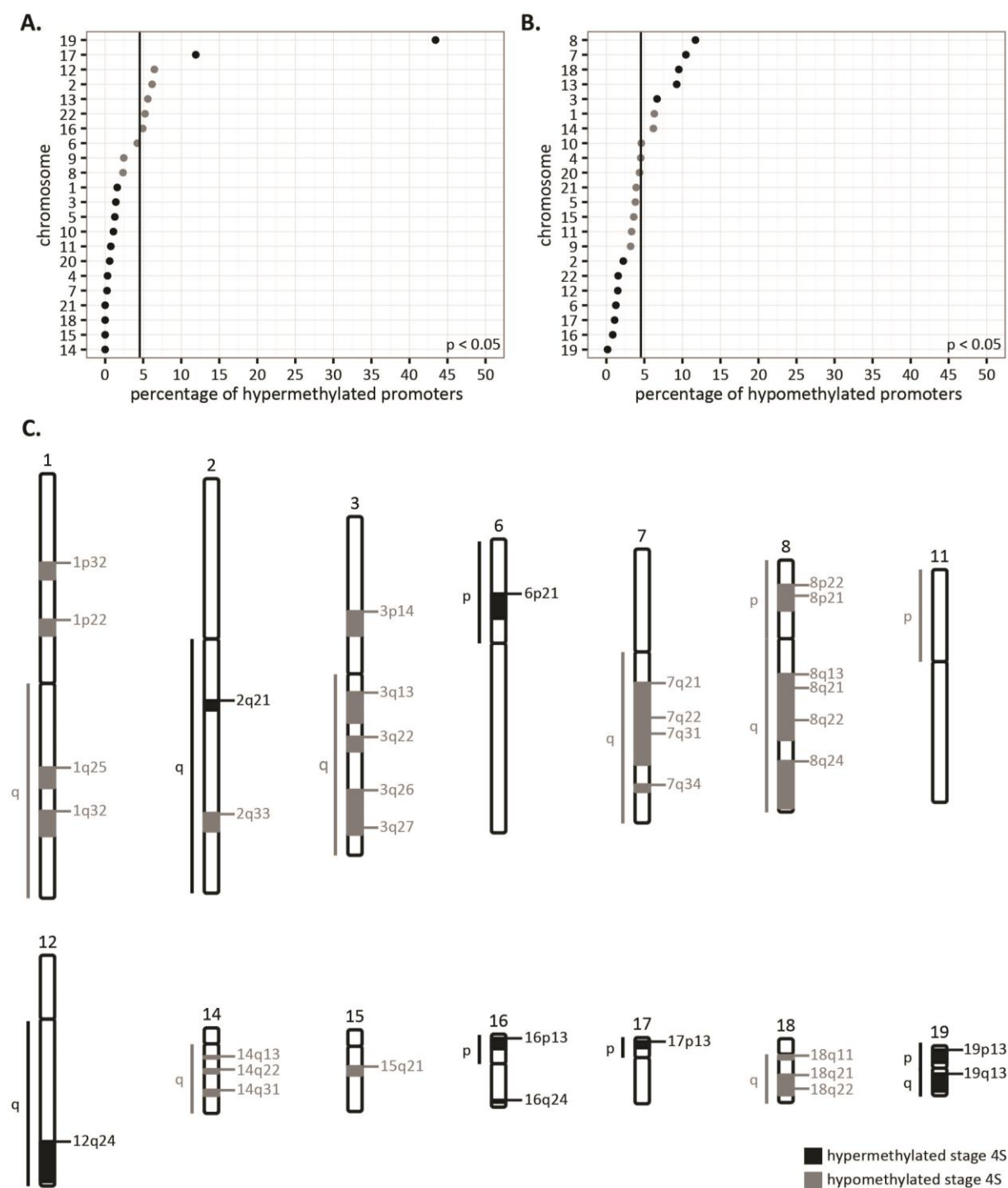


Figure 2. Stage 4S differentially methylated promoters are not randomly distributed across the genome. A. Percentage of stage 4S hypermethylated promoters per chromosome, based on promoter density corrected numbers. Significantly under- and overrepresented chromosomes are indicated with a dark dot. The vertical line represents the percentage if the stage 4S hypermethylated promoters would be randomly distributed across the genome (null hypothesis). **B.** Percentage of stage 4S hypomethylated promoters per chromosome, based on promoter density corrected numbers. Significantly under- and overrepresented chromosomes are indicated with a dark dot. The vertical line represents the percentage if the stage 4S hypomethylated promoters would be randomly distributed across the genome (null hypothesis). **C.** Significantly enriched cytogenetic locations of stage 4S hyper- and hypomethylated promoters. A list of genes located in the enriched regions can be found in Supplemental table 2.

that are located in subtelomeres (1 Mb or 2 Mb regions proximal to the telomeres). Figure 3A and B illustrate that stage 4S hypermethylated and stage 4 hypomethylated promoters have a significantly larger proportion of 1 Mb and 2 Mb subtelomeric promoters ($p < 0.001$), compared to stage 1/2 and stage 4 hypermethylated and stage 4S hypomethylated promoters. Stage 1/2 hypomethylated promoters also comprise a larger proportion of 2 Mb subtelomeric promoters ($p < 0.001$; Figure 3B), compared to stage 1/2 and stage 4 hypermethylated and stage 4S hypomethylated promoters. Next, using the gene lists of the differential methylation analyses, gene set enrichment analyses (GSEA) were performed to evaluate whether specific subtelomeres are enriched in differentially methylated promoters (Table 1). These analyses

firmly demonstrate that multiple subtelomeres are enriched in stage 4S hypermethylated promoters compared to stage 4 and stage 1/2 tumors. Remarkably, not a single subtelomere is enriched in stage 4 hypermethylated promoters. Table 1 further shows that the chr5p 1 Mb and 2 Mb subtelomeres are enriched in promoters hypermethylated in stage 4S compared to both stage 1/2 and stage 4, and in stage 1/2 compared to stage 4. *TERT*, the gene encoding the catalytic subunit of telomerase, is located on the chr5p 2 Mb subtelomere, and its promoter is hypermethylated in stage 4S compared to stage 4 (Supplemental table 1). Using publically available mRNA expression data [17], we could further demonstrate that *TERT* is expressed at lower levels in stage 4S compared to stage 4 ($p < 0.001$; log2 fold change = 0.82).

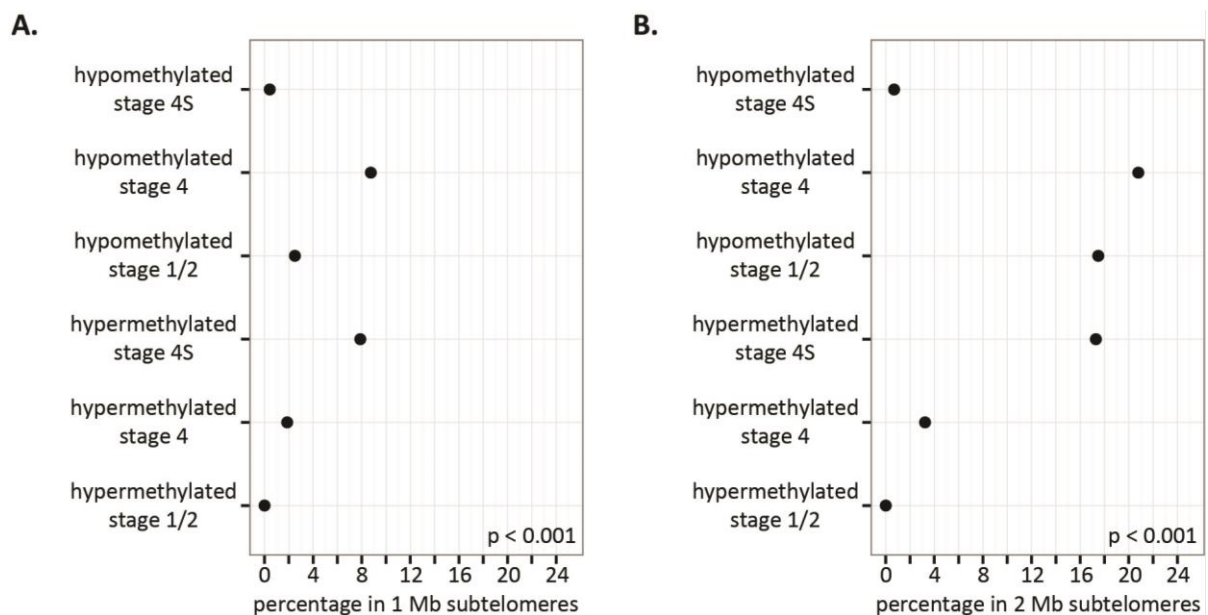


Figure 3. Stage 4S hypermethylated, and stage 4 and 1/2 hypomethylated promoters are frequently located on subtelomeres. Depicted are the percentages of the stage-specific differentially methylated promoters located in the 1 Mb (A.) and 2 Mb (B.) subtelomeres. Stage 4S hypermethylated and stage 4 hypomethylated promoters have a significantly larger proportion of 1 Mb and 2 Mb subtelomeric promoters, compared to stage 1/2 and stage 4 hypermethylated and stage 4S hypomethylated promoters. Stage 1/2 hypomethylated promoters also comprise a larger proportion of 2 Mb subtelomeric promoters, compared to stage 1/2 and stage 4 hypermethylated and stage 4S hypomethylated promoters.

Genes involved in important oncogenic pathways, neural crest development and differentiation, and epigenetic processes are differentially methylated and expressed in stage 4S tumors

To functionally characterize the stage 4S differentially methylated promoters, we first examined whether these promoters share common DNA motifs, by analyzing transcription factor target enrichment in WebGestalt [9,10]. In

Table 1. One third of subtelomeres are enriched in stage 4S hypermethylated promoters.

comparison		enriched 1Mb subtelomeres				enriched 2 Mb subtelomeres			
		ST	NES	nom p-val	FDR q-val	ST	NES	nom p-val	FDR q-val
stage 4S versus stage 4	hypermethylated stage 4S	chr4p	1.58	0.037	0.040	chr1p	2.21	< 0.001	< 0.001
		chr5p	2.29	< 0.001	< 0.001	chr4p	1.92	< 0.001	0.002
		chr10q	1.87	< 0.001	0.002	chr4q	1.47	0.041	0.045
		chr13q	2.12	< 0.001	< 0.001	chr5p	2.60	< 0.001	< 0.001
		chr16p	2.32	< 0.001	< 0.001	chr5q	1.49	0.011	0.042
		chr17q	1.68	0.006	0.022	chr8q	1.35	0.022	0.083
		chr19p	2.28	< 0.001	< 0.001	chr9q	1.58	< 0.001	0.029
						chr10q	1.78	< 0.001	0.003
						chr13q	1.97	< 0.001	0.001
						chr16p	2.50	< 0.001	< 0.001
						chr16q	2.17	< 0.001	< 0.001
						chr17q	1.65	< 0.001	0.020
						chr19p	2.80	< 0.001	< 0.001
						chr20q	1.83	< 0.001	0.004
	hypermethylated stage 4	none				none			
stage 4S versus stage 1/2	hypermethylated stage 4S	chr4q	2.04	< 0.001	< 0.001	chr4q	2.12	< 0.001	< 0.001
		chr5p	1.59	0.022	0.030	chr5p	1.81	< 0.001	0.006
		chr9q	2.30	< 0.001	< 0.001	chr5q	1.96	< 0.001	< 0.001
		chr12q	1.60	0.027	0.031	chr8q	1.79	< 0.001	0.006
		chr13q	1.72	< 0.001	0.013	chr10q	1.73	< 0.001	0.008
		chr17p	1.81	0.008	0.008	chr13q	1.75	< 0.001	0.009
		chr17q	2.22	< 0.001	< 0.001	chr20q	3.56	< 0.001	< 0.001
		chr20q	2.19	< 0.001	< 0.001	chr21q	1.64	< 0.001	0.018
		chr22q	1.53	< 0.001	0.044	chr22q	1.28	0.047	0.109
	hypermethylated stage 1/2	chr1p	1.38	0.035	0.141	none			
stage 1/2 versus stage 4	hypermethylated stage 1/2	chr5p	2.29	< 0.001	< 0.001	chr4p	1.49	0.023	0.054
						chr5p	1.92	< 0.001	< 0.001
						chr10p	1.49	0.035	0.059
	hypermethylated stage 4	none				none			

Note. For each of the three sample group comparisons, the subtelomeres (ST) significantly enriched for promoters hypermethylated in each of the sample groups are indicated, as well as their corresponding normalized enrichment score (NES), nominal p-value (nom p-val < 0.05) and false discovery rate q-value (FDR q-val < 0.25).

total, 24 and 320 motifs were significantly enriched for stage 4S hyper- and hypomethylated promoters, respectively (Supplemental table 3), of which the top 10 (ranked based on the ratio of enrichment) are shown in Table 2. Motifs enriched in stage 4S differentially methylated promoters represent binding sites for transcription factors implicated

in controlling cell cycle, apoptosis, neural crest development and neural differentiation (Table 2). Further, we performed GSEA on the gene lists of the differential methylation analyses, using the oncogenic signatures database containing gene sets that represent signatures of cellular pathways dysregulated in cancer [18]. Enriched gene sets were only identified for promoters

Table 2. Specific transcription factor targets are enriched in stage 4S differentially methylated promoters.

transcription factor target	TF	hypermethylated in stage 4S enrichment analysis						differential expression analysis	
		C	E	O	genes	R	q	stage 4S versus stage 4 q	stage 4S versus stage 1/2 q
CCGNMNNTNACG		75	0.28	3	<i>HMG20B, HSPD1, FOSB</i>	10.58	0.035	0.198	0.777
NTNGCGTGNNN	AHR	209	0.79	6	<i>TSSK6, SLC9A5, FHOD1, FOSB, CIRBP, LPHN1</i>	7.60	0.012	0.051	0.292
TTCNRGNNNTTC		150	0.57	4	<i>MKNK2, HSPD1, FOSB, CCT4</i>	7.06	0.034	0.225	0.944
TAAWWATAG	MEF2A	163	0.62	4	<i>EIF5A, SLC9A5, ALDOA, JSRP1</i>	6.49	0.036	< 0.001	< 0.001
SGCGSSAAA	E2F1 TFDP1 RB1	167	0.63	4	<i>MCM2, MCM7, SLC9A5, FHOD1</i>	6.34	0.038	< 0.001	0.001
TGAMCTTTGMMCYT	HNF4A	260	0.98	6	<i>PRKRA, NDUFA3, PRKCSH, RTN2, C4A, ALDOA</i>	6.11	0.017	0.033	0.147
TTTSGCGS	E2F1	228	0.86	5	<i>GPN3, MCM2, MCM7, SLC9A5, FHOD1</i>	5.80	0.034	< 0.001	0.001
NCSCGCSAAAN	E2F1 TFDP1	234	0.88	5	<i>EIF5A, TNPO2, GPN3, MCM2, MCM7</i>	5.65	0.034	0.053	0.044
GCCATNTTN	YY1	238	0.90	5	<i>NCOR1, PIGL, PRKCSH, PTBP1, LPHN1</i>	5.56	0.034	0.121	0.003
NKCGCGCSAAAN	E2F1 TFDP1	239	0.90	5	<i>EIF5A, TNPO2, MCM2, MCM7, PRKCSH</i>	5.54	0.034	0.053	0.044

(continues)

Note. Shown are the top 10 significantly ($q < 0.05$) enriched transcription factor targets (motifs) in stage 4S hyper- and hypomethylated promoters. For each enriched motif, matching transcription factor (TF) annotation is shown (if available), as well as the number of genes in the category (C), the expected number in the category (E), the number of genes hyper- or hypomethylated in stage 4S in the category (O), the annotation for the genes in O, the ratio of enrichment (R) and the q-value. Global test q-values of the differential expression analyses are also indicated.

Table 2. Specific transcription factor targets are enriched in stage 4S differentially methylated promoters.

