Applied bioinformatics: Genomics of human and murine retinoblastoma

Kooi, E.I.

2017

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
Kooi, E. I. (2017). *Applied bioinformatics: Genomics of human and murine retinoblastoma*. Ipskamp Drukkers B.V.

Chapter 6

Summary, discussion and future perspectives

key findings, data interpretation, prospective studies, epigenetics, cancer genomics recommendations

6

SUMMARY

This thesis describes the results of genomic and transcriptomic profiling of primary enucleated human and murine retinoblastoma samples. Integration of the acquired high-dimensional data matrices from DNA, RNA and phenotype variables provided detailed insights into the genomics of retinoblastoma. Inherent to the complex nature of molecular biology, many of these analyses involve complex methodologies and results. To support a thorough discussion on this intricate matter, first the main findings of our studies are recapitulated in Table 1.

Table 1: Summary of the key findings

| Chapter	Key findings
2: Loss of photoreceptorness and gain of genomic alterations reveal tumor progression	• Retinoblastomas displayed significant variability in DNA and RNA composition • This variability was gradual, not categorical and therefore not indicative for subtypes • Increasing age at diagnosis, tumor volumes and genetic alterations correlated with decreasing differentiation grades and photoreceptor gene expression
3: A meta-analysis of retinoblastoma copy numbers refines the list of possible driver genes involved in tumor progression	• SCNAs concentrated at retinoblastoma-specific genomic regions • A candidate list of genes driving retinoblastoma progression is provided based on a meta-analysis of SCNA frequencies and gene-dosage effects • Between-tumor variability in number of genomic alterations correlated with age at diagnosis and differentiation grades • The amplitude of SCNAs was suggestive of within-tumor variability
4: Somatic genomic alterations in retinoblastoma beyond *RB1* are rare and limited to copy number changes	• Recurrent somatic alterations were restricted to *RB1*, *BCOR* and *CREBBP* • Variant allele frequencies of *BCOR* and *CREBBP* and *SCNA* amplitudes were suggestive of within-tumor heterogeneity for the majority of retinoblastomas. • 13q chromothrypsis targeting *RB1* can be identified by copy number profile data • Increasing age at diagnosis correlated with increasing SCNA frequencies • No evidence was found for the presence of tumor viruses

6

Chapter	Key findings
5: Genomic landscape of retinoblastoma in *Rb^{-/-}p130^{-/-}* mice resembles human retinoblastoma	• Opposed to *Rb^{-/-}p107^{-/-}* mice but similarly to human retinoblastoma, SCNAs are not an absolute requirement for full-blown retinoblastoma development in *Rb^{-/-}p130^{-/-}* mice • Mice that had tumors with SCNAs were diagnosed at later age than mice that had SCNA-neutral tumors. • Focal *Cdh11* (mouse) deletion confirms *CDH11* at 16q (human) as a retinoblastoma suppressor gene

In the individual **chapters 2-5**, the robustness of the results has been rigorously examined and discussed in light of published studies. Some results of individual chapters however, benefit from a joint discussion considering data from all chapters. Furthermore, some results deserve extra points of discussion in light of the newest literature. Also, some research questions remain to be fully answered and/or some results raised new questions and therefore, suggestions for future research are given throughout this chapter.

THREE INDEPENDENT GENE EXPRESSION STUDIES WITH SIMILAR DATA BUT DIFFERENT INTERPRETATIONS

In **chapter 2**, genome-wide gene expression was determined for a set of 76 retinoblastomas in order to study the between-tumor variability. While McEvoy *et al.* suggested that retinoblastomas are highly similar[1], Kapatai *et al.* claimed to have identified two distinct subtypes[2]. We concluded that there is profound between-tumor variability in gene expression, but that the gradual transition from one expression signature to the other and the relation to age at diagnosis, differentiation grades and tumor volumes supports a tumor progression model rather than the existence of distinct subtypes with different etiology. In summary, three genome-wide studies now have been conducted, one concludes "strikingly similar", one says two-subtypes and we say "strikingly different, but no subtypes". For the McEvoy study, gene expression data was publicly available and our re-analysis of this dataset, described in the discussion of **chapter 2**, showed that the McEvoy raw data was strikingly comparable to our data, yet our interpretations were very different. Although the raw data of the Kapatai study is not publicly available, we are confident that the gene expression data of Kaptai *et al.* are very

6

similar to both McEvoy *et al.* and our data. The genes and related ontologies that were differential between the two subtypes defined by Kapatai *et al.* virtually perfectly match the between-tumor variability in our and McEvoy's dataset. In short, although the raw data of three studies is very similar, the interpretations are very different. This discrepancy between data and interpretation requires some extra contemplation.
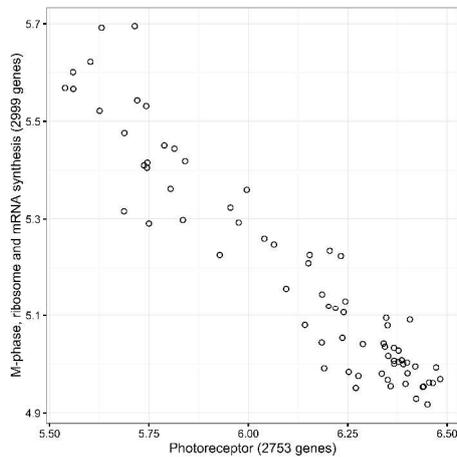
Figure 1: Gradual variability in gene expression signatures. Unsupervised hierarchical clustering was used to separate the tumor cohort into two tumor groups. Differentially expressed genes between these two groups were separated in two gene signatures: photoreceptor (2753 genes) and M-phase, ribosome and mRNA synthesis (2999 genes). For each tumor sample, the mean gene expression for these two signatures is plotted against each other.

In favour of Kapatai *et al*., one could argue that although the between-tumor differences are gradual, separation into two groups is possible. Given the uniform distribution of photoreceptor and M-phase, ribosome and mRNA synthesis expression, there is just as much reason to make 3, 4, 5, or any number of groups, depending on the sample size of the cohort. So, to divide retinoblastoma in an arbitrary number of 2 groups and call these groups cancer subtypes with possibly different etiologies is a misinterpretation of the data. In case of robust subtypes, the similarity of members that belong to the same subtype is large, while the similarity of members that belong to different subtypes is small. For example, an apple is a type of fruit and apples can be further categorized into robust subtypes such as granny smith and red delicious apples. Granny smith apples are green and taste sour, while red delicious apples taste sweet and are obviously red. So apples that belong to the same subtype, say granny smith apples, are very similar, while apples from different subtypes, a granny smith compared to a red delicious apple, are very different. A clear categorical separation is not apparent for gene expression signatures of retinoblastomas, evidenced by the linear, and

6

not sigmoidal, relation of photoreceptor signature with M-phase, ribosome and mRNA synthesis signature (Figure 1). Furthermore, the strong relation of decreasing photoreceptorness with increasing age at diagnosis, tumor volumes and dedifferentiation is highly suggestive for tumor progression.

Reese-Elsworth classification ○ III ○ IV ○ V ○ Not available
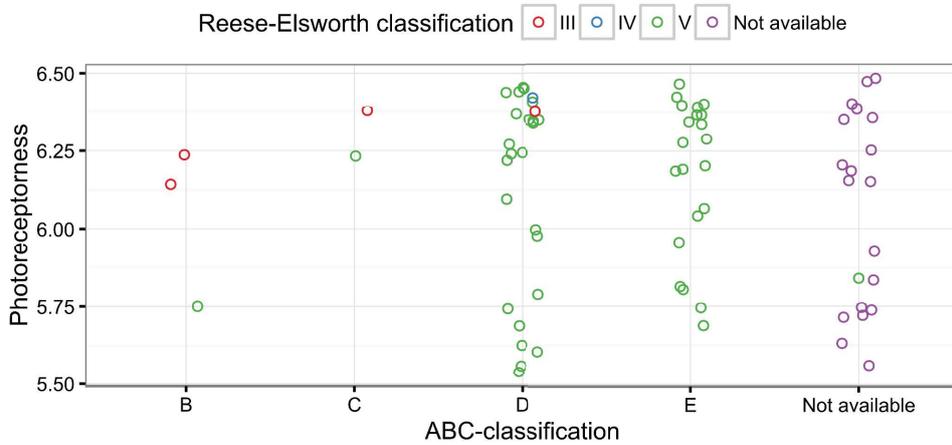
Figure 2: Clinical staging of retinoblastoma patients from chapter 2. Each dot represents a retinoblastoma eye, where the y-axis denotes the mean gene expression of the photoreceptor signature.

In favour of McEvoy *et al.* concluding that retinoblastomas are similar, one could say that although the between-tumor variability in gene expression of our cohort might exist, this is not clinically relevant, because tumors with different clinical staging did not have different photoreceptorness. In our defense, we believe that the clinical staging data of our cohort should be reconsidered. The clinical stages by Reese-Elsworth and the ABC-classification are visualized in the content of photoreceptorness (Figure 2). For 20/72 (28%) tumors, staging is not available for both classification systems. More importantly, 45/47 (96%) of group D and E eyes (X-axis Figure 2) are staged as group V eyes in the Reese-Elsworth classification (green dots Figure 2). This indicates that distinction of in particular group D and E eyes can be challenging. Also because the staging data of our cohort has been collected over the last two decades, we think this data needs to be carefully revised in light of discussion about staging definitions in the last few years[3–9]. At the time of writing, (May 2016) the gene expression cohort is being characterized in more than 25 magnetic resonance imaging features