(continued)

transcription factor target	TF	hypomethylated in stage 4S enrichment analysis						differential expression analysis	
		C	E	O	genes	R	q	stage 4S versus stage 4 q	stage 4S versus stage 1/2 q
NNTGTTACTAAAA ATAGAAMNN		25	0.32	3	<i>SLCO5A1, PPARGC1A, MGST3</i>	9.27	0.011	0.045	0.213
NTGCGTGGGCGK	EGR3	86	1.11	7	<i>CELF4, GPR176, ATP6V1C1, EPHB1, KCNQ5, UBE2W, CLTC</i>	6.29	< 0.001	< 0.001	< 0.001
CARAAGTAGGNCA AAGGTCA	PPARA	37	0.48	3	<i>SS18, MARCH10, PPARGC1A</i>	6.27	0.026	0.046	0.379
KMCATNNWGA		87	1.13	7	<i>CELF4, PDE4D, ADAMTSL1, RUNX1T1, OSBPL6, ZFAT, ARNTL</i>	6.22	< 0.001	0.530	< 0.001
TGATTRY	GFI1	291	3.77	23	<i>RUNX1T1, FOXN3, WDPCP, AGR2, ALPK2, MYOT, SND1, CREB5, CDK14, IGF2BP3, FOXP2, NRXN3, MIR137HG, PTEN, C12orf42, ALDH1A2, JAZF1, ZNF521, MYBPC1, ZBTB20, PPARGC1A, MAPK10, RIMS2</i>	6.11	< 0.001	< 0.001	< 0.001
NNGAATATKCAN NNN	POU2F1	211	2.73	16	<i>ADAMTSL1, NPAS3, RUNX1T1, PHLPP1, NRXN3, FOXP2, MARCH10, PPFA2, ZNF521, ZBTB20, LPL, PDE4D, CACNA1D, PPARGC1A, RGS8, ARPP21</i>	5.86	< 0.001	< 0.001	< 0.001
CTBATTTCARAAW	CEBPG	252	3.26	19	<i>BRAF, RUNX1T1, NR6A1, FOXN3, PTK2, CREB5, WDR64, KCNQ5, CXCL9, FOXP2, SPAG9, PPFA2, ALDH1A2, ST18, ZBTB20, C8orf46, SLIRP, PPARGC1A, PABPC1</i>	5.83	< 0.001	< 0.001	0.007
NNNNNNKCTAW AAATAGMNNNN		226	2.92	17	<i>LARS, SLC8A3, CELF4, ALPK2, CKMT2, KCNQ5, CACNA2D3, ARHGEF38, ARNTL, FBXO40, TSC22D1, TNNT2, PPARGC1A, C8orf46, ADRA1A, MLLT3, ARPP21</i>	5.81	< 0.001	0.040	0.067
CNGTAWNTG	MSX1	173	2.24	13	<i>FOXP2, BCAR3, KCNAB1, ADAMTSL1, ELMO1, C12orf42, RUNX1T1, NR6A1, ARNTL, CBLN2, SND1, INHBA, IGF2BP3</i>	5.81	< 0.001	0.340	0.010
AGATAAGATAA	EVI1	55	0.71	4	<i>CREB5, CACNA1D, PPARGC1A, RUNX1T1</i>	5.62	0.014	< 0.001	0.324

Note. Shown are the top 10 significantly ($q < 0.05$) enriched transcription factor targets (motifs) in stage 4S hyper- and hypomethylated promoters. For each enriched motif, matching transcription factor (TF) annotation is shown (if available), as well as the number of genes in the category (C), the expected number in the category (E), the number of genes hyper- or hypomethylated in stage 4S in the category (O), the annotation for the genes in O, the ratio of enrichment (R) and the q-value. Global test q-values of the differential expression analyses are also indicated.

Table 3. Oncogenic signatures are enriched in stage 4S hypomethylated promoters.

gene set	enrichment analysis						differential expression analysis		description of gene set
	stage 4S versus stage 4			stage 4S versus stage 1/2			stage 4S versus stage 4	stage 4S versus stage 1/2	
	NES	nom p-val	FDR q-val	NES	nom p-val	FDR q-val	q	q	
BCAT_BILD_ET_AL_DN	1.38	0.033	0.053	1.47	0.002	0.026	0.002	0.205	genes downregulated in primary epithelial breast cancer cell culture overexpressing activated <i>CTNNB1</i> (PMID: 16273092)
BMI1_DN.V1_DN	1.30	0.025	0.088	1.30	0.002	0.102	0.002	< 0.001	genes downregulated in DAOY medulloblastoma cells upon knockdown of <i>BMI1</i> by RNAi (PMID: 17452456)
CAHOY_NEURONAL	1.49	0.001	0.042	1.27	0.021	0.131	< 0.001	< 0.001	genes upregulated in neurons (PMID: 18171944)
CRX_NRL_DN.V1_DN	1.38	0.006	0.056	1.23	0.030	0.182	< 0.001	< 0.001	genes downregulated in retina cells from <i>CRX</i> and <i>NRL</i> double knockout mice (PMID: 17653270)
EGFR_UP.V1_DN	1.40	0.001	0.056	1.26	0.007	0.131	< 0.001	< 0.001	genes downregulated in MCF7 breast cancer cells positive for <i>ESR1</i> and engineered to express ligand-activatable <i>EGFR</i> (PMID: 11888208)
EIF4E_DN	1.31	0.035	0.076	1.42	< 0.001	0.036	< 0.001	< 0.001	genes downregulated in HMEC primary mammary epithelium cells upon overexpression of <i>EIF4E</i> (PMID: 17638893)
ERB2_UP.V1_DN	1.25	0.041	0.145	1.41	< 0.001	0.038	< 0.001	0.005	genes downregulated in MCF7 breast cancer cells positive for <i>ESR1</i> and engineered to express ligand-activatable <i>ERBB2</i> (PMID: 11888208)
KRAS.AMP.LUNG_UP.V1_UP	1.37	0.006	0.063	1.30	0.006	0.113	0.003	< 0.001	genes upregulated in epithelial lung cancer cell lines overexpressing <i>KRAS</i> (PMID: 19847166)
KRAS.KIDNEY_UP.V1_UP	1.66	< 0.001	0.004	1.50	< 0.001	0.021	< 0.001	< 0.001	genes upregulated in epithelial kidney cancer cell lines overexpressing an oncogenic form of <i>KRAS</i>
MEK_UP.V1_DN	1.47	< 0.001	0.044	1.37	< 0.001	0.054	< 0.001	< 0.001	genes downregulated in MCF7 breast cancer cells positive for <i>ESR1</i> ; MCF7 cells stably overexpressing constitutively active <i>MAP2K1</i> (PMID: 16585219)

(continues)

Note. For each gene set significantly enriched in promoters hypomethylated in stage 4S compared to stage 4 and stage 1/2, the corresponding normalized enrichment score (NES), nominal p-value (nom p-val < 0.05) and false discovery rate q-value (FDR q-val < 0.25) are shown. Global test q-values of the differential expression analyses are also indicated, as well as a description of the gene set (with PubMed ID of the source publication; as described in the Molecular Signatures Database [18]).

Table 3. Oncogenic signatures are enriched in stage 4S hypomethylated promoters.

(continued)

gene set	enrichment analysis						differential expression analysis		description of gene set
	stage 4S versus stage 4			stage 4S versus stage 1/2			stage 4S versus stage 4	stage 4S versus stage 1/2	
	NES	nom p-val	FDR q-val	NES	nom p-val	FDR q-val	q	q	
MEL18_DN.V1_DN	1.36	0.004	0.062	1.20	0.040	0.244	< 0.001	0.006	genes downregulated in DAOY medulloblastoma cells upon knockdown of <i>PCGF2</i> gene by RNAi (PMID: 17452456)
PIGF_UP.V1_UP	1.31	0.009	0.076	1.38	< 0.001	0.053	< 0.001	0.005	genes upregulated in HUVEC endothelium cells by treatment with <i>PIGF</i> (PMID: 15516835)
PRC2_EDD_UP.V1_DN	1.30	0.014	0.088	1.32	< 0.001	0.106	< 0.001	< 0.001	genes downregulated in TIG3 fibroblast cells upon knockdown of <i>EED</i> (PMID: 16618801)
PRC2_EZH2_UP.V1_DN	1.33	0.003	0.071	1.21	0.034	0.209	0.021	< 0.001	genes downregulated in TIG3 fibroblast cells upon knockdown of <i>EZH2</i> (PMID: 16618801)
RAF_UP.V1_DN	1.39	< 0.001	0.056	1.30	< 0.001	0.105	0.002	< 0.001	genes downregulated in MCF7 breast cancer cells positive for <i>ESR1</i> ; MCF7 cells stably overexpressing constitutively active <i>RAF1</i> (PMID: 16585219)
RAF_UP.V1_UP	1.33	0.007	0.068	1.25	0.009	0.139	0.002	< 0.001	genes upregulated in MCF7 breast cancer cells positive for <i>ESR1</i> ; MCF7 cells stably overexpressing constitutively active <i>RAF1</i> (PMID: 16585219)
STK33_NOMO_UP	1.21	0.040	0.204	1.27	< 0.001	0.127	0.007	< 0.001	genes upregulated in NOMO1 acute myeloid leukemia cells after knockdown of <i>STK33</i> by RNAi (PMID: 19490892)
STK33_SKM_UP	1.27	0.007	0.115	1.28	< 0.001	0.120	0.002	< 0.001	genes upregulated in SKM1 acute myeloid leukemia cells after knockdown of <i>STK33</i> by RNAi (PMID: 19490892)
STK33_UP	1.33	< 0.001	0.070	1.31	< 0.001	0.109	0.004	0.004	genes upregulated in NOMO1 and SKM1 acute myeloid leukemia cells after knockdown of <i>STK33</i> by RNAi (PMID: 19490892)

Note. For each gene set significantly enriched in promoters hypomethylated in stage 4S compared to stage 4 and stage 1/2, the corresponding normalized enrichment score (NES), nominal p-value (nom p-val < 0.05) and false discovery rate q-value (FDR q-val < 0.25) are shown. Global test q-values of the differential expression analyses are also indicated, as well as a description of the gene set (with PubMed ID of the source publication; as described in the Molecular Signatures Database [18]).

hypomethylated in stage 4S compared to stage 4, in stage 4S compared to stage 1/2 and in stage 1/2 compared to stage 4 (Supplemental table 4). Of note, 19 gene sets (Table 3) were specifically enriched in promoters hypomethylated in stage 4S compared to both stage 1/2 and stage 4, including gene sets involved in downstream signaling of important oncogenes and epigenetic actors. Transcription factor target gene sets (Table 2) and the leading edge genes of the oncogenic signatures gene sets (Table 3) significantly enriched in stage 4S differentially methylated promoters were further tested for differential expression, using the global test [19] and publically available mRNA expression data [17]. In total, 25 of the 39 (64.10%) gene sets were identified as differentially expressed in stage 4S compared to both stage 4 and stage 1/2 (Table 2 and Table 3). Finally, using covariates plots [19] and differential gene expression analyses on the individual gene level, we evaluated which genes contributed the most to the significant global test results (Supplemental table 5 and Table 4).

Discussion

Genome-wide characterization studies in NB often combine patients with stage 4S or other low-risk tumors into one group for comparison to high-risk patients, as low-risk patients generally have an excellent outcome, which is in sharp contrast to high-risk patients who often show fatal progression. Patients with stage 4S tumors demonstrate a typical pattern of metastatic spread at diagnosis and have the capacity to undergo spontaneous regression, and these characteristics are not exhibited by stage 1/2 tumors. As genomics alone have not been able to explain this intriguing clinical presentation [6], we aimed to characterize the DNA methylation pattern of stage 4S NB, in order to reveal new biological insights into this special type of NB.

By performing differential methylation analyses

between stage 4S and stage 1/2 and stage 4, we could clearly define a unique stage 4S promoter methylation portrait. We further showed that stage 4S-specific hyper- and hypomethylated promoters are located on specific chromosomal locations. Non-random chromosomal distribution of regions differentially methylated between tumor and normal samples [20,21], and between different tumor stages [21,22], has been previously reported for other cancer types. In this study, we show that specific chromosomal cytogenetic bands are enriched in stage 4S differentially methylated promoters, and most importantly, that stage 4S tumors show characteristic hypermethylation of specific subtelomeric promoters. Telomere length and its regulatory mechanisms are well-studied in NB [11–14,23–25]. Most stage 4S tumors have low telomerase activity or short telomeres [11,12,23], and it has been suggested that the association between stage 4S and spontaneous regression might be related to differences in telomere length regulation [6]. While subtelomeric DNA methylation has been shown to suppress telomere elongation in other cancer types [15,16], such telomere length regulatory mechanism is unexplored in NB. In this study, we have shown that specific subtelomeric promoters are hypermethylated in stage 4S tumors compared to stage 1/2 and stage 4 tumors, which suggests that subtelomeric DNA methylation may represent an additional mechanism by which telomere length and spontaneous regression are controlled in NB. Therefore, the role of DNA methylation in telomere length regulation in stage 4S NB should definitely be further explored.

Further, we have shown that multiple transcription factor target gene sets are enriched in stage 4S differentially methylated promoters. Methylation of cytosines within transcription factor recognition motifs has been shown to cause structural changes influencing DNA-protein interactions (either in a positive or negative manner) and transcription

Table 4. Multiple stage 4S differentially methylated genes are also differentially expressed.

hypermethylated in stage 4S							
enriched gene set	gene	stage 4S versus stage 4			stage 4S versus stage 1/2		
		q	logFC	over-expression in	q	logFC	over-expression in
E2F1, E2F1_TFDP1_RB1 and MEF2A	<i>SLC9A5</i>	< 0.001	0.833	stage 4	0.002	0.500	stage 1/2
hypomethylated in stage 4S							
enriched gene set	gene	stage 4S versus stage 4			stage 4S versus stage 1/2		
		q	logFC	over-expression in	q	logFC	over-expression in
CAHOY_NEURONAL	<i>DYNC1I1</i>	0.010	1.035	stage 4	< 0.001	1.517	stage 1/2
	<i>HS3ST2</i>	0.021	0.680	stage 4	0.031	0.589	stage 1/2
	<i>NSF</i>	0.006	0.470	stage 4	< 0.001	0.538	stage 1/2
	<i>GPR22</i>	< 0.001	1.272	stage 4S	0.016	0.847	stage 4S
	<i>CALB1</i>	< 0.001	1.718	stage 4S	0.048	0.850	stage 4S
CRX_NRL_DN.V1_DN	<i>GPR124</i>	0.011	0.471	stage 4S	0.050	0.344	stage 4S
	<i>PVRL3</i>	< 0.001	1.153	stage 4S	0.009	0.719	stage 4S
EGFR_UP.V1_DN	<i>C9orf3</i>	< 0.001	0.833	stage 4S	0.011	0.574	stage 4S
	<i>DERA</i>	< 0.001	0.897	stage 4S	0.023	0.451	stage 4S
EGR3	<i>ATP6V1C1</i>	0.002	0.380	stage 4	< 0.001	0.444	stage 1/2
	<i>CLTC</i>	0.008	0.391	stage 4	0.002	0.419	stage 1/2
	<i>GPR176</i>	0.031	0.184	stage 4	0.018	0.185	stage 1/2
	<i>UBE2W</i>	0.048	0.177	stage 4	0.031	0.177	stage 1/2
EIF4E_DN	<i>CLTC</i>	0.008	0.391	stage 4	0.002	0.419	stage 1/2
	<i>NPEPPS</i>	0.002	0.348	stage 4	0.004	0.305	stage 1/2
	<i>TP53BP1</i>	0.027	0.287	stage 4	0.010	0.304	stage 1/2
	<i>LMO7</i>	0.001	0.490	stage 4S	0.028	0.323	stage 4S
	<i>SH3YL1</i>	0.004	0.428	stage 4S	0.049	0.282	stage 4S
ERB2_UP.V1_DN	<i>NPEPPS</i>	0.002	0.348	stage 4	0.004	0.305	stage 1/2
	<i>C9orf3</i>	< 0.001	0.833	stage 4S	0.011	0.574	stage 4S
	<i>LAMB1</i>	< 0.001	0.690	stage 4S	0.022	0.354	stage 4S
KRAS.AMP.LUNG_UP.V1_UP	<i>DYNC1I1</i>	0.010	1.035	stage 4	< 0.001	1.517	stage 1/2
	<i>C8B</i>	0.038	0.340	stage 4S	0.006	0.408	stage 4S
	<i>GCKR</i>	0.027	0.240	stage 4S	0.032	0.215	stage 4S
KRAS.KIDNEY_UP.V1_UP	<i>DYNC1I1</i>	0.010	1.035	stage 4	< 0.001	1.517	stage 1/2
	<i>ESRRG</i>	0.021	0.475	stage 4	0.018	0.449	stage 1/2
	<i>CALB1</i>	< 0.001	1.718	stage 4S	0.048	0.850	stage 4S
MEK_UP.V1_DN	<i>NPEPPS</i>	0.002	0.348	stage 4	0.004	0.305	stage 1/2
	<i>C9orf3</i>	< 0.001	0.833	stage 4S	0.011	0.574	stage 4S
	<i>TCF12</i>	0.001	0.370	stage 4S	0.015	0.261	stage 4S
MEL18_DN.V1_DN	<i>DIO2</i>	0.048	0.460	stage 4	0.019	0.497	stage 1/2
PIGF_UP.V1_UP	<i>LRRC17</i>	< 0.001	0.568	stage 4S	0.031	0.283	stage 4S
PRC2_EDD_UP.V1_DN	<i>CBX8</i>	< 0.001	0.815	stage 4	0.002	0.588	stage 1/2

(continues)

Note. For each gene set differentially expressed in stage 4S compared to stage 4 and stage 1/2 (Table 2 and Table 3), differentially expressed individual genes are shown. Q-values are indicated, as well as the log2 fold change (logFC) and the group in which the gene is overexpressed.

Table 4. Multiple stage 4S differentially methylated genes are also differentially expressed.