6

by three independent radiologists. We anticipate that high-risk features
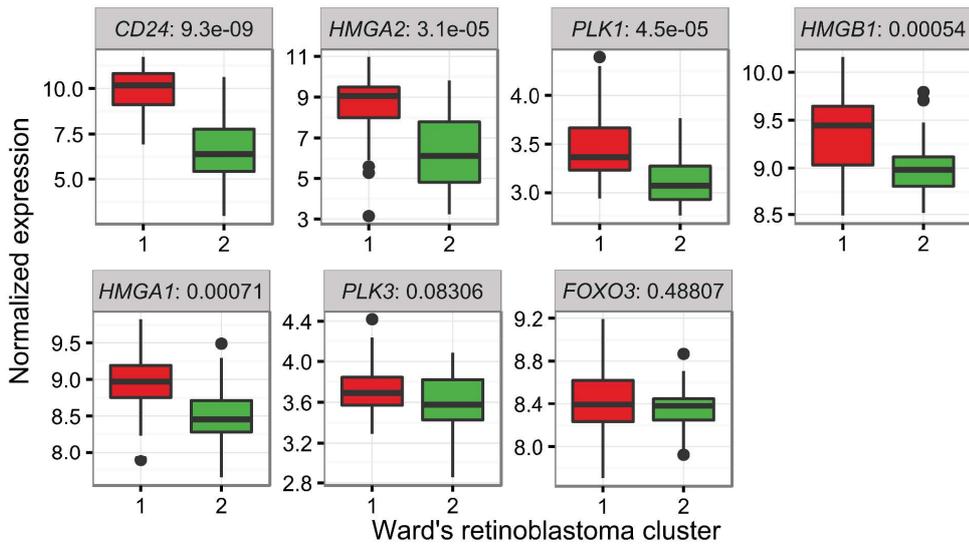
Figure 3: Expression values in our cohort for genes that were studied for the relation between expression and clinicopathological risk factors. Boxplots of normalized gene expression values of tumors described in chapter 2. Ward's retinoblastoma clusters 1 tumors (advanced tumors) have more genomic alterations, low photoreceptorness, are poorly differentiated, are diagnosed at late age and are large in volume relative to Ward's retinoblastoma cluster 2 tumors. P-values denote multiple-testing corrected p-values in gene expression between tumor clusters 1 and 2.
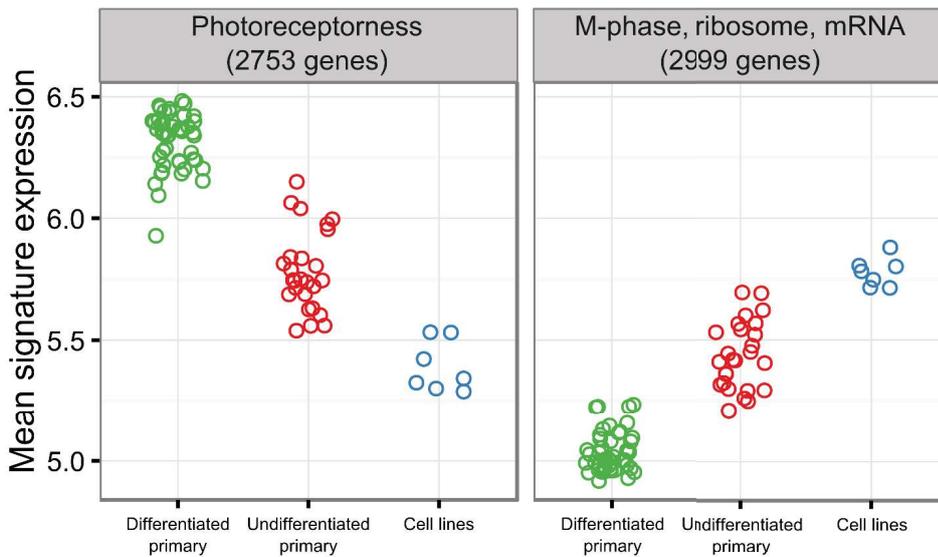
Figure 4: Gene signature expression of retinoblastoma cell lines. In terms of the gene expression signatures, retinoblastoma cell lines (RB176, RB247, RB381, RB383, WERI-RB1, Y79 and three isolations subsequent cell culture passages of *MYCN*-amplified cell line VUMC-RB-26) are at the extreme end of retinoblastoma progression; cell lines have the lowest photoreceptorness and the highest M-phase, ribosome and mRNA synthesis signature expression.

such as large tumor volumes, the number of lesions and high numbers of vitreous and/or subretinal seeds will significantly correlate with decreased photoreceptorness.

CONFIRMATION OF THE RELATION BETWEEN GENE EXPRESSION AND PHENOTYPE VARIABLES RESULTS BY LATER STUDIES

After **chapter 2** was published in July 2015, several studies that correlated clinical and histopathological features with mRNA or protein levels were performed. The results of these studies can now be used to test whether our analyses were validated by these prospective cohorts. Table 2 summarizes the results of these studies. Genes that correlated with advanced-stage clinical or pathological features included *CD24*[10], *PLK1*[11], *HMGA1*[12], *HMGA2*[12], *HMGB1*[13], and *FOXO3*[14] , while no relation could be found for *PLK3*[11]. In agreement with these results, in our dataset, expression of *CD24, PLK1, HMGA1, HMGA2 and HMGB1* was significantly upregulated in retinoblastoma cluster 1 (46 tumors, low photoreceptorness) compared to 2 (26 tumors, high photoreceptorness) (Figure 3). Furthermore, *PLK3* was not differentially expressed in our data set (Figure 3), in agreement with Singh *et al*.[11]. While *FOXO3* expression was found to be borderline significantly (p=0.04) related to increased risk for choroidal invasion[14], *FOXO3* did not show significantly variability in our dataset (Figure 3). These insights suggest that **chapter 2** can predict outcome of other further studies on molecular markers for tumor progression. For example, based on **chapter 2**, we predict that immune staining of *HMGB3*, in our dataset the most significantly upregulated HMGB-gene (FDR 5.79E-09), will demonstrate significantly increased expression in advanced stage retinoblastoma. We suggest that the gene expression differences related to tumor progression described in **chapter 2** can be used to design tumor-tailored therapeutic intervention strategies.

6

Table 2: Studies on the relation between clinicopathological variables and mRNA or protein expression of a selected gene of interest that were published after chapter 2 was published in July 2015.

First author	Title	Publication date
Ishaq[10]	Correlation of *CD24* expression with histological grading and TNM staging of retinoblastoma	Jan 2016
Singh[12]	Role of High-mobility Group Protein A Isoforms and Their Clinicopathologic Significance in Primary Retinoblastoma	Dec 2015
Batra[15]	Expression of *FOXO3a* and Correlation With Histopathologic Features in Retinoblastoma	Nov 2015
Singh[13]	Correlation of High Mobility Group Box-1 Protein (*HMGB1*) with Clinicopathological Parameters in Primary Retinoblastoma	Sep 2015
Singh[11]	Prognostic significance of polo-like kinases in retinoblastoma: correlation with patient outcome, clinical and histopathological parameters.	Aug 2015