(continued)

enriched gene set	gene	hypomethylated in stage 4S stage 4S versus stage 4			stage 4S versus stage 1/2		
		p	logFC	over-expression in	p	logFC	over-expression in
RAF_UP.V1_DN	<i>C9orf3</i>	< 0.001	0.833	stage 4S	0.011	0.574	stage 4S
	<i>PALLD</i>	0.042	0.330	stage 4S	0.019	0.347	stage 4S
	<i>TCF12</i>	0.001	0.370	stage 4S	0.015	0.261	stage 4S
RAF_UP.V1_UP	<i>CABYR</i>	0.003	0.462	stage 4	< 0.001	0.589	stage 1/2
	<i>GATA2</i>	0.032	0.711	stage 4S	0.012	0.760	stage 4S
STK33_NOMO_UP	<i>KIAA1467</i>	< 0.001	0.560	stage 4	0.002	0.447	stage 1/2
	<i>NPEPPS</i>	0.002	0.348	stage 4	0.004	0.305	stage 1/2
	<i>UBE2W</i>	0.048	0.177	stage 4	0.031	0.177	stage 1/2
STK33_SKM_UP	<i>UBE2W</i>	0.048	0.177	stage 4	0.031	0.177	stage 1/2
	<i>PALLD</i>	0.042	0.330	stage 4S	0.019	0.347	stage 4S
STK33_UP	<i>UBE2W</i>	0.048	0.177	stage 4	0.031	0.177	stage 1/2

Note. For each gene set differentially expressed in stage 4S compared to stage 4 and stage 1/2 (Table 2 and Table 3), differentially expressed individual genes are shown. Q-values are indicated, as well as the log2 fold change (logFC) and the group in which the gene is overexpressed.

factor activity [26]. As such, DNA methylation serves as a regulatory mechanism for transcription factor activity. Therefore, the transcription factor target gene sets enriched in stage 4S differentially methylated promoters may indicate differential activity of the respective transcription factors in stage 4S tumors compared to stage 1/2 and stage 4 tumors. Of note, this assumption is reinforced by the fact that several of the identified transcription factors of which multiple target genes are differentially methylated in stage 4S show proven dependency on the DNA methylation status of the binding site for transcription factor- DNA interaction. For example, DNA methylation inhibits the binding of E2F1 and CEBPG [26–28]. Additionally, the differentially methylated target gene sets are differentially expressed in stage 4S NB, further strengthening the possibility that DNA methylation affects the activity of these transcription factors. Of note, E2F1 also binds to the *TERT* promoter [29], which is hypermethylated in stage 4S compared to stage 4, and *TERT* is expressed at lower levels in stage 4S compared to stage 4. This indicates that *TERT* DNA methylation might also regulate telomerase

activity in stage 4S NB.

Furthermore, our transcription factor target enrichment analyses particularly demonstrate that the DNA methylation portrait of stage 4S NB is dominated by differential methylation of target genes of transcription factors involved in neural crest development and (sympathetic) neural differentiation (MSX1 [30,31], EVI1 [32,33], EGR3 [34], AHR [35,36], MEF2A [37,38], YY1 [39], PPARA [40], POU2F1 [41] and GF11 [42,43]). These findings suggest that differences in the pathogenic mechanisms between stage 4S and stage 1/2 and stage 4 tumors are at least in part related to differences in transcriptional activity during neural crest development and differentiation, which might indicate that stage 4S, stage 1/2 and stage 4 tumors arise during different phases of neural crest cell development. This is in line with the previous postulation that stage 4S NB originates from defective premigratory neural crest cells, which explains its clinical symptomology, as stage 4S tumor sites mirror developmental neural crest cell migration [44].

One of the genes contributing the most to the significant differential expression of the target genes of E2F1 and MEF2A is *SLC9A5*. The *SLC9A5*

promoter is hypermethylated in stage 4S and these tumors show low *SLC9A5* expression levels compared to stage 4 and stage 1/2 tumors. *SLC9A5* is a member of the Na(+)/H(+) exchanger family and is abundantly expressed in neurons. In rat pheochromocytoma, a neuroendocrine tumor of the adrenal medulla, Diering et al. have shown that *SLC9A5* modulates trafficking of TrkA between recycling endosomes and the plasma membrane, neurotrophin signaling and neural differentiation [45]. The authors demonstrate that upon knockdown of *SLC9A5* in these tumor cells, there is no difference in the total abundance of TrkA compared to control cells, suggesting that reduced expression of *SLC9A5* has little effect on TrkA degradation or synthesis. However, although the overall TrkA protein levels did not change, the cell surface abundance of TrkA was reduced upon knockdown of *SLC9A5* [45]. Stage 4S NB tumors generally show high levels of TrkA, and depending on the presence or absence of nerve growth factor (NGF) in the tumor microenvironment, the tumors undergo neural differentiation or spontaneous regression, respectively. Therefore, neurotrophin deprivation (TrkA without NGF) and activation of apoptosis have been proposed to explain the phenomenon of spontaneous regression in NB [6]. Given the crucial role of *SLC9A5* in regulating the cellular distribution of TrkA and NGF signaling [45], our findings suggest that over time stage 4S tumor cell surfaces might get depleted of TrkA (with an accompanying intracellular accumulation of TrkA) due to the low expression levels of *SLC9A5*, ultimately leading to apoptosis and regression.

Our gene set enrichment analyses also highlight differential methylation of genes involved in important oncogenic pathways, especially of genes implicated in EGFR and RAS signaling. Although the exact mechanisms by which DNA methylation contributes to differential regulation of these pathways in stage 4S specifically cannot be determined in our study, our findings indicate that it is crucial to separate

stage 4S tumors from other low-stage (stage 1/2) tumors in studies focusing on NB biology, as this might reveal new insights into the stage 4S NB pathogenesis. The importance hereof is highlighted by the fact that all but one differentially methylated oncogenic signatures are also differentially expressed in stage 4S NB, confirming differential regulation of these oncogenic pathways in stage 4S on the one hand and stage 4 and stage 1/2 on the other hand. Finally, the enrichment analyses also illustrate that stage 4S tumors present differential activity of important epigenetic players, such as BMI1, PCGF2, EED and EZH2. Potentially, differential regulation of epigenetic mechanisms might also contribute to the regression phenotype of stage 4S NB, as epigenetic alterations can more easily be reversed. Clearly, these epigenetic mechanisms need to be further studied in NB.

In conclusion, this study is the first to describe the stage 4S DNA methylome and demonstrates that stage 4S NB tumors show characteristic DNA methylation patterns compared to stage 1/2 and stage 4 tumors. Genomic localization and functional characterization analyses of this stage 4S DNA methylation portrait suggested new mechanisms that may contribute to the stage 4S-specific tumor biology and spontaneous regression. Future studies must confirm these observations and hypotheses.

Materials and methods

Neuroblastoma primary tumors

Sample collection for MBD sequencing (n = 41) and expression profiling (n = 180) is described in Decock et al. [7] and Oberthuer et al. [17], respectively. All tumor samples were assigned to one of the three following subgroups, based on the INSS tumor stage [3] and *MYCN* amplification status: (1) *MYCN* non-amplified stage 4S (MBD sequencing: n = 14; expression profiling: n = 28), (2) *MYCN* non-amplified stage 4 (MBD sequencing: n = 14; expression profiling: n = 46),

or (3) *MYCN* non-amplified stage 1 or 2 (MBD sequencing: $n = 13$; expression profiling: $n = 106$). Detailed criteria for patient inclusion in the study are described in Decock et al. [8,46]. *MYCN* amplified tumor samples were excluded. Detailed clinicobiological characteristics of the different sample cohorts are shown in Supplemental table 6. The study was approved by the ethical committee of the Ghent University Hospital (approval number: B67020109912).

DNA methylome and transcriptome profiling

A detailed description of the generation and raw data analysis of the MBD sequencing data is provided in Decock et al. [7]. Data can be accessed through the Gene Expression Omnibus numbers GSE69243 and GSE69268. For the mRNA expression analyses, microarray data of 180 NB tumors were used [17].

Bioinformatics and statistical analyses

Differential methylation analyses on the MBD sequencing data were performed as described in Decock et al. [8], using DESeq [47] and promoter region count matrices with data on all Ensembl transcripts (release 68). Promoters located on X and Y chromosomes were excluded to avoid gender-related confounding. To estimate the background signal and to be able to exclude possible falsely identified differentially methylated regions (DMRs) in the DESeq analyses, a filtering procedure based on two correction steps was applied on the resulting DMR lists. A first correction step consisted of excluding DMRs which were also significantly identified in the same comparison on the input (non-MBD-enriched) sample data (input correction) [8]. Secondly, computational simulation analyses were performed, in which 100 random patient annotation swaps were executed to create two artificial sample groups on which the same statistics were applied as on the biologically relevant groups. Only biologically

relevant DMRs with a π -value ($= -\ln p\text{-val} * \log_2$ fold change) [48] that equals or outperforms the π -value of the corresponding simulation DMR were considered for further evaluation (swap correction). DESeq normalized read counts of the stage-specific DNA methylation portraits were used for principal component analysis in R (package *factoextra*).

Cytogenetic band enrichment analysis and transcription factor target enrichment analysis were performed using WebGestalt [9,10] with gene symbols as input list. The entire human genome was used as reference list. Benjamini-Hochberg multiple testing correction was applied [49] and only statistically enriched terms ($q < 0.05$) with at least two genes were considered.

Non-corrected transcript lists, including all transcripts included in the differential methylation analyses (preranked on the π -value) were used for GSEA [18,50]. To this purpose, Ensembl transcript IDs were converted to gene symbols. For genes with multiple transcripts, the data of the Ensembl transcript with the highest absolute π -value was retained. Tested gene set databases are the oncogenic signatures (version 5.1) of the Molecular Signatures Database [18] and two in-house created subtelomere gene set lists. These two subtelomere gene set lists contain all genes in the 1 Mb or 2 Mb subtelomeres, which are defined as the 1 Mb or 2 Mb genomic regions [24] proximal to the telomeres (10 kb chromosomal ends; hg19). GSEA was also performed on the input data to exclude falsely identified enriched gene sets.

The sets of differentially methylated transcription factor targets and the leading edge genes of the enriched oncogenic signature gene sets were further tested for differential expression using the *globaltest* R package [19] and publically available mRNA expression data [17]. For significantly differentially expressed gene sets, covariates plots [19] were generated to inspect the contributions of each gene to the test result. Differential expression analyses on

the individual gene level were performed using the R package limma [51]. Benjamini-Hochberg multiple testing correction was applied [49].

Acknowledgments

This study is supported by the Fournier-Majoie Foundation (FFM), the Belgian National Lottery, the Flemish League Against Cancer (VLK), UGent GOA (grant number 01G01910) and the ISCIII Spain & FEDER (grant number PI14/01008). AD was a recipient of a grant of the Research Foundation Flanders (FWO) and an Emmanuel Van der Schueren research grant (VLK).

Supplemental material

Supplemental materials are available via the following link and password:

<https://www.dropbox.com/sh/bf9651jk3uplg14/AABiVI2Nq1vyZTvEbuceRTBUa?dl=0>
password: PhDthesisADC

Supplemental table 1. Neuroblastoma tumors are characterized by a stage-specific DNA methylation pattern. For each Ensembl transcript ID, gene and promoter annotation is given, as well as the corresponding p-value, log2 fold change and π -value ($= -\ln p\text{-value} * \log_2 \text{fold change}$) of the differential methylation analyses. For transcripts considered differentially methylated upon input and swap correction (see materials and methods), it is indicated which sample group is hypermethylated and whether the transcript is part of a stage-specific DNA methylation portrait.

Supplemental table 2. Specific chromosomal cytogenetic bands are enriched in the stage-specific differentially methylated promoters. Shown are the significantly ($q < 0.05$) enriched cytogenetic bands in stage 4S, stage 4 and stage 1/2 hyper- and hypomethylated promoters. For each enriched gene set, the differentially methylated genes are indicated. C: the number

of reference genes in the category (cytogenetic band); E: the expected number in the category; O: the number of genes hyper- or hypomethylated in the corresponding stage in the category; genes: annotation for the genes in O; R: ratio of enrichment.

Supplemental table 3. Specific transcription factor targets are enriched in stage 4S differentially methylated promoters. Shown are the significantly ($q < 0.05$) enriched transcription factor targets (motifs) in stage 4S hyper- and hypomethylated promoters. C: the number of reference genes in the category (transcription factor target); E: the expected number in the category; O: the number of genes hyper- or hypomethylated in stage 4S in the category; genes: annotation for the genes in O; R: ratio of enrichment.

Supplemental table 4. Gene set enrichment analyses identify oncogenic signatures in stage-related differentially methylated promoters. Shown are the gene sets significantly (nom p-val < 0.05 and FDR q-val < 0.25) enriched in promoters hypomethylated in stage 4S compared to stage 4, in stage 4S compared to stage 1/2 or in stage 1/2 compared to stage 4. Gene sets also identified in the corresponding input data comparisons are excluded. NES: normalized enrichment score; nom p-val: nominal p-value; FDR q-val: false discovery rate q-value.

Supplemental table 5. Transcription factor target and oncogenic signatures gene sets are differentially expressed between stage 4S and stage 4 and stage 1/2. For each gene set significantly differentially expressed ($q < 0.05$; global test) between stage 4S and stage 4 and/or stage 1/2, data extracted from the covariates plots (p-values and sample group associations of individual genes) is shown. For each of the genes in all gene sets, also the results (q-value, log2 fold change and sample group with

overexpression) of the differential expression analysis on the individual gene level (limma) are shown.

Supplemental table 6. In total, 221 annotated primary neuroblastoma samples were included in the study.

A. Detailed characteristics. Each sample is characterized by a unique patientID and is assigned to a biological subgroup (stage 4S, stage 1/2 or stage 4) and subcohort (MBD sequencing or expression). Clinical characteristics given are the age of the patient at diagnosis in months, International Neuroblastoma Staging System (INSS) stage, and overall survival (OS) and event-free survival (EFS) status and time after diagnosis in days. The OS status indicates whether the patient is alive (0) at the last known follow-up or died of disease (1). Similarly, the EFS status indicates events such as relapse, progression or death. All samples are *MYCN* non-amplified. NAs represent missing values. **B.** Summary. Per subcohort an overview of the clinical characteristics is given.

References

- [1] Evans AE et al. (1971). A proposed staging for children with neuroblastoma. Children's cancer study group A. *Cancer*; 27(2):374-378.
- [2] Evans AE et al. (1976). Factors influencing survival of children with nonmetastatic neuroblastoma. *Cancer*; 38(2):661-666.
- [3] Brodeur GM et al. (1993). Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *Journal of Clinical Oncology*; 11(8):1466-1477.
- [4] London WB et al. (2005). Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's Oncology Group. *Journal of Clinical Oncology*; 23(27):6459-6465.
- [5] Monclair T et al. (2009). The International Neuroblastoma Risk Group (INRG) staging system: an INRG Task Force report. *Journal of Clinical Oncology*; 27(2):298-303.
- [6] Brodeur GM et al. (2014). Mechanisms of neuroblastoma regression. *Nature Reviews. Clinical Oncology*; 11(12):704-713.
- [7] Decock A et al. (2016). DNA methylation profiling of primary neuroblastoma tumors using methyl-CpG-binding domain sequencing. *Scientific Data*; 3:160004.
- [8] Decock A et al. (2016). Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncotarget*; 7(2):1960-1972.
- [9] Zhang B et al. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Research*; 33:W741-748.
- [10] Wang J et al. (2013). WEB-based GENE SeT Analysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Research*; 41:W77-83.
- [11] Hiyama E et al. (1995). Correlating telomerase activity levels with human neuroblastoma outcomes. *Nature Medicine*; 1(3):249-255.
- [12] Onitake Y et al. (2009). Telomere biology in neuroblastoma: telomere binding proteins and alternative strengthening of telomeres. *Journal of Pediatric Surgery*; 44(12):2258-2266.
- [13] Valentijn LJ et al. (2015). TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nature Genetics*; 47(12):1411-1414.
- [14] Peifer M et al. (2015). Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature*; 526(7575):700-704.
- [15] Blasco MA (2007). The epigenetic regulation of mammalian telomeres. *Nature Reviews. Genetics*; 8(4):299-309.
- [16] Vera E et al. (2008). Epigenetic regulation of telomeres in human cancer. *Oncogene*; 27(54):6817-6833.
- [17] Oberthuer A et al. (2006). Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. *Journal of Clinical Oncology*; 24(31):5070-5078.
- [18] Subramanian A et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the USA*; 102(43):15545-15550.
- [19] Goeman JJ et al. (2004). A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics*; 20(1):93-99.
- [20] Gómez S et al. (2015). DNA methylation fingerprint of neuroblastoma reveals new biological and clinical insights. *Epigenomics*; 7(7):1137-1153.
- [21] Figueroa ME et al. (2009). MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. *Blood*; 114(16):3448-3458.

- [22] Heuck CJ et al. (2013). Myeloma is characterized by stage-specific alterations in DNA methylation that occur early during myelomagenesis. *Journal of Immunology*; 190(6):2966-2975.
- [23] Krams M et al. (2003). Full-length telomerase reverse transcriptase messenger RNA is an independent prognostic factor in neuroblastoma. *American Journal of Pathology*; 162(3):1019-1026.
- [24] Buckley PG et al. (2011). Genome-wide DNA methylation analysis of neuroblastic tumors reveals clinically relevant epigenetic events and large-scale epigenomic alterations localized to telomeric regions. *International Journal of Cancer*; 128(10):2296-2305.
- [25] Hertwig F et al. (2016). Telomere maintenance is pivotal for high-risk neuroblastoma. *Cell Cycle*; 15(3):311-312.
- [26] Blattler A et al. (2013). Cross-talk between site-specific transcription factors and DNA methylation states. *Journal of Biological Chemistry*; 288(48):34287-34294.
- [27] Campanero MR et al. (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proceedings of the National Academy of Sciences of the USA*; 97(12):6481-6486.
- [28] Mann IK et al. (2013). CG methylated microarrays identify a novel methylated sequence bound by the CEBPB|ATF4 heterodimer that is active in vivo. *Genome Research*; 23(6):988-997.
- [29] Alonso MM et al. (2005). Expression of transcription factor E2F1 and telomerase in glioblastomas: mechanistic linkage and prognostic significance. *Journal of the National Cancer Institute*; 97(21):1589-1600.
- [30] Tribulo C et al. (2004). A balance between the anti-apoptotic activity of Slug and the apoptotic activity of msx1 is required for the proper development of the neural crest. *Developmental Biology*; 275(2):325-342.
- [31] Monsoro-Burq AH et al. (2005). Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during *Xenopus* neural crest induction. *Developmental Cell*; 8(2):167-178.
- [32] Hoyt PR et al. (1997). The Evi1 proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development. *Mechanisms of Development*; 65(1-2):55-70.
- [33] Kazama H et al. (1999). Ecotropic viral integration site-1 is activated during, and is sufficient for, neuroectodermal P19 cell differentiation. *Cell Growth and Differentiation*; 10(8):565-573.
- [34] Quach DH et al. (2013). A sympathetic neuron autonomous role for Egr3-Mediated gene regulation in dendrite morphogenesis and target tissue innervation. *Journal of Neuroscience*; 33(10):4570-4583.
- [35] Akaoshi E et al. (2006). Over-expression of AhR (aryl hydrocarbon receptor) induces neural differentiation of Neuro2a cells: neurotoxicology study. *Environmental Health*; 5:24.
- [36] Wu PY et al. (2014). Aryl hydrocarbon receptor downregulates MYCN expression and promotes cell differentiation of neuroblastoma. *PLoS One*; 9(2):e88795.
- [37] Ruffle RA et al. (2006). Distribution of constitutively expressed MEF2A in adult rat and human nervous systems. *Synapse*; 59(8):513-520.
- [38] Mao Z et al. (1996). Functional and physical interactions between mammalian achaete-scute homolog 1 and myocyte enhancer factor 2A. *Journal of Biological Chemistry*; 271(24):14371-14375.
- [39] He Y et al. (2008). The Yin and Yang of YY1 in the nervous system. *Journal of Neurochemistry*; 106(4):1493-1502.
- [40] Bento-Abreu A et al. (2007). Peroxisome proliferator-activated receptor- α is required for the neurotrophic effect of oleic acid in neurons. *Journal of Neurochemistry*; 103(3):871-881.
- [41] Dick T et al. (1991). Two closely linked *Drosophila* POU domain genes are expressed in neuroblasts and sensory elements. *Proceedings of the National Academy of Sciences of the USA*; 88(17):7645-7649.
- [42] Ossovskaya VS et al. (2009). Loss of function genetic screens reveal MTGR1 as an intracellular repressor of β 1 integrin-dependent neurite outgrowth. *Journal of Neuroscience Methods*; 177(2):322-333.
- [43] Cameron S et al. (2002). PAG-3, a Zn-finger transcription factor, determines neuroblast fate in *C. elegans*. *Development*; 129(7):1763-1774.
- [44] van Noesel MM (2012). Neuroblastoma stage 4S: a multifocal stem-cell disease of the developing neural crest. *Lancet Oncology*; 13(3):229-230.
- [45] Diering GH et al. (2013). Endosomal acidification by Na⁺/H⁺ exchanger NHE5 regulates TrkA cell-surface targeting and NGF-induced PI3K signaling. *Molecular Biology of the Cell*; 24(21):3435-3448.
- [46] Decock A et al. (2012). Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. *Genome Biology*; 13(10):R95

- [47] Anders S et al. (2010). Differential expression analysis for sequence count data. *Genome Biology*; 11(10):R106.
- [48] Xiao Y et al. (2014). A novel significance score for gene selection and ranking. *Bioinformatics*; 30(6):801-807.
- [49] Benjamini Y et al. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*; 57(1):289-300.
- [50] Mootha VK et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*; 34(3):267-273.
- [51] Ritchie ME et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*; 43(7):e47.

4



**Discussion and
future perspectives**

4.1 Defeating pitfalls and limitations in biomarker research: the successful establishment of a prognostic DNA methylation signature for neuroblastoma

Aiming at personalized medicine, molecular biomarker research is extremely important, especially for diseases showing a high degree of clinical heterogeneity such as NB. Revolutions in molecular characterization of NB tumors have led to the identification of a large amount of putative prognostic biomarkers, but only very few have shown clinical validity and utility; only *MYCN* oncogene amplification (§1.2.1.4), chromosome 11q aberration (§1.2.1.5.2.2) and DNA ploidy (§1.2.1.5.1) status are standardly used in patient stratification (§1.2.4) [154]. One of the main reasons for this is that many discovery studies suffer from pitfalls and limitations, such as an inadequate study design, insufficient sample numbers and lack of biomarker validation. In our search for prognostic DNA methylation biomarkers for NB, we aimed to defeat these issues. An overview of the research results is given in Figure 14.

At the time of project initiation, most DNA methylation biomarker studies in NB were candidate gene-based; only few made use of whole-genome DNA methylation detection methodologies (re-expression profiling upon treatment with DAC, MeDIP chip and Illumina 27 K methylation array analysis (§review paper; [228]); Figure 14). Though, whole-genome sequencing technologies had not yet been applied to study DNA methylation patterns in NB. By the start of our project, the implementation of the Illumina HM450 array (§1.1.4.2) was not yet fully optimized and whole-genome bisulfite sequencing (§1.1.4.4) was not yet cost-efficient. Therefore, given the expertise of the UGent sequencing reference center NXTGNT (www.nxtgnt.ugent.be) in MBD sequencing (§1.1.4.3), this NGS-based technology was selected as the method of choice for DNA methylation biomarker discovery in our project. As proof of principle, we set up a pilot study (§paper 1; [229]; Figure 14) in which eight NB cell lines were profiled using microarray re-expression analysis and MBD sequencing. Using these cell line data, a bioinformatics pipeline for the analysis of MBD sequencing data was developed and methylated promoters were identified. To further select candidate prognostic biomarkers, these data were integrated with publically available primary tumor expression data and locus-specific DNA methylation analysis in a limited number of primary tumors. Subsequently, fast and accurate validation of the candidate prognostic biomarkers was obtained through the implementation of a semi-automated, high-throughput MSP pipeline [229]. Through the successful establishment of this validation technology, we ensured the possibility to translate the identified biomarkers into a clinically useful assay, as at present, MSP is still considered the gold standard methodology for measuring DNA methylation in clinical assays [120]. Although NGS may hold the future for clinical assay testing, NGS-based methodologies are currently immature to be clinically implemented for DNA methylation biomarker testing [120, 232–234]. Encouraged by the results of these initial experiments [229] and having specified the technological aspects of our study design, we set out a more profound DNA methylation biomarker discovery, verification and validation trajectory, a valorization project supported by the Fournier-Majoie Foundation (www.fournier-majoie.org).

In the extended discovery phase, MBD sequencing was applied to 87 primary tumor DNA samples, to allow differential methylation analysis between prognostic patient groups and candidate biomarker prioritization (§paper 2; [230]; Figure 14). Next, more than 350 primary tumor DNA samples were collected for biomarker verification and validation. It should be noted that such large numbers of

samples are unprecedented in NB DNA methylation biomarker research. For rare diseases such as NB, one of the bottlenecks biomarker research faces, is the small number of available biospecimens. Consequently, our biomarker project depended on the establishment of international collaborations, in order to collect sample sizes that were large enough to be of statistical significance. During the

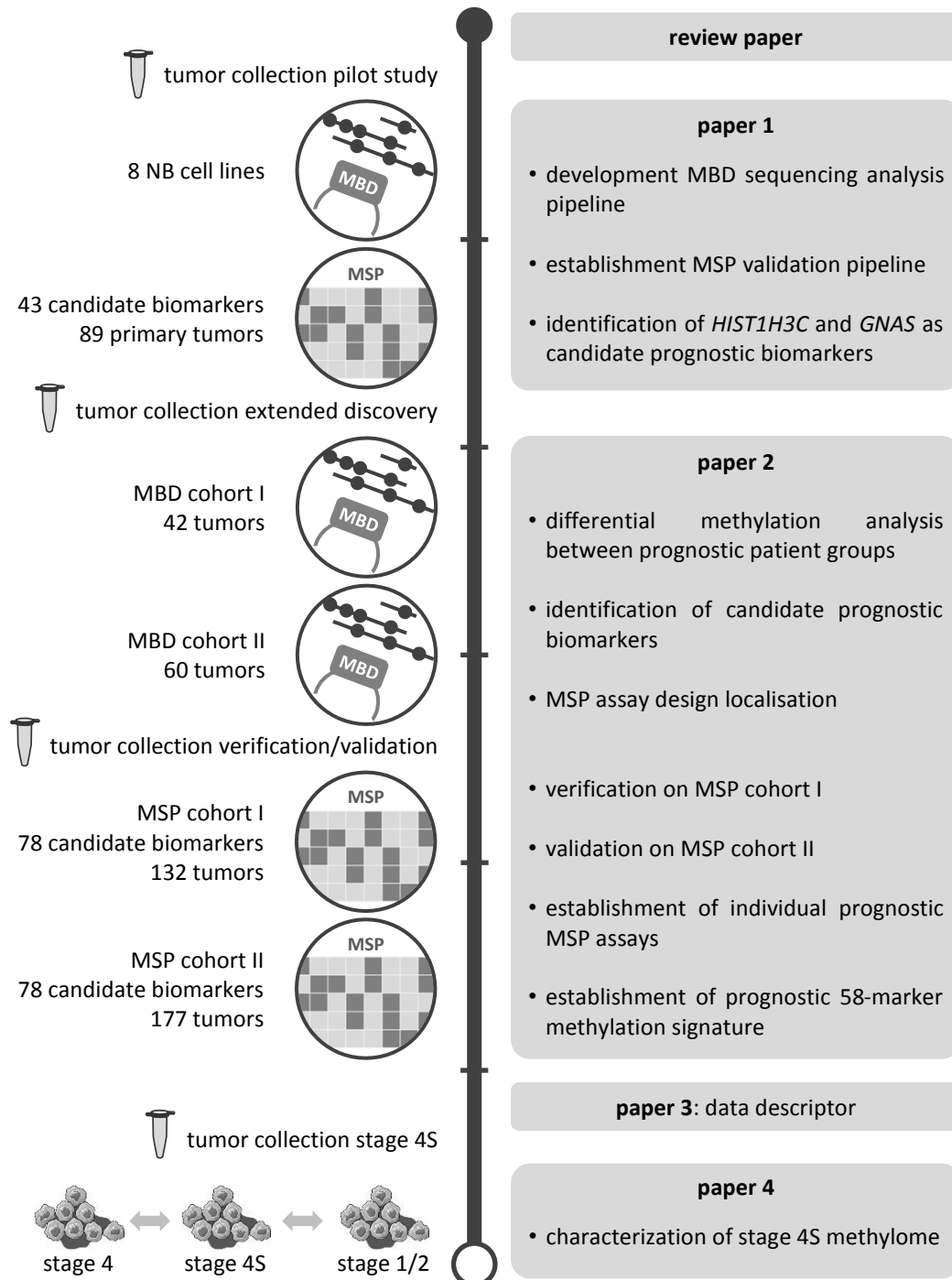


Figure 14. The research project resulted in five publications. At the start of the project, the literature on prognostic DNA methylation biomarkers in neuroblastoma (NB) was reviewed [228]. The first original research paper represents a pilot study in which the technological aspects of the project were specified [229]. The second paper describes the results of the more profound discovery, verification and validation trajectory [230]. The methyl-CpG-binding domain (MBD) sequencing data that were used in this second paper were also described in detail in a data descriptor [110]. Finally, a fourth paper was written on the characterization of the stage 4S NB methylome [231]. MSP: methylation-specific PCR.

project, multiple research centers and biobanks across the world were contacted, sample request applications were filed, material transfer agreements were set up, biospecimens were moved across national borders and sample annotation was gathered. In total, nine national biobank resource centers and reference centers for the study and diagnosis of NB contributed to sample collection for this biomarker project and enabled access to biospecimens: the Ghent University Hospital (Belgium), the Centre Léon Bérard (Lyon, France), the Hospital Clínico Universitario (Valencia, Spain), the Sydney Children's Hospital (Australia), the Institut Curie (Paris, France), the Children's Cancer and Leukemia Group (Leicester, UK), the Our Lady's Children's Hospital Dublin (Ireland), the University Hospital Brno (Czech Republic) and the University Children's Hospital Essen (Germany). Of note, the investments it took to build up this sample collection prolonged the project's time frame to a considerable degree, especially due to administrative hurdles that needed to be tackled. To speed up NB biomarker studies in the future, it is important to intensify the effort and resources put in expanding the infrastructure of international biobanks and databases, such as the interactive International Neuroblastoma Risk Group database (iINRGdb) [154]. Such repositories of high quality samples with updated clinical and treatment annotation should remarkably facilitate sample application and collection procedures for research projects.

Partly as a consequence of difficulties in sample collection, validation of the clinical significance of the identified biomarkers is often overlooked in NB studies, especially in relatively new and rapidly evolving research fields such as epigenetics. Unfortunately, this is a general trend in cancer biomarker research and has far-reaching consequences researchers often do not take into account when publishing unvalidated results [235, 236]. For example, publication of a new, high-profile biomarker is frequently accompanied by press releases generating high expectations about the new biomarker. However, follow-up reports demonstrating that the same biomarker failed to be validated for clinical use are generally ignored by the media, leaving the general public with skewed information. To improve biomarker research, it might be appropriate not to call a molecule a biomarker until it has passed at least one independent validation study, avoiding misinterpretations and highly misleading conclusions [235]. In our project, the DNA methylation biomarker candidates identified by primary tumor MBD sequencing were additionally tested on independent tumor cohorts using MSP, allowing to validate their clinical significance (paper 2; [230]; Figure 14). Importantly, we also reported in detail on patient and biospecimen characteristics, the assay methods, study design and statistical analysis methods [110, 230]. This is extremely important, as in addition to lack of independent validation, inadequate reporting also hinders biomarker research progress [237].

As such, we were able to identify and validate multiple individual prognostic MSP assays (located in *CCDC177*, *NXPH1*, *Inc-MRPL3-2*, *Inc-TREX1-1*, *SPRED3*, *TNFAIP2*, *NPM2*, *CYYR1* and one on a region from chromosome 8 with no further annotation), as well as a robust prognostic 58-marker methylation signature of which a methylation score cutoff of 25% allows accurate OS and EFS prediction in the global NB patient cohort. Importantly, in the validation cohort, this signature was an independent predictor of OS after controlling for known NB risk factors, clearly indicating its clinical relevance. Further investigations must confirm these results and integration with other biomarkers should be explored [230] (Figure 14).

4.2 Finding prognostic biomarkers for high-risk neuroblastoma: a tough nut to crack

Having collected a considerable number of both high-risk survivors and high-risk non-survivors for our study, we also tested the prognostic relevance of the selected MSP assays within the high-risk cohorts (Spaper 2; [230]). This analysis is very valuable for the reason that the need for prognostic biomarkers is the highest within this group of patients. Ideally, high-risk prognostic biomarkers would allow pretreatment identification of ultra-high-risk patients, i.e. patients currently dying of the disease although intensively being treated. Recent comprehensive whole-genome sequencing studies have shown that high-risk NB tumors can be genetically subdivided into three major subgroups, based on the presence of *MYCN* amplification, *TERT* rearrangements and *ATRX* alterations, but currently these findings have not shown utility as prognostic biomarker [238, 239]. Previously, our center developed a 59-mRNA [222] and 25-miRNA signature [224] using RNA from fresh frozen tumors, each enabling identification of patients with increased risk in the current risk groups (§1.2.4), including the high-risk patient group, but these signatures await further validation on formalin-fixed paraffin-embedded tumor samples. Next to tumor mRNAs and miRNAs, also tumor DNA methylation patterns have been searched for prognostic biomarkers for high-risk NB [240, 241]. However, these DNA methylation studies could not validate the biomarkers [241] or incorrectly included low-risk samples in the so-called validation cohort [240]. Unfortunately, also in our project we could not validate prognostic high-risk biomarkers using MSP, although differential methylation analyses on the MBD sequencing data illustrated that high-risk survivors and high-risk non-survivors show different methylation patterns. Importantly, this does not mean that high-risk DNA methylation biomarkers cannot be found. It only indicates that the methylation differences in the differentially methylated regions (of 2 kb or 5 kb) in the MBD sequencing data between these high-risk groups are too subtle to be easily translated in an MSP assay which only interrogates a few CpGs. Therefore, the possibility of establishing high-risk methylation biomarkers based on genome-wide bisulfite sequencing, which allows analysis of the methylome at the single CpG level, should be addressed in the future. These future studies might also benefit from focusing on more homogeneous high-risk patient groups, for example by only studying *MYCN* amplified or non-amplified samples, as the heterogeneity within our high-risk cohorts might also have counteracted the possibility of establishing high-risk DNA methylation biomarkers [230].