POSITIONING OF RETINOBLASTOMA CELL LINES IN THE SPECTRUM OF PRIMARY RETINOBLASTOMAS

In **chapter 2** it is described that in primary retinoblastomas, decreasing photoreceptor signature is correlated with increasing total genomic disruption. In **chapter 3**, it is described that the total genomic disruption of retinoblastoma cell lines is comparable to the 25% most disrupted primary tumors. This implies that cell lines represent a late stage of retinoblastoma progression and should therefore have a relatively low expression of the photoreceptor signature and high expression of M-phase, ribosome and mRNA synthesis signature. Although we also profiled gene expression profiles of retinoblastoma cell lines, these data have not been discussed. In Figure 4 it is shown that cell lines indeed have low expression of the photoreceptor and high expression of the M-phase, ribosome and mRNA synthesis signature relative to the primary tumors. So both on DNA and RNA level, cell lines appear to represent advanced-stage retinoblastoma and therefore can be a useful model for the pre-clinical evaluation of drug sensitivity in advanced-stage retinoblastoma. The reason why cell lines represent late-stage retinoblastoma can be that biopsies from advanced-stage retinoblastomas are easier to put in culture. Alternatively, biopsies

from primary retinoblastomas might contain a mixture of early and advanced-staged cells, where the advantage-staged cells are selected for during in vitro culturing.

RELEVANCE OF SECONDARY ALTERATIONS

In **chapter 2** we present support for the hypothesis that retinoblastoma cells can progress through accumulation of genomic alterations. Similarly to others, we hypothesized that the identification of the genomic alterations that drive tumor progression can help design targeted therapy. **Chapter 3** describes a prioritization for candidate genes that are likely to drive tumor progression, based on the frequency of copy number alterations and its impact on gene expression. Importantly, **chapter 3** also describes that SCNAs were not evident in each individual tumor, and if present, not in each tumor cell. This challenged the hypothesis that secondary alterations are an absolute requirement for cancer development. However, we could not rule out that tumor samples without SCNAs were contaminated with non-cancer cells or had alternative genomic alterations, such as single nucleotide variants. This is however ruled out in **chapter 4**, showing that the majority (>95%) of retinoblastoma samples contain very little non-cancer cell contamination (<5%). Furthermore, aside from *RB1*, recurrently mutated genes are very rare in retinoblastoma. In addition, **chapter 4** again showed that not all retinoblastoma samples displayed SCNAs and if present, SCNAs were mostly subclonal. Since secondary alterations could not be detected in all tumors and also often appeared to be subclonal, targeted inhibition of these alterations might not target all tumor cells and therefore might not be sufficient for complete tumor control. On the other hand, the identified secondary alterations might indicate what molecular circuitry is most beneficial for retinoblastoma cells, which might be applicable to all retinoblastoma cells. For example, it could be that retinoblastoma cells require the proper function of *MYCN* and benefit from *MYCN* amplification. The fact that in only some tumors and only some cells, *MYCN* might have been accidently amplified and been selected for by clonal evolution doesn't discredit anti-*MYCN* treatment for retinoblastoma cells with normal number

6

of *MYCN* copies. In fact, it may be quite the opposite, where *MYCN*-amplified cells are more tolerant to *MYCN* inhibition since they have so many *MYCN* copies. In all, molecular characterization of retinoblastomas described in **chapter 3, 4 and 5** has provided several candidate genes that allegedly drive tumor progression. Given the frequency of occurrence, these alterations are likely to be relevant for retinoblastoma cells. Whether these gene candidates are useful targets for efficacious retinoblastoma treatment now has to be determined in pre-clinical experiments.

A ROLE FOR EPIGENETICS?

During the highly regulated development of a fertilized oocyte to an adult organism, cells experience a massive reorganization of their epigenetic landscapes[16,17]. During maturation, the ability of cells to adopt any cell fate becomes more and more restricted through epigenetic changes[18–20]. Epigenetics can also play a role in the cellular transition into a malignant state[21–23]. The most well-known form of epigenetic regulation is DNA-methylation[24], which can control the level of gene expression. Hypermethylation of a gene promoter region can form a physical block for transcription factors, thereby interfering with gene function. Silencing of *RB1* expression through promotor hypermethylation is described as one of the retinoblastoma initiating events[25,26]. Similarly to genomic alterations, epigenomic alterations beyond *RB1* might further drive tumor progression. For 15 retinoblastoma samples that were exome-sequenced and described in **chapter 4**, we also performed genome-wide DNA methylation quantification by array, although this data has not been described in this thesis until now. Unsupervised hierarchical clustering with Ward, Complete-linkage, average-linkage and McQuitty consistently divided the 15 samples into 4 samples that had focal *MYCN* amplification and 11 samples without *MYCN* amplification (Figure 5A-D). Therefore, we performed supervised clustering of the 15 tumors with regions that were differentially methylated between tumors with and without *MYCN* amplification (Figure 5E). We concluded that tumors with *MYCN* amplification were hypomethylated and that loss of methylation was correlated with increased age at diagnosis and thereby an increased number

of genomic alterations (Figure 5F). It is difficult to disseminate what is cause or consequence of genomic and epigenomic alterations. Thereby, possibly deregulation of both genomics and epigenomics jointly affect retinoblastoma development. Transfection of a MYCN-amplified vector in retinoblastoma cell lines might reveal the epigenetic consequence of *MYCN* amplification. Thereby, an extra layer of complexity in the analysis of epigenetics in primary tissue is that epigenetics changes over time during tissue normal development[27–29]. Therefore, age-matched controls might be required for a proper analysis of the relations between (epi)genomics and retinoblastoma development.