4.3 Sharing is caring

Along with the 87 tumors in our prognostic biomarker project, also 15 stage 4S NB tumors were profiled by MBD sequencing. The MBD sequencing data of these 102 primary NB tumors are unique in the NB research field, as it represents the first sample cohort in which the full tumor heterogeneity is being assessed by genome-wide methylation analysis using NGS. Given the rare occurrence of NB, we are therefore convinced that this data set is a valuable resource of methylation information and that other researchers must be able to use it. Although there is a growing international agreement on the need of sharing research data, especially for rare diseases, open science is not the standard [242, 243]. Next to ethical and legal frameworks hindering open science, also researchers themselves are sometimes not that keen on sharing, as they fear that the valuable work they have put into their sample collection and data generation will not be recognized. As a consequence, researchers often hold on to data until they have extracted every last possible publication from it. Fortunately, things are changing and particularly scientific publishing is at a turning point [242, 243]. This is evidenced by

the growing number of scientific journals specifically focusing on data sharing, for example Matters (www.sciencematters.io), Genomics Data (www.journals.elsevier.com/genomics-data) and Scientific Data (www.nature.com/sdata). To make our data more discoverable, interpretable and reusable for other researchers (inexperienced with MBD sequencing), the MBD sequencing data were published as a data descriptor in Scientific Data (Spaper 3; [110]; Figure 14). In this data descriptor, we provide details on sample annotation (including accession IDs to matching expression and aCGH data), methodological approach and bioinformatics analyses, as well as easy access to the (analyzed) MBD sequencing data and analysis tools and scripts [110]. Additionally, this data descriptor is accompanied by an ISA-Tab metadata file, providing a machine readable overview of the study, which allows exploration via the ISA-explorer tool (<http://scientificdata.isa-explorer.org>). Shortly upon publication of our data descriptor we were contacted by another research group willing to make use of the MBD sequencing data to analyze the methylation status of a certain gene [244]. This request clearly illustrates the opportunities data sharing has to offer and why every researcher should embrace it.

4.4 The stage 4S DNA methylome sheds light on mechanisms contributing to spontaneous regression

To further exemplify the possibilities of reusing our MBD sequencing data, we characterized the DNA methylome of stage 4S NB (§1.2.1.3). This NB tumor stage is a very intriguing clinical presentation; despite patients present with metastatic disease, they usually have a favorable prognosis, often caused by spontaneous tumor regression. Previous studies have shown that most stage 4S NB tumors are *MYCN* non-amplified and hyperdiploid (near-triploid with whole chromosome gains), and demonstrate characteristic mRNA and protein expression patterns [245–248]. For example, Bénard et al. have developed a stage 4S NB classifier by comparing mRNA expression levels of stage 4S and stage 4 NB tumors [246], and Lavarino et al. have shown that distinct gene expression profiles correlate with distinct genomic abnormalities between these two NB tumor stages [247]. Based on these characterization studies, several mechanisms have been proposed to explain the regression phenotype of stage 4S NB (summarized in [245]), but the precise molecular determinants underlying this phenomenon remain to be elucidated. This is partly caused by the fact that recent genome-wide characterization studies in NB have combined patients with stage 4S or other low-risk tumors with favorable prognosis (stage 1/2) into one group for comparison to high-risk patients with unfavorable prognosis (stage 4), hindering the opportunity to gain new insights into the stage 4S NB biology. By comparing MBD sequencing data of *MYCN* non-amplified stage 4S tumors to data of *MYCN* non-amplified stage 1/2 and stage 4 tumors, we described for the first time stage 4S-specific hyper- and hypomethylated promoters, which were characterized for genomic location and function (Spaper 4; [231]; Figure 14). As such, we found that stage 4S tumors show characteristic hypermethylation of specific subtelomeric (§1.1.2.1) promoters and based on the literature we suggested that subtelomeric DNA methylation might represent an additional regulatory mechanism by which telomeres are shortened and spontaneous regression is induced in stage 4S NB. One of the first studies describing a link between subtelomeric DNA methylation and telomere length homeostasis is published by Gonzalo et al. [249]. Using mice embryonic stem cells, they showed that *Dnmt1* or *Dnmt3a* and *Dnmt3b* knockdown leads to a global decrease in subtelomeric DNA methylation levels, and demonstrated that dramatic telomere elongation is present in these DNMT deficient cells. Furthermore, they prove that loss of DNA methylation leads to increased telomeric sister-chromatid

exchange events at telomeres, possibly resulting in telomere lengthening. As such, subtelomeric DNA methylation can be considered a negative regulator of telomere recombination and telomere length [249]. Later on, this has been confirmed by Vera et al. in human cancer cells [250]. Nevertheless, the effect of subtelomeric DNA methylation on telomere length might also directly result from altered expression of other telomere length regulators located in subtelomeres. For example, in humans, the TERT and TERRA promoters are located in subtelomeric regions, more specifically on chr5q and chr20q [251], respectively. TERT is hypermethylated in stage 4S NB compared to stage 4 NB, and this might lead to downregulation of telomerase levels, contributing to telomere shortening in stage 4S NB. Although no differential methylation of the TERRA promoter between the different NB tumor stages could be detected in our study (data not shown), this telomere length regulator should definitely be further analyzed in NB, as it has been shown that TERRA plays a crucial role in alternative lengthening of telomeres (ALT) [252], a process which is also described in NB [253]. Additionally, we also found differential methylation and expression of genes involved in important oncogenic pathways, neural crest development and differentiation, and epigenetic processes. One of these genes is *SLC9A5*, modulating the cellular distribution of TrkA (NTRK1) and nerve growth factor (NGF) signaling [254]. Importantly, TrkA expression and NGF signaling have been suggested to play an important role in NB spontaneous regression [245]. As such, hypermethylation and lower expression levels of *SLC9A5* in stage 4S NB might also contribute to the stage 4S regression phenotype. Importantly, it should be stressed that this study is the first to describe the stage 4S DNA methylome and thus that independent validation of stage 4S-specific differential methylation (of individual genes) is needed. Nevertheless, exploring the DNA methylome of stage 4S tumors has opened new avenues to further unravel its intriguing biology.

4.5 Latent mysteries in neuroblastoma epigenetics may decipher the clinical enigma

In contrast to the impressive amount of knowledge that has been acquired on the NB genome and transcriptome during the last decades, profound DNA methylation studies in NB have lagged behind. Nevertheless, recent findings have triggered the NB research field to dig deeper into this relatively unexplored discipline.

Our project mainly focused on the establishment of prognostic tumor DNA methylation biomarkers. However, applications of DNA methylation biomarkers in cancer management are versatile (§1.2.3) and these should definitely be further explored in the context of NB. For example, the possibilities of identifying DNA methylation biomarkers for minimal residual disease monitoring should be further examined and methylation detection methodologies should be further optimized to allow non-invasive biomarker measurement in blood. It has previously been shown that methylation of *RASSF1A* in bone marrow samples can be used for minimal residual disease testing in NB [255] and methylation of this gene was detected in pretreatment serum of NB patients [256]. Therefore, using the recently described MethyLight digital droplet PCR technology [68], it is worthwhile to test whether serum *RASSF1A* methylation can be used for minimal residual disease monitoring, as well as to investigate whether our identified prognostic DNA methylation biomarkers can also be detected in the blood of NB patients.

Additionally, the link between *MYCN* and DNA methylation needs to be further clarified. Literature suggests an important function of *MYCN* in regulating the NB epigenome (and vice versa), as it

directly interacts with important epigenetic factors, such as DNA methyltransferases DNMT1 and DNMT3A and the MBD protein MeCP2, and colocalizes with them at certain gene promoters, of which some are known to be hypermethylated in NB, for example *RASSF1A* [257, 258]. Additionally, it has been shown that *BM11*, encoding a member of the Polycomb group family that effectuates chromatin modifications, is a direct target of MYCN [259, 260]. Furthermore, we and others have shown that MYCN non-amplified and amplified NB cell lines and primary tumors show differential methylation patterns [229, 240, 261]. For example, in our pilot study, we have demonstrated that the *HIST1H3C* and *ACSS3* promoter is specifically hypermethylated in MYCN amplified NB [229]. Nevertheless, more in-depth analyses are needed to truly characterize the complex interplay between MYCN and the different epigenetic actors.

Other epigenetic players frequently deregulated in cancer but unexplored in NB are the 5-hydroxymethylcytosine base (and its oxidation products 5-formylcytosine and 5-carboxylcytosine) and the TET enzyme family (§1.1.1.1) [9]. 5-hydroxymethylcytosine is present in high levels in neural cells and therefore might have a crucial role in the pathogenesis of NB tumors, which originate from precursor cells of the sympathetic nervous system. So far, it has only been shown that hypoxia in NB cells results in transcriptional activation of *TET1*, leading to an accumulation of 5-hydroxymethylcytosine density at hypoxia-responsive genes [262], and that TET proteins negatively regulate neural differentiation in NB cell lines [263], but more exhaustive studies are currently lacking. Clearly, possible mechanisms deregulating the levels of the oxidation products of 5-methylcytosine in NB should be further examined.

Finally, given that epigenetic processes are highly influenced by environmental factors, it needs to be determined whether DNA methylation differences between primary tumors, metastases and relapse tumors can be detected. An initial study of 16 matching primary and relapse NB tumors showed that primary tumor DNA methylation patterns are globally preserved in relapse tumors, but differential methylation analyses were not assessed [264]. Nevertheless, these analyses might unravel new aspects of NB metastasis and tumor evolution, and might offer new possibilities for treatment [121].

4.6 Conclusions

By exploring the NB DNA methylome using MBD sequencing and by validating site-specific DNA methylation using MSP, we successfully established new prognostic DNA methylation biomarkers for NB. Nevertheless, the quest for prognostic NB markers did not yet come to an end, and especially in the high-risk patient group, the need for additional markers persists. To speed up NB (biomarker) research, sharing data should become the standard. To contribute to the policy of open science, the MBD sequencing data were made available via a data descriptor and its reusability was demonstrated by portraying the stage 4S NB DNA methylome. As such, we discovered new biological mechanisms possibly contributing to the NB pathology. Further exploration of the NB DNA methylome and other epigenetic factors will unravel new aspects of NB biology and will open new horizons in NB biomarker research.

REFERENCES

- [1] Allis CD et al. (2007). Overview and concepts. In Allis CD et al. (Ed.). *Epigenetics* (pp. 23-61). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- [2] Jurkowska RZ et al. (2011). Structure and function of mammalian DNA methyltransferases. *ChemBioChem*; 12(2):206-222.
- [3] Illingworth RS et al. (2009). CpG islands: a rough guide. *FEBS Letters*; 583(11):1713-1720.
- [4] Smallwood SA et al. (2012). De novo DNA methylation: a germ cell perspective. *Trends in Genetics*; 28(1):33-42.
- [5] Wu SC et al. (2010). Active DNA demethylation: many roads lead to Rome. *Nature Reviews. Molecular Cell Biology*; 11(9):607-620.
- [6] Breiling A et al. (2015). Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics and Chromatin*; 8:24.
- [7] Zhang L et al. (2012). Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nature Chemical Biology*; 8(4):328-330.
- [8] He YF et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*; 333(6047):1303-1307.
- [9] Kroeze LI et al. (2015). 5-Hydroxymethylcytosine: an epigenetic mark frequently deregulated in cancer. *Biochimica et Biophysica Acta*; 1855(2):144-154.
- [10] Guo JU et al. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*; 145(3):423-434.
- [11] Branco MR et al. (2011). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nature Reviews. Genetics*; 13(1):7-13.
- [12] Iurlaro M et al. (2016). In vivo genome-wide profiling reveals a tissue-specific role for 5-formylcytosine. *Genome Biology*; 17(1):141.
- [13] Raiber EA et al. (2015). 5-Formylcytosine alters the structure of the DNA double helix. *Nature Structural and Molecular Biology*; 22(1):44-49.
- [14] Bachman M et al. (2015). 5-Formylcytosine can be a stable DNA modification in mammals. *Nature Chemical Biology*; 11(8):555-557.
- [15] Wang L et al. (2015). Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature*; 523(7562):621-625.
- [16] Delatte B et al. (2016). RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science*; 351(6270):282-285.
- [17] Hobert O (2008). Gene regulation by transcription factors and microRNAs. *Science*; 319(5871):1785-1786.
- [18] Jones PA (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews. Genetics*; 13(7):484-492.
- [19] Baylin SB et al. (2011). A decade of exploring the cancer epigenome - biological and translational implications. *Nature Reviews. Cancer*; 11(10):726-734.
- [20] Satterlee JS et al. (2010). Tackling the epigenome: challenges and opportunities for collaboration. *Nature Biotechnology*; 28(10):1039-1044.
- [21] Bae JB (2013). Perspectives of international human epigenome consortium. *Genomics and Informatics*; 11(1):7-14.
- [22] Blasco MA (2007). The epigenetic regulation of mammalian telomeres. *Nature Reviews. Genetics*; 8(4):299-309.
- [23] Slotkin RK et al. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews. Genetics*; 8(4):272-285.
- [24] Friedli M et al. (2015). The developmental control of transposable elements and the evolution of higher species. *Annual Review of Cell and Developmental Biology*; 31:429-451.
- [25] Chaligné R et al. (2014). X-chromosome inactivation in development and cancer. *FEBS Letters*; 588(15):2514-2522.
- [26] Peters J (2014). The role of genomic imprinting in biology and disease: an expanding view. *Nature Reviews. Genetics*; 15(8):517-530.
- [27] Portela A et al. (2010). Epigenetic modifications and human disease. *Nature Biotechnology*; 28(10):1057-

references

- 1068.
- [28] Rodríguez-Paredes M et al. (2011). Cancer epigenetics reaches mainstream oncology. *Nature Medicine*; 17(3):330-339.
- [29] Berdasco M et al. (2010). Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Developmental Cell*; 19(5):698-711.
- [30] You JS et al. (2012). Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*; 22(1):9-20.
- [31] Esteller M (2008). Epigenetics in cancer. *New England Journal of Medicine*; 358(11):1148-1159.
- [32] Liu J et al. (2010). Folate and one-carbon metabolism and its impact on aberrant DNA methylation in cancer. In Herceg Z et al. (Ed.). *Epigenetics and cancer, part B* (p 98). Amsterdam: Elsevier.
- [33] Laird PW (2010). Principles and challenges of genome-wide DNA methylation analysis. *Nature Reviews. Genetics*; 11(3):191-203.
- [34] Nagarajan RP et al. (2013). Methods for cancer epigenome analysis. *Advances in Experimental Medicine and Biology*; 754:313-338.
- [35] Hatada I et al. (1991). A genomic scanning method for higher organisms using restriction sites as landmarks. *Proceedings of the National Academy of Sciences of the USA*; 88(21):9523-9527.
- [36] Gonzalgo ML et al. (1997). Identification and characterization of differentially methylated regions of genomic DNA by methylation-sensitive arbitrarily primed PCR. *Cancer Research*; 57(4):594-599.
- [37] Huang TH et al. (1997). Identification of DNA methylation markers for human breast carcinomas using the methylation-sensitive restriction fingerprinting technique. *Cancer Research*; 57(6):1030-1034.
- [38] Toyota M et al. (1999). Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Research*; 59(10):2307-2312.
- [39] Frigola J et al. (2002). Methylome profiling of cancer cells by amplification of inter-methylated sites (AIMS). *Nucleic Acids Research*; 30(7):e28.
- [40] Nygren AO et al. (2005). Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucleic Acids Research*; 33(14):e128.
- [41] Huang TH et al. (1999). Methylation profiling of CpG islands in human breast cancer cells. *Human Molecular Genetics*; 8(3):459-470.
- [42] Nouzova M et al. (2004). Epigenomic changes during leukemia cell differentiation: analysis of histone acetylation and cytosine methylation using CpG island microarrays. *Journal of Pharmacology and Experimental Therapeutics*; 311(3):968-981.
- [43] Lippman Z et al. (2005). Profiling DNA methylation patterns using genomic tiling microarrays. *Nature Methods*; 2(3):219-224.
- [44] Ibrahim AE et al. (2006). MMASS: an optimized array-based method for assessing CpG island methylation. *Nucleic Acids Research*; 34(20):e136.
- [45] Yuan E et al. (2006). A single nucleotide polymorphism chip-based method for combined genetic and epigenetic profiling: validation in decitabine therapy and tumor/normal comparisons. *Cancer Research*; 66(7):3443-3451.
- [46] Estécio MR et al. (2007). High-throughput methylation profiling by MCA coupled to CpG island microarray. *Genome Research*; 17(10):1529-1536.
- [47] Irizarry RA et al. (2008). Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Research*; 18(5):780-790.
- [48] Khulan B et al. (2006). Comparative isoschizomer profiling of cytosine methylation: The HELP assay. *Genome Research*; 16(8):1046-1055.
- [49] Oda M et al. (2009). High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. *Nucleic Acids Research*; 37(12):3829-3839.
- [50] Hu M et al. (2006). Methylation-specific digital karyotyping. *Nature Protocols*; 1(3):1621-1636.
- [51] Brunner AL et al. (2009). Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Research*; 19(6):1044-1056.
- [52] Ball MP et al. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology*; 27(4):361-368.
- [53] Edwards JR et al. (2010). Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. *Genome Research*; 20(7):972-980.
- [54] Maunakea AK et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative

references

- promoters. *Nature*; 466(7303):253-257.
- [55] Xiong Z et al. (1997). COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Research*; 25(12):2532-2534.
 - [56] Meissner A et al. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Research*; 33(18):5868-5877.
 - [57] Guo H et al. (2013). Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Research*; 23(12):2126-2135.
 - [58] Neri F et al. (2016). Methylation-assisted bisulfite sequencing to simultaneously map 5fC and 5caC on a genome-wide scale for DNA demethylation analysis. *Nature Protocols*; 11(7):1191-1205.
 - [59] Gebhard C et al. (2006). Rapid and sensitive detection of CpG-methylation using methyl-binding (MB)-PCR. *Nucleic Acids Research*; 34(11):e82.
 - [60] Herman JG et al. (1996). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proceedings of the National Academy of Sciences of the USA*; 93(18):9821-9826.
 - [61] Aggerholm A et al. (1999). Extensive intra- and interindividual heterogeneity of p15INK4B methylation in acute myeloid leukemia. *Cancer Research*; 59(2):436-441.
 - [62] Eads CA et al. (2000). MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Research*; 28(8):e32.
 - [63] Worm J et al. (2001). In-tube DNA methylation profiling by fluorescence melting curve analysis. *Clinical Chemistry*; 47(7):1183-1189.
 - [64] Gonzalgo ML et al. (2002). Quantitative methylation analysis using methylation-sensitive single-nucleotide primer extension (Ms-SNuPE). *Methods*; 27(2):128-133.
 - [65] Dobrovic A et al. (2002). Screening for and analysis of methylation differences using methylation-sensitive single-strand conformation analysis. *Methods*; 27(2):134-138.
 - [66] Cottrell SE et al. (2004). A real-time PCR assay for DNA-methylation using methylation-specific blockers. *Nucleic Acids Research*; 32(1):e10.
 - [67] Wojdacz TK et al. (2007). Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Research*; 35(6):e41.
 - [68] Yu M et al. (2015). MethyLight droplet digital PCR for detection and absolute quantification of infrequently methylated alleles. *Epigenetics*; 10(9):803-809.
 - [69] Ekram MB et al. (2014). High-throughput targeted repeat element bisulfite sequencing (HT-TREBS): genome-wide DNA methylation analysis of IAP LTR retrotransposon. *PLoS One*; 9(7):e101683.
 - [70] Bernstein DL et al. (2015). The BisPCR(2) method for targeted bisulfite sequencing. *Epigenetics and Chromatin*; 8:27.
 - [71] Gitan RS et al. (2002). Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. *Genome Research*; 12(1):158-164.
 - [72] Bibikova M et al. (2006). High-throughput DNA methylation profiling using universal bead arrays. *Genome Research*; 16(3):383-393.
 - [73] Bibikova M et al. (2009). Genome-wide DNA methylation profiling using Infinium® assay. *Epigenomics*; 1(1):177-200.
 - [74] Bibikova M et al. (2011). High density DNA methylation array with single CpG site resolution. *Genomics*; 98(4):288-295.
 - [75] Nazor KL et al. (2014). Application of a low cost array-based technique - TAB-Array - for quantifying and mapping both 5mC and 5hmC at single base resolution in human pluripotent stem cells. *Genomics*; 104(5):358-367.
 - [76] Hong KM et al. (2005). Semiautomatic detection of DNA methylation at CpG islands. *BioTechniques*; 38(3):354, 356, 358.
 - [77] Cokus SJ et al. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*; 452(7184):215-219.
 - [78] Smallwood SA et al. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nature Methods*; 11(8):817-820.
 - [79] Farlik M et al. (2015). Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Reports*; 10(8):1386-1397.
 - [80] Lister R et al. (2008). Highly integrated single-base resolution maps of the epigenome in arabidopsis. *Cell*; 133(3):523-526.
 - [81] Deng J et al. (2009). Targeted bisulfite sequencing reveals changes in DNA methylation associated with

references

- nuclear reprogramming. *Nature Biotechnology*; 27(4):353-360.
- [82] Wang J et al. (2011). High resolution profiling of human exon methylation by liquid hybridization capture-based bisulfite sequencing. *BMC Genomics*; 12:597.
 - [83] Hajj NE et al. (2015). Limiting dilution bisulfite pyrosequencing®: a method for methylation analysis of individual DNA molecules in a single or a few cells. *Methods in Molecular Biology*; 1315:221-239.
 - [84] Weber M et al. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genetics*; 37(8):853-862.
 - [85] Zhang X et al. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell*; 126(6):1189-1201.
 - [86] Gebhard C et al. (2006). Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. *Cancer Research*; 66(12):6118-6128.
 - [87] Rauch TA et al. (2010). DNA methylation profiling using the methylated-CpG island recovery assay (MIRA). *Methods*; 52(3):213-217.
 - [88] Down TA et al. (2008). A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nature Biotechnology*; 26(7):779-785.
 - [89] Brinkman AB et al. (2010). Whole-genome DNA methylation profiling using MethylCap-seq. *Methods*; 52(3):232-236.
 - [90] Serre D et al. (2010). MBD-isolated genome sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Research*; 38(2):391-399.
 - [91] Hayatsu H (2008). Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis - a personal account. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences*; 84(8):321-330.
 - [92] Oakeley E (1999). DNA methylation analysis a review of current methodologies. *Pharmacology and Therapeutics*; 84(3):389-400.
 - [93] Jin SG et al. (2010). Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Research*; 38(11):e125.
 - [94] Dedeurwaerder S et al. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics*; 3(6):771-784.
 - [95] Bock et al. (2010). Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nature Biotechnology*; 28(10):1106-1114.
 - [96] Stirzaker C et al. (2014). Mining cancer methylomes: prospects and challenges. *Trends in Genetics*; 30(2):75-84.
 - [97] Gu H et al. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols*; 6(4):468-481.
 - [98] Wang J et al. (2013). Double restriction-enzyme digestion improves the coverage and accuracy of genome-wide CpG methylation profiling by reduced representation bisulfite sequencing. *BMC Genomics*; 14:11.
 - [99] Meissner A (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Research*; 33(18):5868-5877.
 - [100] Lister R et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*; 462(7271):315-322.
 - [101] Hansen KD et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nature Genetics*; 43(8):768-775.
 - [102] Berman BP et al. (2011). Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nature Genetics*; 44(1):40-46.
 - [103] Krueger F et al. (2012). DNA methylome analysis using short bisulfite sequencing data. *Nature Methods*; 9(2):145-151.
 - [104] Li H et al. (2010). A survey of sequence alignment algorithms for next-generation sequencing. *Briefings in Bioinformatics*; 11(5):473-483.
 - [105] Wilbanks EG et al. (2010). Evaluation of algorithm performance in ChIP-seq peak detection. *PLoS One*; 5(7):e11471.
 - [106] Minoche AE et al. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biology*; 12(11):R112.
 - [107] Su J et al. (2012). Advances in bioinformatics tools for high-throughput sequencing data of DNA methylation. *Hereditary Genetics*; 1:107.

references

- [108] Langmead B et al. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*; 9(4):357-359.
- [109] Zhang Y et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*; 9(9):R137.
- [110] Decock A et al. (2016). DNA methylation profiling of primary neuroblastoma tumors using methyl-CpG-binding domain sequencing. *Scientific Data*; 3:160004.
- [111] Anders S et al. (2010). Differential expression analysis for sequence count data. *Genome Biology*; 11(10):R106.
- [112] Benjamini Y et al. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*; 57(1):289-300.
- [113] How Kit A et al. (2012). DNA methylation based biomarkers: practical considerations and applications. *Biochimie*; 94(11):2314-2337.
- [114] Palmisano WA et al. (2000). Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer Research*; 60(21):5954-5958.
- [115] Scesnaite A et al. (2012). Similar DNA methylation pattern in lung tumours from smokers and never-smokers with second-hand tobacco smoke exposure. *Mutagenesis*; 27(4):423-429.
- [116] Laird PW (2003). The power and the promise of DNA methylation markers. *Nature*; 3(4):253-266.
- [117] Shen L et al. (2007). Association between DNA methylation and shortened survival in patients with advanced colorectal cancer treated with 5-fluorouracil based chemotherapy. *Clinical Cancer Research*; 13(20):6093-6098.
- [118] Widschwendter M et al. (2004). Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. *Cancer Research*; 64(11):3807-3813.
- [119] Brock MV et al. (2003). Prognostic importance of promoter hypermethylation of multiple genes in esophageal adenocarcinoma. *Clinical Cancer Research*; 9(8):2912-2919.
- [120] Noehammer C et al. (2014). Strategies for validation and testing of DNA methylation biomarkers. *Epigenomics*; 6(6):603-622.
- [121] Nebbioso A et al. (2012). Trials with “epigenetic” drugs: an update. *Molecular Oncology*; 6(6):657-682.
- [122] Juo YY et al. (2015). Epigenetic therapy for solid tumors: from bench science to clinical trials. *Epigenomics*; 7(2):215-235.
- [123] Stresemann C et al. (2008). Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *International Journal of Cancer*; 123(1):8-13.
- [124] Juergens RA et al. (2011). Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. *Cancer Discovery*; 1(7):598-607.
- [125] Fu S et al. (2011). Phase 1b-2a study to reverse platinum resistance through use of a hypomethylating agent, azacitidine, in patients with platinum-resistant or platinum-refractory epithelial ovarian cancer. *Cancer*; 117(8):1661-1669.
- [126] Yoo CB et al. (2004). Zebularine: a new drug for epigenetic therapy. *Biochemical Society Transactions*; 32(6):910-912.
- [127] Lee BH et al. (2005). Procainamide is a specific inhibitor of DNA methyltransferase 1. *Journal of Biological Chemistry*; 280(49):40749-40756.
- [128] Candelaria M et al. (2007). A phase II study of epigenetic therapy with hydralazine and magnesium valproate to overcome chemotherapy resistance in refractory solid tumors. *Annals of Oncology*; 18(9):1529-1538.
- [129] Lin YS et al. (2011). Identification of novel DNA methylation inhibitors via a two-component reporter gene system. *Journal of Biomedical Science*; 18:3.
- [130] Brueckner B (2005). Epigenetic reactivation of tumor suppressor genes by a novel small-molecule inhibitor of human DNA methyltransferases. *Cancer Research*; 65(14):6305-6311.
- [131] Datta J et al. (2009). A new class of quinoline-based DNA hypomethylating agents reactivates tumor suppressor genes by blocking DNA methyltransferase 1 activity and inducing its degradation. *Cancer Research*; 69(10):4277-4285.
- [132] Stewart DJ et al. (2003). A phase I pharmacokinetic and pharmacodynamic study of the DNA methyltransferase 1 inhibitor MG98 administered twice weekly. *Annals of Oncology*; 14(5):766-774.
- [133] Plummer R et al. (2009). Phase I study of MG98, an oligonucleotide antisense inhibitor of human DNA methyltransferase 1, given as a 7-day infusion in patients with advanced solid tumors. *Clinical Cancer Research*; 15(9):3177-3183.
- [134] Baud MG et al. (2012). Defining the mechanism of action and enzymatic selectivity of psammaplin A against its epigenetic targets. *Journal of Medical Chemistry*; 55(4):1731-1750.

references

- [135] Nandakumar V et al. (2011). (-)-Epigallocatechin-3-gallate reactivates silenced tumor suppressor genes, Cip1/p21 and p16INK4a, by reducing DNA methylation and increasing histones acetylation in human skin cancer cells. *Carcinogenesis*; 32(4):537-544.
- [136] Shilpi et al. (2015). Mechanisms of DNA methyltransferase-inhibitor interactions: Procyanidin B2 shows new promise for therapeutic intervention of cancer. *Chemico-biological Interactions*; 233:122-138.
- [137] Berthold F et al. (2005). Clinical presentation. In Cheung NKV et al. (Ed.). *Neuroblastoma* (pp. 63-64). Heidelberg: Springer-Verlag.
- [138] Pepper W (1901). A study of congenital sarcoma of the liver and suprarenal with report of a case. *American Journal of the Medical Sciences*; 121(3):287-299.
- [139] Hutchison R (1907). On suprarenal sarcoma in children with metastases in the skull. *Quarterly Journal of Medicine*; 1:33-38.
- [140] Wright J (1910). Neurocytoma or neuroblastoma, a kind of tumor not generally recognized. *Journal of Experimental Medicine*; 12:556-561.
- [141] Rothenberg AB et al. (2009). Neuroblastoma-remembering the three physicians who described it a century ago: James Homer Wright, William Pepper, and Robert Hutchison. *Pediatric Radiology*; 39(2):155-160.
- [142] Bronner-Fraser M et al. (2008). Neurogenesis and migration. In Squire L et al. (Ed.). *Fundamental neuroscience* (pp. 352-361). Amsterdam: Elsevier.
- [143] Monclair T et al. (2009). The International Neuroblastoma Risk Group (INRG) staging system: an INRG Task Force report. *Journal of Clinical Oncology*; 27(2):298-303.
- [144] Marieb EN et al. (2005). Basic embryology. In Murray MA et al. (Ed.). *Human anatomy* (pp. 51-68). San Francisco: Pearson Education.
- [145] Cotterill SJ et al. (2000). Clinical prognostic factors in 1277 patients with neuroblastoma: results of The European Neuroblastoma Study Group "Survey" 1982-1992. *European Journal of Cancer*; 36(7):901-908.
- [146] Vo KT et al. (2014). Clinical, biologic, and prognostic differences on the basis of primary tumor site in neuroblastoma: a report from the International Neuroblastoma Risk Group project. *Journal of Clinical Oncology*; 32(28):3169-3176.
- [147] Dubois SG et al. (1999). Metastatic sites in stage IV and IVS neuroblastoma correlate with age, tumor biology, and survival. *Journal of Pediatric Hematology/Oncology*; 21(3):181-189.
- [148] Strenger V et al. (2007). Diagnostic and prognostic impact of urinary catecholamines in neuroblastoma patients. *Pediatric Blood and Cancer*; 48(5):504-509.
- [149] Cangemi G et al. (2012). Prognostic value of ferritin, neuron-specific enolase, lactate dehydrogenase, and urinary and plasmatic catecholamine metabolites in children with neuroblastoma. *Oncotargets and Therapy*; 5:417-423.
- [150] Massaron S et al. (1998). Neuron-specific enolase evaluation in patients with neuroblastoma. *Tumour Biology*; 19(4):261-268.
- [151] Meany HJ et al. (2014). Significance of clinical and biologic features in stage 3 neuroblastoma: a report from the International Neuroblastoma Risk Group project. *Pediatric Blood and Cancer*; 61(11):1932-1939.
- [152] Hsiao RJ et al. (1990). Chromogranin A in children with neuroblastoma. Serum concentration parallels disease stage and predicts survival. *Journal of Clinical Investigation*; 85(5):1555-1559.
- [153] Hayashi Y et al. (1991). Significance of plasma neuropeptide Y (NPY) in diagnosis and prognosis of neuroblastoma. *Progress in Clinical and Biological Research*; 366:359-365.
- [154] Cohn SL et al. (2009). The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. *Journal of Clinical Oncology*; 27(2):289-297.
- [155] Howlader N et al. (Ed.). *SEER Cancer Statistics Review, 1975-2010*, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.
- [156] Ora I et al. (2011). Progress in treatment and risk stratification of neuroblastoma: impact on future clinical and basic research. *Seminars in Cancer Biology*; 21(4):217-228.
- [157] Breslow N et al. (1971). Statistical estimation of prognosis for children with neuroblastoma. *Cancer Research*; 31(12):2098-2103.
- [158] London WB et al. (2005). Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's Oncology Group. *Journal of Clinical Oncology*; 23(27):6459-6465.
- [159] Schmidt ML et al. (2005). Favorable prognosis for patients 12 to 18 months of age with stage 4