Figure 5: Epigenetics of primary retinoblastoma. (A-D) Unsupervised hierarchial clustering of methylation values (M-values < 0: no methylation, > 0 methylation) with Ward, Complete-linkage, average-linkage and McQuitty algorithms separated tumors based on *MYCN* amplification (Samples, T7, T11, T14 and T72, red colored branches of the dendrograms). (E) Heatmap of M-values for differentially methylated regions between *MYCN* amplified and non-amplified tumors. Color-coded samples information is given on top of the heatmap for MYCN amplification (red=amplification, black = non-amplification) and age at diagnoses (green=youngest, red=oldest). (F) Analysis of variance of M-values in relation to MYCN amplification and age at diagnosis. MYCN-amplification was very significantly related to decreased methylation values (p-value 6.67E-08). Adjusting for the *MYCN*-amplification relation with methylation values, increased age at diagnosis was significantly related to decreased methylation values (p-value 0.0374).

FUTURE RESEARCH ON RETINOBLASTOMA DRIVER IDENTIFICATION AND PRECISION MEDICINE

This thesis describes molecular profiling of retinoblastoma samples in order to understand oncogenesis and to design precision medicine that targets tumor cells specifically. We discovered that retinoblastomas progress from photoreceptor-like cells to undifferentiated cells through SCNAs. We and others hypothesized that the SCNAs include many passenger genes and only a limited set of driver genes, whose expression is changed through a gene-dosage effect and thereby promote tumor progression. In the last decade, several studies were performed aimed at identifying the minimal region of gain or loss for commonly altered chromosome arms 1q, 2p, 6p and 16q[30,31]. The strategy was to include as many retinoblastoma samples as possible and define the genomic intersect of the identified SCNAs. Since retinoblastoma is a rare disease, cohorts were however usually small (<50 tumor samples)[32–38], limiting the study power. Therefore, we performed a meta-analysis using all available data and integrated these data with new high-resolution data, matched with gene expression. Although this strategy was successful to a limited degree, it was not so simple that one or two genes per commonly altered chromosome arm could be assigned as candidates, as is the case for focally amplified *MYCN* at 2p[34,35,39,40]. Since we now know that tumors diagnosed at higher age, with larger volumes and more poorly differentiated tumors contain more alterations with higher amplitudes (**chapter 3 and 4**), we suggest that future studies may select highly advanced staged tumors, to maximize the chance to detect focal amplifications. Preferably, samples of extra-ocular retinoblastoma or metastasis should be used for such analysis. Furthermore, since we believe that within-tumor heterogeneity can exist, we suggest spatial sampling prior to copy number determination, in particular for advanced-staged tumors. Alternatively, multi-dimensional fluorescent activated cell sorting might be used to sort separate cells from different tumor clones. Since retinoblastoma biopsies can be relatively small, sub-sampling might only yield very limited quantities of DNA. In that case, single-cell genomics should be considered[41].

6

Previously, SCNAs of metastases of murine retinoblastoma samples were determined[42–44], reasoning that in these advanced-stage tumor cells, alterations that drive proliferation have been highly selected for. This yielded multiple focal high-level gains and losses that included single genes only such as *Mdm2*[44], *Cdkn2A*[44], *Mycn*[45] and *Mir17-92*[43]. These single-gene SCNAs overlapped with the larger SCNAs that were found in the matching primary tissues, and were therefore considered the most likely target genes of the non-focal copy number alterations. These results are a showcase of how careful sample selection can drastically improve identification rates of driving forces of tumor progression.

On the other hand, it might not even be necessary to identify the driving genes of SCNA regions in order to design retinoblastoma targeting therapy. Differential gene expression analysis of photoreceptor-like and photoreceptor-unlike tumors in **chapter 2** provides a list of genes that are deregulated during tumor progression and might prove to be very resourceful. For example, the second-most significantly upregulated gene was the somatostatin receptor 2 (*SSTR2*). Studies that successfully exploited *SSTR2* expression for targeted delivery of somatostatin analogous conjugated to anti-tumor effector molecules are plentiful[46–50] and could be a promising lead for retinoblastoma as well. Furthermore, expression of *SSTR2* in retinoblastoma might be used as a target for nuclear imaging using somatostatin-based radiopeptides, similarly as currently being used for imaging of neuroendocrine tumors[51–54]. Instead of selecting the most significantly upregulated genes, a cross-reference of any significantly differentially expressed gene from **chapter 2** with well-estabilished targets for anti-cancer treatment might reveal promising leads. The inhibitory effect on retinoblastoma cell viability could be tested in retinoblastoma cells or mouse models using for example siRNAs, shRNAs or CRISPR-Cas9 knockdown/out screens. In addition to evidence-based selection of anti-cancer targets by genomics profiling, improvements in drug delivery might be essential for effective retinoblastoma treatment. For example, through (epi-)genomic analysis, inhibition of *SYK* was identified as a lead for retinoblastoma therapy[36]. In a follow-up study, an attempt to treat retinoblastoma mice with a SYK antagonist failed[55] due to insufficient drug delivery to the vitreous.

6

Considering the high success rates for eye-preserving treatments using selective intra-arterial chemotherapy and intra-vitreal chemotherapy with conventional chemotherapeutics[56–58], sufficient drug delivery indeed appears to be essential. Future studies on the late effects in terms of efficacy, toxicity and complications of these recently (re-)explored treatments in combination with non-invasive clinical and radiological predictive markers are now urgently required to further improve retinoblastoma care.