references

- nonamplified MYCN neuroblastoma: a Children's Cancer Group study. *Journal of Clinical Oncology*; 23(27):6474-6480.
- [160] Rich JT et al. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology: Head and Neck Surgery*; 143(3):331-336.
 - [161] Shimada H et al. (1999). The International Neuroblastoma Pathology Classification (the Shimada System). *Cancer*; 86(2):364-372.
 - [162] Beckwith JB et al. (1968). Observations on the histopathology of neuroblastomas. *Journal of Pediatric Surgery*; 3(1):106-110.
 - [163] Mäkinen J (1972). Microscopic patterns as a guide to prognosis of neuroblastoma in childhood. *Cancer*; 29(6):1637-1646.
 - [164] Hughes M et al. (1974). Histologic patterns of neuroblastoma related to prognosis and clinical staging. *Cancer*; 34(5):1706-1711.
 - [165] Shimada H et al. (1984). Histopathologic prognostic factors in neuroblastic tumors: definition of subtypes of ganglioneuroblastoma and an age-linked classification of neuroblastomas. *Journal of the National Cancer Institute*; 73(2):405-416.
 - [166] Shimada H et al. (1999). Terminology and morphologic criteria of neuroblastic tumors: recommendations by the International Neuroblastoma Pathology Committee. *Cancer*; 86(2):349-363.
 - [167] Joshi VV et al. (1992). Age-linked prognostic categorization based on a new histologic grading system of neuroblastomas. A clinicopathologic study of 211 cases from the Pediatric Oncology Group. *Cancer*; 69(8):2197-2211.
 - [168] Joshi VV et al. (1996). Modified histologic grading of neuroblastomas by replacement of mitotic rate with mitosis karyorrhexis index. A clinicopathologic study of 223 cases from the Pediatric Oncology Group. *Cancer*; 77(8):1582-1588.
 - [169] Shimada H et al. (2005). Pathology of peripheral neuroblastic tumors. In Cheung NKV et al. (Ed.). *Neuroblastoma* (p 89). Heidelberg: Springer-Verlag.
 - [170] Carlsen NL et al. (1986). Prognostic value of different staging systems in neuroblastomas and completeness of tumour excision. *Archives of Disease in Childhood*; 61(9):832-842.
 - [171] Smith EI et al. (1989). A surgical perspective on the current staging in neuroblastoma - The International Neuroblastoma Staging System proposal. *Journal of Pediatric Surgery*; 24(4):386-390.
 - [172] James DH et al. (1967). Proposed classification of neuroblastoma. *Journal of Pediatrics*; 71(5):764.
 - [173] Pinkel D et al. (1968). Survival of children with neuroblastoma treated with combination chemotherapy. *Journal of Pediatrics*; 73(6):928-931.
 - [174] Green AA et al. (1976). Total-body sequential segmental irradiation and combination chemotherapy for children with disseminated neuroblastoma. *Cancer*; 38(6):2250-2257.
 - [175] Cohen D (1969). Neuroblastoma: factors affecting survival. *Zeitschrift für Kinderchirurgie*; 6:389-395.
 - [176] Thurman W et al. (1969). Current concept in the management of neuroblastoma. In: Anderson Hospital MD (Ed.). *Neoplasia in childhood* (pp. 175-181). Chicago: Year Book Medical Publishers.
 - [177] Wippold FJ 2nd et al. (2006). Neuropathology for the neuroradiologist: rosettes and pseudorosettes. *American Journal of Neuroradiology*; 27(3):488-492.
 - [178] Evans AE et al. (1971). A proposed staging for children with neuroblastoma. Children's cancer study group A. *Cancer*; 27(2):374-378.
 - [179] Evans AE et al. (1976). Factors influencing survival of children with nonmetastatic neuroblastoma. *Cancer*; 38(2):661-666.
 - [180] In Harmer MH (Ed.; 1982). *International Union Against Cancer: TNM classification of malignant tumours* (pp. 135-146). Geneva: International Union Against Cancer.
 - [181] Hartmann O et al. (1983). Neuroblastomas treated at the Gustave-Roussy Institute from 1975 to 1979. 173 cases. *Archive Françaises de Pédiatrie*; 40(1):15-21.
 - [182] Hayes FA et al. (1983). Surgicopathologic staging of neuroblastoma: prognostic significance of regional lymph node metastases. *Journal of Pediatrics*; 102(1):59-62.
 - [183] Nitschke R et al. (1988). Localized neuroblastoma treated by surgery: a Pediatric Oncology Group Study. *Journal of Clinical Oncology*; 6(8):1271-1279.
 - [184] de Bernardi B et al. (1987). Localized neuroblastoma. Surgical and pathologic staging. *Cancer*; 60(5):1066-1072.
 - [185] Brodeur GM et al. (1988). International criteria for diagnosis, staging, and response to treatment in patients with neuroblastoma. *Journal of Clinical Oncology*; 6(12):1874-1881.

references

- [186] Brodeur GM et al. (1993). Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *Journal of Clinical Oncology*; 11(8):1466-1477.
- [187] Bombardieri E et al. (2010). 131I/123I-metaiodobenzylguanidine (mIBG) scintigraphy: procedure guidelines for tumour imaging. *European Journal of Nuclear Medicine and Molecular Imaging*; 37(12):2436-2446.
- [188] Matthay KK et al. (2010). Criteria for evaluation of disease extent by (123)I-metaiodobenzylguanidine scans in neuroblastoma: a report for the International Neuroblastoma Risk Group (INRG) Task Force. *British Journal of Cancer*; 102(9):1319-1326.
- [189] Wong KK et al. (2013). Dynamic bone imaging with 99mTc-labeled diphosphonates and 18F-NaF: mechanisms and applications. *Journal of Nuclear Medicine*; 54(4):590-599.
- [190] Schwab M et al. (2003). Neuroblastoma: biology and molecular and chromosomal pathology. *Lancet Oncology*; 4(8):472-480.
- [191] Schwab M et al. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. *Nature*; 305(5931):245-248.
- [192] Beltran H (2014). The N-myc oncogene: maximizing its targets, regulation, and therapeutic potential. *Molecular Cancer Research*; 12(6):815-822.
- [193] Schwab M et al. (1985). Human N-myc gene contributes to neoplastic transformation of mammalian cells in culture. *Nature*; 316(6024):160-162.
- [194] Grandori C et al. (2000). The Myc/Max/Mad network and the transcriptional control of cell behavior. *Cancer Research*; 16:653-699.
- [195] Hasan MK et al. (2013). ALK is a MYCN target gene and regulates cell migration and invasion in neuroblastoma. *Scientific Reports*; 3:3450.
- [196] Schulte JH et al. (2012). MYCN and ALKF1174L are sufficient to drive neuroblastoma development from neural crest progenitor cells. *Oncogene*; 32(8):1059-1065.
- [197] Ambros PF et al. (2009). International consensus for neuroblastoma molecular diagnostics: report from the International Neuroblastoma Risk Group (INRG) Biology Committee. *British Journal of Cancer*; 100(9):1471-1482.
- [198] Moreau LA et al. (2006). Does MYCN amplification manifested as homogeneously staining regions at diagnosis predict a worse outcome in children with neuroblastoma? A Children's Oncology Group study. *Clinical Cancer Research*; 12(19):5693-5697.
- [199] De Preter K et al. (2002). Quantification of MYCN, DDX1, and NAG gene copy number in neuroblastoma using a real-time quantitative PCR assay. *Modern Pathology*; 15(2):159-166.
- [200] Kallioniemi A et al. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*; 258(5083):818-821.
- [201] Ambros IM et al. (2011). A multilocus technique for risk evaluation of patients with neuroblastoma. *Clinical Cancer Research*; 17(4):792-804.
- [202] Schouten JP et al. (2002). Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Research*; 30(12):e57.
- [203] Brodeur GM et al. (1984). Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science*; 224(4653):1121-1124.
- [204] Seeger RC et al. (1985). Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *New England Journal of Medicine*; 313(18):1111-1116.
- [205] Chan HS et al. (1997). MYCN protein expression as a predictor of neuroblastoma prognosis. *Clinical Cancer Research*; 3(10):1699-1706.
- [206] Bordow SB et al. (1998). Prognostic significance of MYCN oncogene expression in childhood neuroblastoma. *Journal of Clinical Oncology*; 16(10):3286-3294.
- [207] Cohn SL et al. (2000). MYCN expression is not prognostic of adverse outcome in advanced-stage neuroblastoma with nonamplified MYCN. *Journal of Clinical Oncology*; 18(21):3604-3613.
- [208] Fredlund E et al. (2008). High Myc pathway activity and low stage of neuronal differentiation associate with poor outcome in neuroblastoma. *Proceedings of the National Academy of Sciences of the USA*; 105(37):14094-14099.
- [209] Westermann F et al. (2008). Distinct transcriptional MYCN/c-MYC activities are associated with spontaneous regression or malignant progression in neuroblastomas. *Genome Biology*; 9(10):R150.
- [210] George RE et al. (2005). Hyperdiploidy plus nonamplified MYCN confers a favorable prognosis in children 12 to 18 months old with disseminated neuroblastoma: a Pediatric Oncology Group study. *Journal of*

references

- Clinical Oncology; 23(27):6466-6473.
- [211] Vandesompele J et al. (2005). Unequivocal delineation of clinicogenetic subgroups and development of a new model for improved outcome prediction in neuroblastoma. *Journal of Clinical Oncology*; 23(10):2280-2299.
 - [212] Tomioka N et al. (2008). Novel risk stratification of patients with neuroblastoma by genomic signature, which is independent of molecular signature. *Oncogene*; 27(4):441-449.
 - [213] Janoueix-Lerosey I et al. (2009). Overall genomic pattern is a predictor of outcome in neuroblastoma. *Journal of Clinical Oncology*; 27(7):1026-1033.
 - [214] Schleiermacher G et al. (2011). Segmental chromosomal alterations lead to a higher risk of relapse in infants with MYCN-non-amplified localised unresectable/disseminated neuroblastoma (a SIOPEN collaborative study). *British Journal of Cancer*; 105(12):1940-1948.
 - [215] Brodeur GM et al. (1981). Cytogenetic features of human neuroblastomas and cell lines. *Cancer Research*; 41(11):4678-4686.
 - [216] White PS et al. (2001). Detailed molecular analysis of 1p36 in neuroblastoma. *Medical and Pediatric Oncology*; 36(1):37-41.
 - [217] Attiyeh EF et al. (2005). Chromosome 1p and 11q deletions and outcome in neuroblastoma. *New England Journal of Medicine*; 353(21):2243-2253.
 - [218] Plantaz D et al. (2001). Comparative genomic hybridization (CGH) analysis of stage 4 neuroblastoma reveals high frequency of 11q deletion in tumors lacking MYCN amplification. *International Journal of Cancer*; 91(5):680-686.
 - [219] Bown N et al. (1999). Gain of chromosome arm 17q and adverse outcome in patients with neuroblastoma. *New England Journal of Medicine*; 340(25):1954-1961.
 - [220] Nakagawara A (1993). Association between high levels of expression of the TRK gene and favorable outcome in human neuroblastoma. *New England Journal of Medicine*; 328(12):847-854.
 - [221] Oberthuer A et al. (2006). Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. *Journal of Clinical Oncology*; 24(31):5070-5078.
 - [222] Vermeulen J et al. (2009). Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/COG/GPOH study. *Lancet Oncology*; 10(7):663-671.
 - [223] Schulte JH et al. (2010). Accurate prediction of neuroblastoma outcome based on miRNA expression profiles. *International Journal of Cancer*; 127(10):2374-2385.
 - [224] De Preter K et al. (2011). miRNA expression profiling enables risk stratification in archived and fresh neuroblastoma tumor samples. *Clinical Cancer Research*; 17(24):7684-7692.
 - [225] Beiske K et al. (2009). Consensus criteria for sensitive detection of minimal neuroblastoma cells in bone marrow, blood and stem cell preparations by immunocytology and QRT-PCR: recommendations by the International Neuroblastoma Risk Group Task Force. *British Journal of Cancer*; 100(10):1627-1637.
 - [226] Pinto NR et al. (2015). Advances in risk classification and treatment strategies for neuroblastoma. *Journal of Clinical Oncology*; 33(27):3008-3017.
 - [227] De Preter K et al. (2010). Accurate outcome prediction in neuroblastoma across independent data sets using a multigene signature. *Clinical Cancer Research*; 16(5):1532-1541.
 - [228] Decock A et al. (2011). Neuroblastoma epigenetics: from candidate gene approaches to genome-wide screenings. *Epigenetics*; 6(8):962-970.
 - [229] Decock A et al. (2012). Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. *Genome Biology*; 13(10):R95.
 - [230] Decock A et al. (2016). Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncotarget*; 7(2):1960-1972.
 - [231] Decock A et al. (2016). Stage 4S neuroblastoma tumors show a characteristic DNA methylation portrait. Accepted for publication in *Epigenetics*.
 - [232] Schrijver I et al. (2012). Opportunities and challenges associated with clinical diagnostic genome sequencing: a report of the Association for Molecular Pathology. *Journal of Molecular Diagnostics*; 14(6):525-540.
 - [233] Deverka PA et al. (2014). Clinical integration of next generation sequencing: coverage and reimbursement challenges. *Journal of Law, Medicine and Ethics*; 42(1):22-41.
 - [234] Damodaran S et al. (2015). Clinical tumor sequencing: opportunities and challenges for precision cancer medicine. *American Society of Clinical Oncology Educational Book*; e175-182.

references

- [235] Diamandis EP (2010). Cancer biomarkers: can we turn recent failures into success? *Journal of the National Cancer Institute*; 102(19):1462-1467.
- [236] Marchiò C et al. (2011). Revisiting the technical validation of tumour biomarker assays: how to open a Pandora's box. *BMC Medicine*; 9:41.
- [237] Altman DG et al. (2012). Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *PLoS Medicine*; 9(5):e1001216.
- [238] Valentijn LJ et al. (2015). TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nature Genetics*; 47(12):1411-1414.
- [239] Peifer M et al. (2015). Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature*; 526(7575):700-704.
- [240] Yáñez Y et al. (2015). Two independent epigenetic biomarkers predict survival in neuroblastoma. *Clinical Epigenetics*; 7(1):16.
- [241] Banelli B et al. (2013). Clinical potentials of methylator phenotype in stage 4 high-risk neuroblastoma: an open challenge. *PLoS One*; 8(5):e63253.
- [242] Watson M (2015). When will "open science" become simply "science"? *Genome Biology*; 16:101.
- [243] Mascalzoni D et al. (2015). International charter of principles for sharing bio-specimens and data. *European Journal of Human Genetics*; 23(6):721-728.
- [244] Fabian J et al. (2016). MYCN and HDAC5 transcriptionally repress CD9 to trigger an invasion-metastasis cascade in neuroblastoma. Accepted for publication in *Oncotarget*.
- [245] Brodeur GM et al. (2014). Mechanisms of neuroblastoma regression. *Nature Reviews. Clinical Oncology*; 11(12):704-713.
- [246] Bénard J et al. (2008). MYCN-non-amplified metastatic neuroblastoma with good prognosis and spontaneous regression: a molecular portrait of stage 4S. *Molecular Oncology*; 2(3):261-271.
- [247] Lavarino C et al. (2009). Specific gene expression profiles and chromosomal abnormalities are associated with infant disseminated neuroblastoma. *BMC Cancer*; 9:44.
- [248] Yu F et al. (2011). Proteomics-based identification of spontaneous regression-associated proteins in neuroblastoma. *Journal of Pediatric Surgery*; 46(10):1948-1955.
- [249] Gonzalo S et al. (2006). DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nature Cell Biology*; 8(4):416-424.
- [250] Vera E et al. (2008). Epigenetic regulation of telomeres in human cancer. *Oncogene*; 27(54):6817-6833.
- [251] Montero JJ et al. (2016). Telomeric RNAs are essential to maintain telomeres. *Nature Communications*; 7:12534.
- [252] Arora R et al. (2014). RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nature Communications*; 5:5220.
- [253] Onitake Y et al. (2009). Telomere biology in neuroblastoma: telomere binding proteins and alternative strengthening of telomeres. *Journal of Pediatric Surgery*; 44(12):2258-2266.
- [254] Diering GH et al. (2013). Endosomal acidification by Na⁺/H⁺ exchanger NHE5 regulates TrkA cell-surface targeting and NGF-induced PI3K signaling. *Molecular Biology of the Cell*; 24(21):3435-3448.
- [255] Stutterheim J et al. (2012). Methylated RASSF1a is the first specific DNA marker for minimal residual disease testing in neuroblastoma. *Clinical Cancer Research*; 18(3):808-814.
- [256] Misawa A et al. (2009). RASSF1A hypermethylation in pretreatment serum DNA of neuroblastoma patients: a prognostic marker. *British Journal of Cancer*; 100(2):399-404.
- [257] Charlet J et al. (2014). MYCN is recruited to the RASSF1A promoter but is not critical for DNA hypermethylation in neuroblastoma. *Molecular Carcinogenesis*; 53(5):413-420.
- [258] Murphy DM et al. (2011). Co-localization of the oncogenic transcription factor MYCN and the DNA methyl binding protein MeCP2 at genomic sites in neuroblastoma. *PLoS One*; 6(6):e21436.
- [259] Ochiai H et al. (2010). Bmi1 is a MYCN target gene that regulates tumorigenesis through repression of KIF1Bbeta and TSLC1 in neuroblastoma. *Oncogene*; 29(18):2681-2690.
- [260] Huang R et al. (2011). MYCN and MYC regulate tumor proliferation and tumorigenesis directly through BMI1 in human neuroblastomas. *FASEB Journal*; 25(12):4138-4149.
- [261] Gómez S et al. (2015). DNA methylation fingerprint of neuroblastoma reveals new biological and clinical insights. *Epigenomics*; 7(7):1137-1153.
- [262] Mariani CJ et al. (2014). TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma. *Cell Reports*; 7(5):1343-1352.
- [263] Gao J et al. (2016). Non-catalytic roles for TET1 protein negatively regulating neuronal differentiation

references

- through srGAP3 in neuroblastoma cells. *Protein and Cell*; 7(5):351-361.
- [264] Schramm A et al. (2015). Mutational dynamics between primary and relapse neuroblastomas. *Nature Genetics*; 47(8):872-877.