RECOMMENDATIONS FOR CANCER GENOMICS STUDIES

Through the analysis of various types of high-dimensional (epi-)genomic and transcriptomic data, I not only learnt about the complex molecular circuitry of retinoblastoma, I also experienced the challenges that come into play when designing, performing and analysing cancer genomics experiments, in particular studies using next-generation sequencing approaches. Based on my experiences, here I would like to provide some recommendations for future cancer genomics studies that might help to design and perform the experiments effectively. Expressions in this paragraph should be considered as my personal opinion. References to scientific research articles and reviews are therefore omitted.

- **Multidisciplinary:** Cancer genomics should be contemplated as a multidisciplinary research field. It requires active participation of at least a pathologist, a clinician that treats the patient group, a biologist, a statistician and a bio-informatician who all should have minimally 4 years of relevant working experience.

- **Study design:** After the research aims have been defined by the multidisciplinary cancer genomics team, discussion of study design prior to the execution of any experiment should be top priority. What sample preparation is required? Which genomics technique is most applicable? How much sequencing data is required? What are the endpoints of quality control? How many samples are required for an experiment to have sufficient power? What are relevant comparisons?

Are technical replicates required? Is it possible to perform paired experiments? What negative and positive controls are essential to interpret the results?

- **Pilot study:** In some cases, it can be hard to answer some of the questions related to the study design. For example, the number of sequencing reads required for accurate detection of transcript structures in complex cancer samples depends on many factors, which cannot all be predicted beforehand. Therefore it can be very useful to perform a series of pilot experiments. For example, instead of sequencing an NGS library at the estimated desired depth instantly, the library can be sequenced in series where the number of sequencing reads is gradually increased until a plateau in quality is achieved. By doing so, not only the cost-quality balance is optimized, this also warrants high quality data acquisition.

- **Between-tumor heterogeneity:** Cancer is a heterogeneous disease, where tumors not only differ between cancer types, but also within cancer types. Studies that aim to identify common events need to realize that in an unselected heterogeneous cohort, a large sample size will be required to identify commonalities. Alternatively, selection of samples in order to decrease the sample population diversity can drastically increase the chances of finding common traits. Conversely, studies aimed to relate genotypes to phenotypes might benefit from large population diversity and should therefore randomly sample or perform stratified subsampling to warrant sample diversity to increase the study power.

- **Within-tumor heterogeneity:** Next to between-tumor variability, significant diversity between tumor cells within the same tumor has been described[59]. Although there is increasing appreciation of within-tumor heterogeneity, still many cancer genomic studies depend on a single sample per tumor, or even per patient. A single sample of a polyclonal tumor in a patient with multiple lesions might therefore

6

be inadequate to base either fundamental or clinical conclusions on.

- **Non-tumor cell contamination:** In my opinion, one of the most common pitfalls of cancer genomics studies is that non-tumor cell contamination is not reported or even assessed. For example it was recently reported that for whole genome sequencing from 14,000 patients by the International Cancer Genome Consortium, in 92% of the cases, data about non-tumor cell contamination was not available[60]. Particularly for DNA analysis, contamination with diploid wild-type non-tumor DNA confounds the identification of tumor variants and should be accounted for. Depending on the research questions, non-tumor cells should either be removed from the sample prior to profiling or should be accounted for during analysis. Of note, there are numerous bioinformatic tools that claim to accurately determine tumor cellularity using copy number and allele frequency data. However, they often depend on assumptions about ploidy and heterogeneity that cannot always be guaranteed, in my believe. The best way to determine non-cancer cell contamination is to evaluate the variant allele frequency of the tumor initiating mutation, which should be present in all tumor cells and not in any non-tumor cells. Admittedly, this method also depends on the assumption that tumor evolution is hierarchical, where all tumor cells descended from a single initiating tumor cell and daughter cells inherited and maintained the initiating mutation.

- **Sample identification:** Whatever can go wrong, will go wrong. During the process of collecting tumor samples for genomic profiling, it is not unlikely that samples can get swapped accidently. Therefore it can be very useful to be able to validate the identity of the individual samples. For example, in case independent genetic data is available about included samples, this can be used to verify sample identities. To the least, in DNA or RNA studies, respectively copy number or gene expression of sex chromosomes should be correlated with sex phenotypes. In case tumor-normal DNA or RNA profiling is

performed, hierarchical clustering of single nucleotide polymorphism (SNP) genotypes can be used to validate that the appropriate sample pairings were used.

- **Get best out of data:** Although a genomics approach might be designated to collect a particular kind of data, it might also be used for other purposes. For example, SNP arrays were first only used for SNP genotyping but are now widely used for copy number determination additionally. Similarly, exome-sequencing is primarily performed for SNV/INDEL detection but can be used for copy number, loss of heterozygosity or even virus quantification analysis. Another example is RNA sequencing, which is mostly used for determination of transcript abundance or structure, but can also be used for SNV/INDEL analysis.

- **Data sharing:** There is increasing awareness that sharing high-dimensional genomics data is essential for the cancer genomics field to make sustainable translational contributions. This way, safe long-term data storage is ensured, independent researchers can reproduce and thereby validate analyses, perform powerful meta-analysis or use publicly available data to help interpret new data.

6

- **Phenotype data:** Well-documented, complete and validated phenotype data is absolutely essential for the translational interpretation of genomic data. Considerable time and effort should be devoted to collect high-quality phenotype data. Most importantly, similarly to genomics data, sharing of phenotype data with due consideration of patient consent can greatly enhance the force of cancer genomics in the war on cancer.

As a final remark, I would like to stress that although cancer genomics studies yielded many insights into the molecular pathways that drive cancer, cancer genomics researchers might want to temper any optimism about the contribution of cancer genomics to improved health care. Surely, next-

generation sequencing revolutionized genetics, leading to an exponential increase in sequencing throughput. However, although numerous clinical trials with targeted therapies based on cancer genomics are currently in progress, the revolutionizing impact of cancer genomics on long-term survival rates remains to be demonstrated in the coming years. In combination with targeted drug delivery, immune therapy and combinatorial medicine, identification of the tumor-specific molecular essentials will hopefully open up new avenues for the save and effective control of cancer.

REFERENCES

1. McEvoy, J. *et al.* Coexpression of normally incompatible developmental pathways in retinoblastoma genesis. *Cancer cell* **20,** 260–75 (2011).
2. Kapatai, G. *et al.* Gene expression profiling identifies different sub-types of retinoblastoma. *British journal of cancer* **109,** 512–25 (2013).
3. Shields, C. L. *et al.* The International Classification of Retinoblastoma predicts chemoreduction success. *Ophthalmology* **113,** 2276–80 (2006).
4. Linn Murphree, A. Intraocular retinoblastoma: the case for a new group classification. *Ophthalmology clinics of North America* **18,** 41–53, viii (2005).
5. Reese, A. B. & Ellsworth, R. M. The evaluation and current concept of retinoblastoma therapy. *Transactions - American Academy of Ophthalmology and Otolaryngology. American Academy of Ophthalmology and Otolaryngology* **67,** 164–72 (1963).
6. Chantada, G. *et al*. A proposal for an international retinoblastoma staging system. *Pediatric blood & cancer* **47,** 801–5 (2006).
7. Sastre, X. *et al.* Proceedings of the consensus meetings from the International Retinoblastoma Staging Working Group on the pathology guidelines for the examination of enucleated eyes and evaluation of prognostic risk factors in retinoblastoma. *Archives of pathology & laboratory medicine* **133,** 1199–202 (2009).
8. Novetsky, D. E., Abramson, D. H., Kim, J. W. & Dunkel, I. J. Published international classification of retinoblastoma (ICRB) definitions contain inconsistencies--an analysis of impact. *Ophthalmic genetics* **30,** 40–4 (2009).
9. Shields, C. L. & Shields, J. A. Basic understanding of current classification and management of retinoblastoma. *Current Opinion in Ophthalmology* **17,** 228–234 (2006).
10. Ishaq, S. M., Kehar, S. I., Zafar, S. & Hasan, S. F. U. Correlation of CD24 expression with histological grading and TNM staging of retinoblastoma. *Pakistan journal of medical sciences* **32,** 160–4 (2016).
11. Singh, L. *et al.* Prognostic significance of polo-like kinases in retinoblastoma: correlation with patient outcome, clinical and histopathological parameters. *Clinical & experimental ophthalmology* **43,** 550–7 (2015).
12. Singh, M. K. *et al.* Role of High-mobility Group Protein A Isoforms and Their Clinicopathologic Significance in Primary Retinoblastoma. *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* (2015). doi:10.1097/PAI.0000000000000295
13. Singh, M. K. *et al*. Correlation of High Mobility Group Box-1 Protein (HMGB1) with Clinicopathological Parameters in Primary Retinoblastoma. *Pathology oncology research : POR* **21,** 1237–42 (2015).

6

14. Batra, A., Kashyap, S., Singh, L. & Bakhshi, S. Expression of FOXO3a and Correlation With Histopathologic Features in Retinoblastoma. *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* (2015). doi:10.1097/PAI.0000000000000278

15. Batra, A., Kashyap, S., Singh, L. & Bakhshi, S. Sirtuin1 Expression and Correlation with Histopathological Features in Retinoblastoma. *Ocular Oncology and Pathology* **2,** 86–90 (2015).

16. Rodríguez-Rodero, S., Fernández-Morera, J. L., Fernandez, A. F., Menendez-Torre, E. & Fraga, M. F. Epigenetic Regulation of Aging. *Discovery Medicine* **10,** 225–233 (2010).

17. Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447,** 425–32 (2007).

18. Khavari, D. A., Sen, G. L. & Rinn, J. L. DNA methylation and epigenetic control of cellular differentiation. *Cell Cycle* **9,** 3880–3883 (2014).

19. Lotem, J. & Sachs, L. Epigenetics and the plasticity of differentiation in normal and cancer stem cells. *Oncogene* **25,** 7663–72 (2006).

20. Mohn, F. & Schübeler, D. Genetics and epigenetics: stability and plasticity during cellular differentiation. *Trends in genetics : TIG* **25,** 129–36 (2009).

21. Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nature reviews. Cancer* **4,** 143–53 (2004).

22. Rodríguez-Paredes, M. & Esteller, M. Cancer epigenetics reaches mainstream oncology. *Nature medicine* **17,** 330–9 (2011).

23. Brower, V. Epigenetics: Unravelling the cancer code. *Nature* **471,** S12–3 (2011).

24. Jones, P. A. The Role of DNA Methylation in Mammalian Epigenetics. *Science* **293,** 1068–1070 (2001).

25. Greger, V., Passarge, E., Hopping, W., Messmer, E. & Horsthemke, B. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Human Genetics* **83,** 155–158 (1989).

26. Sakai, T. *et al.* Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *American journal of human genetics* **48,** 880–8 (1991).

27. Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nature reviews. Genetics* **3,** 662–73 (2002).

28. Feng, S., Jacobsen, S. E. & Reik, W. Epigenetic reprogramming in plant and animal development. *Science (New York, N.Y.)* **330,** 622–7 (2010).

29. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nature reviews. Genetics* **9,** 129–40 (2008).

30. Thériault, B. L., Dimaras, H., Gallie, B. L. & Corson, T. W. The genomic landscape of retinoblastoma: a review. *Clinical & experimental ophthalmology* **42,** 33–52

31. McEvoy, J. *et al.* RB1 gene inactivation by chromothripsis in human retinoblastoma. *Oncotarget* **5,** 438–50 (2014).