CURRICULUM VITAE

PERSONAL DETAILS

name Anneleen Decock
address Ottergemsesteenweg 375A, 9000 Ghent, Belgium
telephone +32-479-902039
e-mail anneleen.decock@gmail.com
born on August 14, 1987 in Ronse
nationality Belgian



EDUCATION

2010 - present higher education
PhD fellow in Health Sciences
topic: epigenetic biomarkers for prognosis in children with neuroblastoma
Ghent University, De Pintelaan 185, 9000 Ghent, Belgium

2005 - 2010 higher education
master's degree of Biomedical Sciences - discipline Neurosciences
thesis: the role of inositol 1,4,5-trisphosphate in the spread of apoptosis via gap junctions in C6-glioma
Ghent University, De Pintelaan 185, 9000 Ghent, Belgium

1999 - 2005 general secondary education
direction Sciences - Mathematics
KSO Glorieux, Glorieuxlaan 30, 9600 Ronse, Belgium

EXPERIENCE

2010 - present technical training in: cell culturing, western blotting, tissue slicing using microtome, DNA isolation, methylations-specific PCR (MSP; DNA bisulfite treatment and pre-amplification, qPCR on Roche LightCycler480; sizing on Caliper LabChip GX), qPCR, methyl-CpG-binding domain (MBD) sequencing and (chromatin) immunoprecipitation
statistical analysis software: SPSS, SAS PSS and R
Department of Pediatrics and Medical Genetics - Center for Medical Genetics
Faculty of Medicine and Health Sciences, De Pintelaan 185 (2MRB1), 9000 Ghent, Belgium

2009 - 2010 technical training in: cell culturing, SDS-PAGE, DNA electrophoresis, *in situ* electroporation, PCR, fluorescence microscopy (intracellular Ca²⁺ measurements, immuno- and apoptosis stainings), confocal laser scanning microscopy (FRAP), transformation of bacteria, plasmid purification, cell transfections with plasmid DNA

Department of Basic Medical Sciences - Physiology
Faculty of Medicine and Health Sciences, De Pintelaan 185 (Blok B), 9000 Ghent,
Belgium

COURSES AND CERTIFICATES

Power and sample size calculations - Faculty of Medicine and Health Sciences, Ghent University (December 12, 19 and 20, 2013, Ghent, Belgium).

Biomarker training - Fournier-Majoie Foundation (November 29 and December 13, 2013, Brussels, Belgium).

Statistical analyses using SPSS: advanced courses - Faculty of Medicine and Health Sciences, Ghent University (August 30, September 6, 13 and 20, 2012, Ghent, Belgium).

qPCR experiment design and data analysis - Biogazelle (May 23 and 24, 2011, Ghent, Belgium).

Statistical analyses using SPSS: starter courses - Faculty of Medicine and Health Sciences, Ghent University (January 11, 18 and 25, and February 1, 2011, Ghent, Belgium).

Literature mining techniques: tips and tricks - VIB/BITS (December 2, 2010, Leuven, Belgium).

Laboratory animal science - Faculty of Medicine and Health Sciences, Ghent University (September 29 - October 10, 2008, Belgium). The course has been recognized by the Belgian Federal Public Service of Public Health, Food Chain Safety and Environment as being in compliance with Annex VIII (education of persons responsible for directing animal experiments) of the Royal Decree of the 14th of November 1993 covering the protection of experimental animals.

AWARDS AND GRANTS

Best storm session presentation of Oncopoint 2016. Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. March 2, 2016, Ghent, Belgium.

Emmanuel Van der Schueren research grant of VLK. Exploring the neuroblastoma DNA methylome: identifying differences between biologically distinct patient groups and establishing biomarker detection in plasma samples. November 25, 2015, Mechelen, Belgium.

FWO PhD fellowship renewal grant. Epigenetic biomarkers for prognosis in children with neuroblastoma. June 26, 2013, Brussels, Belgium.

Fournier-Majoie Foundation Valorization Award 2013. Knowledge for Growth, May 30, 2013, Ghent, Belgium.

FWO PhD fellowship grant. An epigenetic DNA modification signature as a biomarker for prognosis in children with neuroblastoma. June 22, 2011, Brussels, Belgium.

Joël Vandekerckhove Price, supported by Devgen, for the best master thesis biomedical sciences (academic year 2009-2010). September 28, 2010, Ghent, Belgium.

MENTORSHIP OF BACHELOR AND MASTER STUDENTS

Chloë De Meyere (2016). Promising epigenetic medicines for the treatment of non-small cell lung cancer. Z-line paper, bachelor of Medicine (Ghent University, Faculty of Medicine and Health Sciences). Promotor: Vandesompele J, mentor: Decock A.

Lucas Deseins (2016). Mechanisms of action of SGI-110 on solid tumors. Z-line paper, bachelor of Medicine (Ghent University, Faculty of Medicine and Health Sciences). Promotor: Vandesompele J, mentor: Decock A.

Stefan De Praeter (2015). Inhibition of Polycomb Repressive Complex 1 and 2 in prostate cancer therapy. Thesis submitted to obtain the degree of master of Medicine (Ghent University, Faculty of Medicine and Health Sciences). Promotor: Speleman F, co-promotor: Ongenaert M, mentor: Decock A.

Lise Van den Haute (2014). Prognostic DNA methylation markers for neuroblastoma. Thesis submitted to obtain the degree of FBT (Hogeschool Gent, Faculty of Education, Health and Social Work). Promotor: Vandesompele J, mentor: Decock A.

Sherly Verbauwhede (2013). Epigenetic biomarkers for neuroblastoma. Thesis submitted to obtain the degree of master of Biomedical Sciences (Ghent University, Faculty of Medicine and Health Sciences). Promotor: Ongenaert M, mentor: Decock A.

PUBLICATIONS

Decock A, Ongenaert M, De Wilde B, Brichard B, Noguera R, Speleman F, Vandesompele J. Stage 4S neuroblastoma tumors show a characteristic DNA methylation portrait. Accepted for publication in Epigenetics.

Fabian J, Opitz D, Althoff K, Lodrini M, Hero B, Volland R, Beckers A, De Preter K, Decock A, Patil N, Abba M, Astrahantseff K, Wünschel J, Ercu M, Künkele M, Hu J, Thole T, Schweizer L, Mechttersheimer G, Carter D, Cheung BB, Popanda O, von Deimling A, Koster J, Versteeg R, Schwab M, Marshall GM, Speleman F, Erb U, Zoeller M, Allgayer H, Simon T, Fischer M, Kulozik AE, Eggert A, Witt O, Schulte JH, Deubzer HE. MYCN and HDAC5 transcriptionally repress CD9 to trigger an invasion-metastasis cascade in neuroblastoma. Accepted for publication in Oncotarget.

Decock A, Ongenaert M, Van Criekinge W, Speleman F, Vandesompele J (2016). DNA methylation profiling of primary neuroblastoma tumors using methyl-CpG-binding domain sequencing. Scientific Data 3:160004.

Decock A, Ongenaert M, Cannoodt R, Verniers K, De Wilde B, Laureys G, Van Roy N, Berbegall AP,

Bienertova-Vasku J, Bown N, Clément N, Combaret V, Haber M, Hoyoux C, Murray J, Noguera R, Pierron G, Schleiermacher G, Schulte JH, Stallings RL, Tweddle DA for the Children's Cancer and Leukaemia Group (CCLG), De Preter K, Speleman F, Vandesompele J (2016). Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncotarget* 7(2):1960-1972.

Decock A*, Ongenaert M*, Hoebeeck J, De Preter K, Van Peer G, Van Criekinge W, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F, Vandesompele J (2012). Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. *Genome Biology* 13(10):R95. *contributed equally.

Decock A*, Ongenaert M*, Vandesompele J, Speleman F (2011). Neuroblastoma epigenetics: from candidate gene approaches to genome-wide screenings. *Epigenetics* 6(8):962-970. *contributed equally.

ORAL PRESENTATIONS

Decock A, Ongenaert M, Cannoodt R, Verniers K, De Wilde B, Laureys G, Van Roy N, Berbegall AP, Bienertova-Vasku J, Bown N, Clément N, Combaret V, Haber M, Hoyoux C, Murray J, Noguera R, Pierron G, Schleiermacher G, Schulte JH, Stallings RL, Tweddle DA for the Children's Cancer and Leukaemia Group (CCLG), De Preter K, Speleman F, Vandesompele J. Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncopoint*. March 2, 2016, Ghent, Belgium.

Decock A, Zeka F, Ongenaert M, De Preter K, Speleman K, Vandesompele J. Improved outcome prediction of children with neuroblastoma using an integrated miRNA and DNA methylation signature. 3rd scientific progress meeting Villa Joep!. December 9, 2013, Amsterdam, The Netherlands.

ABSTRACTS AND CONGRESS REPORTS

Decock A, Ongenaert M, Cannoodt R, Verniers K, De Wilde B, Laureys G, Van Roy N, Berbegall AP, Bienertova-Vasku J, Bown N, Clément N, Combaret V, Haber M, Hoyoux C, Murray J, Noguera R, Pierron G, Schleiermacher G, Schulte JH, Stallings RL, Tweddle DA for the Children's Cancer and Leukaemia Group (CCLG), De Preter K, Speleman F, Vandesompele J. Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncopoint*, March 2, 2016, Ghent, Belgium.

Decock A, Ongenaert M, Cannoodt R, Verniers K, De Wilde B, Laureys G, Van Roy N, Berbegall AP, Bienertova-Vasku J, Bown N, Clément N, Combaret V, Haber M, Hoyoux C, Murray J, Noguera R, Pierron G, Schleiermacher G, Schulte JH, Stallings RL, Tweddle DA for the Children's Cancer and Leukaemia Group (CCLG), De Preter K, Speleman F, Vandesompele J. Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. 1st European Cancer Epigenetics Conference, November 11-13, 2015, Maastricht, The Netherlands.

Decock A, Ongenaert M, Berbegall A, Cannoodt R, Noguera R, Murray J, Haber M, Stallings RL, Schulte JH, Combaret V, Schleiermacher G, Clément N, Tweddle D, De Preter K, Speleman F and Vandesompele J. Genome-wide methylation analysis unravels differential DNA methylation patterns between (prognostic) neuroblastoma patient subgroups. Advances in Neuroblastoma Research 2014. May 13-16, 2014, Cologne, Germany.

Beckers A, Van Looke W, Decock A, Ongenaert M, Vandesompele J, Speleman F and De Preter K. Promoter DNA methylation in the TH-MYCN neuroblastoma mouse model. Advances in Neuroblastoma Research 2014. May 13-16, 2014, Cologne, Germany.

Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Van Criekinge W, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F and Vandesompele J. Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. Cell Symposium: Cancer Epigenomics. October 6-8, 2013, Sitges, Spain.

Beckers A, Decock A, Cheung B, Carter D, Van Peer G, De Brouwer S, Laureys G, Vandesompele J, Schulte JH, Marshall G, Ongenaert M, De Preter K and Speleman F. The MYCN driven murine and human neuroblastoma methylome and dynamic regulation of chromatin regulators as a guide for epigenetic therapy. Cell Symposium: Cancer Epigenomics. October 6-8, 2013, Sitges, Spain.

Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Van Criekinge W, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F and Vandesompele J. Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. IUAP - P7/03 Progress Meeting. May 29, 2013, Leuven, Belgium.

Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Van Criekinge W, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F and Vandesompele J. Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers. EMBO Conference Series: Chromatin and Epigenetics. May 8-12, 2013, Heidelberg, Germany.

Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Van Criekinge W, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F and Vandesompele J. Identifying and validating DNA methylation biomarkers - the example of neuroblastoma. 13th Annual Meeting of the Belgian Society of Human Genetics: genetics of human development EXPOsed. March 15, 2013, Brussels, Belgium.

Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Van Criekinge W, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F and Vandesompele J. Promoter methylation analysis identifies prognostic methylation biomarkers in neuroblastoma. Advances in Neuroblastoma Research 2012. June 18-21, 2012, Toronto, Canada.

Ongenaert M, Decock A, Laureys G, Vandesompele J and Speleman F. Mapping the neuroblastoma epigenome: perspectives for improved prognostic biomarkers. Advances in Neuroblastoma Research

2012. June 18-21, 2012, Toronto, Canada.

Beckers A, Ongenaert M, Decock A, Kumps C, Pattyn F, the Neuroblastoma Research Consortium (NRC), Mestdagh P, Schulte JH, Speleman F and De Preter K. MiR-137 is epigenetically silenced in MYCN amplified neuroblastoma and targets the polycomb repressive complex 2 (PRC2) component EZH2. Advances in Neuroblastoma Research 2012. June 18-21, 2012, Toronto, Canada.

Beckers A, Ongenaert M, Decock A, Kumps C, Pattyn F, the Neuroblastoma Research Consortium (NRC), Mestdagh P, Schulte JH, Speleman F and De Preter K. MiR-137 is epigenetically silenced in MYCN amplified neuroblastoma and targets the polycomb repressive complex 2 (PRC2) component EZH2. Oncopoint. May 23, 2012, Ghent, Belgium.

Ongenaert M, Decock A, Vandesompele J and Speleman F. Mapping the neuroblastoma epigenome: perspectives for improved prognostic biomarkers. 12th Annual Meeting of the Belgian Society of Human Genetics: next generation sequencing and recent advances in genetics. March 2, 2012, Liège, Belgium.

Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Vandesompele J and Speleman F. Semi-high-throughput epigenetic validation using MSP on a real-time qPCR platform. Keystone Symposium Epigenomics (J3). January 17-22, 2012, Keystone, Colorado, USA.

Ongenaert M, Decock A, Vandesompele J and Speleman F. Mapping the neuroblastoma epigenome. Keystone Symposium Epigenomics (J3). January 17-22, 2012, Keystone, Colorado, USA.

Decock A, Ongenaert M, Vandesompele J and Speleman F. Semi-high-throughput epigenetic validation using MSP on a real-time qPCR platform. INSERM Workshop 211 High-throughput approaches in epigenomics. October 10-12, 2011, Bordeaux, France.

Ongenaert M, Decock A, Vandesompele J and Speleman F. Mapping the neuroblastoma epigenome. INSERM Workshop 211 High-throughput approaches in epigenomics. October 10-12, 2011, Bordeaux, France.

Decock A, Ongenaert M, Vandesompele J and Speleman F. Semi-high-throughput epigenetic validation using MSP on a real-time qPCR platform. EMBL Cancer Genomics. September 17-19, 2011, Heidelberg, Germany.

Ongenaert M, Decock A, Vandesompele J and Speleman F. Exploring the neuroblastoma methylome using MBD2-seq. EMBL Cancer Genomics. September 17-19, 2011, Heidelberg, Germany.

CONGRESSES, WORKSHOPS AND MEETINGS

OncoPoint. March 2, 2016, Ghent, Belgium.

1st European Cancer Epigenetics Conference. November 11-13, 2015, Maastricht, The Netherlands.

2015 Fournier-Majoie Foundation Friends Day. Valorisation. September 24, 2015, Brussels, Belgium.

OncoPoint. February 11, 2015, Ghent, Belgium.

4th scientific progress meeting Villa Joep!. December 12, 2014, Amsterdam, The Netherlands.

Advances in Neuroblastoma Research 2014. May 13-16, 2014, Cologne, Germany.

2014 Fournier-Majoie Foundation Friends Day and Award Ceremony. Essential feedback to guide the development of products to a significantly benefit for cancer patients. March 19, 2014, Ukkel, Belgium.

3rd scientific progress meeting Villa Joep!. December 9, 2013, Amsterdam, The Netherlands.

Cell Symposium: Cancer Epigenomics. October 6-8, 2013, Sitges, Spain.

IUAP - P7/03 Progress Meeting. May 29, 2013, Leuven, Belgium.

EMBO Conference Series: Chromatin and Epigenetics. May 8-12, 2013, Heidelberg, Germany.

13th Annual Meeting of the Belgian Society of Human Genetics: genetics of human development EXPOsed. March 15, 2013, Brussels, Belgium.

2013 Fournier-Majoie Foundation Friends Day. The biomarker steeple chase: how industry requirements impact early development of a biomarker. March 9, 2013, Brussels, Belgium.

Advances in Neuroblastoma Research 2012. June 18-21, 2012, Toronto, Canada.

Science Day Ghent University 2012. March 14, 2012, Gent, Belgium.

12th Annual Meeting of the Belgian Society of Human Genetics: next generation sequencing and recent advances in genetics. March 2, 2012, Liège, Belgium.

Keystone Symposium Epigenomics (J3). January 17-22, 2012, Keystone, Colorado, USA.

INSERM Workshop 211 High-throughput approaches in epigenomics. October 10-12, 2011, Bordeaux, France.

WOUD mini-symposium: bioinformatics: tools in research. September 28, 2011, Gent, Belgium.

2011 Fournier-Majoie Foundation Friends Day. Valorisation II: good practices in project management and intellectual property protection. September 24, 2011, Ukkel, Belgium.

EMBL Cancer Genomics. September 17-19, 2011, Heidelberg, Germany.

11th Annual Meeting of the Belgian Society of Human Genetics: dynamic DNA. March 4, 2011, Louvain-la-Neuve, Belgium.

Belgian Association for Cancer Research: hallmarks of cancer revisited and its implications for treatment. February 5, 2011, Liège, Belgium.