32. van der Wal, J. E. *et al.* Comparative genomic hybridisation divides retinoblastomas into a high and a low level chromosomal instability group. *Journal of Clinical Pathology* **56,** 26–31 (2003).

33. Zielinski, B. *et al.* Detection of chromosomal imbalances in retinoblastoma by matrix-based comparative genomic hybridization. *Genes, chromosomes & cancer* **43,** 294–301 (2005).

34. Bowles, E. *et al.* Profiling genomic copy number changes in retinoblastoma beyond loss of RB1. *Genes, chromosomes & cancer* **46,** 118–129 (2007).

35. Mol, B. M. *et al.* High resolution SNP array profiling identifies variability in retinoblastoma genome stability. *Genes, chromosomes & cancer* **53,** 1–14 (2014).

36. Zhang, J. *et al.* A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature* **481,** 329–34 (2012).

37. Herzog, S. *et al.* Marked differences in unilateral isolated retinoblastomas from young and older children studied by comparative genomic hybridization. *Hum.Genet.* **108,** 98–104 (2001).

6

38. Sampieri, K. *et al.* Array comparative genomic hybridization in retinoma and retinoblastoma tissues. *Cancer science* **100,** 465–471 (2009).
39. Rushlow, D. E. *et al.* Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *The Lancet. Oncology* **14,** 327–34 (2013).
40. Mairal, A. *et al.* Detection of chromosome imbalances in retinoblastoma by parallel karyotype and CGH analyses. *Genes, chromosomes & cancer* **28,** 370–379 (2000).
41. Macaulay, I. C. & Voet, T. Single cell genomics: advances and future perspectives. *PLoS genetics* **10,** e1004126 (2014).
42. Macpherson, D. Insights from mouse models into human retinoblastoma. *Cell division* **3,** 9 (2008).
43. Conkrite, K. *et al.* miR-17~92 cooperates with RB pathway mutations to promote retinoblastoma. *Genes & development* **25,** 1734–1745 (2011).
44. Conkrite, K., Sundby, M., Mu, D., Mukai, S. & MacPherson, D. Cooperation between Rb and Arf in suppressing mouse retinoblastoma. *Journal of Clinical Investigation* **122,** 1726–1733 (2012).
45. MacPherson, D. *et al.* Murine bilateral retinoblastoma exhibiting rapid-onset, metastatic progression and N-myc gene amplification. *The EMBO journal* **26,** 784–94 (2007).
46. Guillermet, J. *et al.* Somatostatin receptor subtype 2 sensitizes human pancreatic cancer cells to death ligand-induced apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 155–60 (2003).
47. Chinol, M., Bodei, L., Cremonesi, M. & Paganelli, G. Receptor-mediated radiotherapy with 90Y-DOTA-DPhe1-Tyr3-octreotide: The experience of the european institute of oncology group. *Seminars in Nuclear Medicine* **32,** 141–147 (2002).
48. Imhof, A. *et al.* Response, survival, and long-term toxicity after therapy with the radiolabeled somatostatin analogue [90Y-DOTA]-TOC in metastasized neuroendocrine cancers. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29,** 2416–23 (2011).
49. Zhang, J. *et al.* A novel octreotide modified lipid vesicle improved the anticancer efficacy of doxorubicin in somatostatin receptor 2 positive tumor models. *Molecular pharmaceutics* **7,** 1159–68 (2010).
50. Sun, L.-C. & H. Coy, D. Somatostatin Receptor-Targeted Anti-Cancer Therapy.
51. Maecke, H. R. & Reubi, J. C. Somatostatin receptors as targets for nuclear medicine imaging and radionuclide treatment. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* **52,** 841–4 (2011).
52. Kwekkeboom, D. J. *et al.* Somatostatin-receptor-based imaging and therapy of gastroenteropancreatic neuroendocrine tumors. *Endocrine-related cancer* **17,** R53–73 (2010).
53. de Herder, W. W. *et al.* Neuroendocrine tumors and somatostatin: imaging techniques. *Journal of endocrinological investigation* **28,** 132–6 (2005).
54. de Jong, M., Breeman, W. A. P., Kwekkeboom, D. J., Valkema, R. & Krenning, E. P. Tumor imaging and therapy using radiolabeled somatostatin analogues. *Accounts of chemical research* **42,** 873–80 (2009).
55. Pritchard, E. M. *et al.* Pharmacokinetics and efficacy of the spleen tyrosine kinase inhibitor r406 after ocular delivery for retinoblastoma. *Pharmaceutical research* **31,** 3060–72 (2014).
56. Abramson, D. H. *et al.* Advanced Unilateral Retinoblastoma: The Impact of Ophthalmic Artery Chemosurgery on Enucleation Rate and Patient Survival at MSKCC. *PloS one* **10,** e0145436 (2015).
57. Suzuki, S., Aihara, Y., Fujiwara, M., Sano, S. & Kaneko, A. Intravitreal injection of melphalan for intraocular retinoblastoma. *Japanese journal of ophthalmology* **59,** 164–72 (2015).
58. Munier, F. L., Gaillard, M.-C., Balmer, A. & Beck-Popovic, M. Intravitreal chemotherapy for vitreous seeding in retinoblastoma: Recent advances and perspectives. *Saudi journal of ophthalmology : official journal of the Saudi Ophthalmological Society* **27,** 147–50 (2013).

6

59.   Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nature reviews. Cancer* **12,** 323–34 (2012).

60.   Rubin, M. A. Health: Make precision medicine work for cancer care. *Nature* **520,** 290–1 (2015).

6