Neuroblastoma and Whole Genome Searches

Mario Capasso
University of Naples Federico II, Department of Biochemistry and Medical Biotechnology, CEINGE Biotecnologie Avanzate, Italy

1. Introduction

Neuroblastoma is the most frequent solid cancer of early childhood, its incidence is 10.2 cases per million children under 15 years of age. Neuroblastoma is thought to be an embryonal tumor that is derived from precursor cells of the peripheral (sympathetic) nervous system (Hoehner et al., 1996; Maris, 2010). The tumor can arise anywhere along the sympathetic chain but is most frequently in the adrenal medulla and paraspinal ganglia (Maris, 2010).

The most fascinating feature of neuroblastoma is its remarkable biological heterogeneity, which becomes apparent in the breadth of the clinical courses of the disease. While some patients experience spontaneous regression or differentiation of the tumor into benign ganglioneuroma, others are affected by rapid and fatal tumor progression despite increasingly intensive strategies. This highly heterogeneous clinical behavior of disease makes the prediction of each patient’s individual risk at the time of diagnosis the major goal to choose the most adequate therapeutic approach.

So far, the well-known tumor-specific genetic markers (amplification of the MYCN oncogene, DNA ploidy, chromosome 1p and 11q losses, and gains of chromosome 17q), patient age at diagnosis with two cut-offs of 12 and 18 months, the histopathological assessment (Shimada system), disease stage defined by the international neuroblastoma staging system (INSS) have been crucial determinants for the assignment of the risk class and the treatment plan. Currently, clinical trials stratify patients into four prognostic subgroups with expected very low risk, low risk, intermediate risk, and high risk of death from disease, and sixteen pretreatment designations (Cohn et al., 2009).

However, according to the Surveillance, Epidemiology, and End Results databases (www.seer.cancer.gov), the survival rates, during the period from 1999 through 2005, have improved only for patients with the more benign form of the disease, the rates among children with high-risk neuroblastoma have shown only modest improvement, despite dramatic escalations in the intensity of therapy provided. The incorporation in clinical decision of additional genetic prognostic markers, including gene copy number aberrations, common genetic polymorphisms as well as gene expression and epigenetic markers is
believed to be crucial in the development of a more accurate classification of patients in distinct prognostic risk categories and so in making treatment decisions.

The completion of the human genome sequence and the development of high throughput technology present exciting opportunities for the study of cancer. Recent technological advances based on whole genome approaches have made possible to discover a lot of new genetic markers in neuroblastoma which could be of substantial clinical importance in the relatively near future (Capasso and Diskin, 2010; Maris, 2010). Genome-wide analysis of amplification, deletion of genomic regions, somatically acquired genetic variations, common predisposing genetic variants, and mRNA expression profiles is a critical step to resolving the mechanisms of neuroblastoma tumorigenesis, might have implications towards better risk assessment and offer possibilities for target therapy. In this chapter, we will summarize the new discoveries related to the underlying molecular pathogenesis of this tumour with a special focus on advances that are translatable to the clinic. We will also present some recent, exciting findings from studies that have used a whole genome analysis technique such as comparative genomic hybridization (CGH), genome-wide association studies (GWAS), gene expression profiling (GEP) and next generation sequencing (NGS). We will finally discuss the clinical application of these whole genome searches and, the challenges and future directions for the genome-wide approaches to study the genetics of neuroblastoma.

2. Comparative Genomic Hybridization (CGH) of neuroblastoma

Genetic alternations such as amplifications and deletions frequently contribute to tumorigenesis. These alternations change the level of gene expression, which modify normal growth control and survival pathways. Characterization of these DNA copy-number changes is important for both the basic understanding of cancer and its diagnosis (Pinkel and Albertson, 2005). CGH was developed to survey DNA copy-number variations across a whole genome. With CGH, differentially labeled test (i.e. tumor) and reference (i.e. normal individual) genomic DNAs are cohybridized to normal metaphase chromosomes, and fluorescence ratios along the length of chromosomes provide a cytogenetic representation of the relative DNA copy-number variation. Chromosomal CGH resolution is limited to 10-20 Mb – therefore, anything smaller than that will not be detected. In the few years, diverse studies have used in-house bacterial artificial chromosome array or cDNA microarray to identify novel cancer related genes or crucial genome copy number alterations that determine distinct genetic subgroups for risk stratification of neuroblastoma (Mosse et al., 2005; Wang et al., 2006).

In a most recent paper, investigating 165 primary neuroblastomas, Caren and colleagues have compared the survival of patient subgroups defined by genomic alterations (Caren et al., 2010). They have demonstrated that patients with only numerical chromosomal aberrations and no other alteration had a favorable long-term outcome. On the other hand, the survival of patients characterized by MYCN amplification, loss of 11q or gain of 17q was considerably worse, whereas no death or disease was observed in patients with tumors harboring segmental chromosome alterations other than those previously mentioned. The main significance of these results is that a small number of predictive genomic alterations is sufficient for risk assessment of neuroblastoma patients.
Another study conducted by Janoueix-Lerosey et al. examined almost 500 primary neuroblastomas by bacterial artificial chromosome array-CGH and indicated that global genomic profiles may add significant prognostic information to current neuroblastoma risk estimation (Janoueix-Lerosey et al., 2009). The analysis showed that two distinct genetic classes of neuroblastoma might be related to two different mechanisms of instability. One, characterized by whole chromosome changes without any segmental alteration, is associated with an excellent outcome even in patients older than 18 months or with advanced stages of disease. The other class presents segmental chromosome imbalances that mainly arise through unbalanced chromosome translocations leading to concomitant loss and gain of chromosome fragments. Patients included in this last class, regardless the MYCN amplification, have an increased risk of relapse and in general they show a worse prognosis.

Another interesting CGH analysis has been performed on 236 primary neuroblastomas in Japan (Tomioka et al., 2008). The authors demonstrated that three major groups of genomic alterations in sporadic neuroblastomas could well define the prognoses of neuroblastomas:

- **GG-S (n=23):** a genetic group of silent with no obvious losses and gains except MYCN amplification (5-year cumulative survival rate: 68%; DNA ploidy: 87% diploidy);
- **GG-P (n=53):** a genetic group of partial chromosomal gains and/or losses (43% survival; 77% diploidy);
- **GG-W (n=36):** a genetic group of whole chromosomal gains and/or losses (80% survival; 22% diploidy).

In general, even if the results of these three excellent studies highlight the importance of prognostic classification of neuroblastoma using CGH, the clinical significance of global genomic alterations needs to be further evaluated in independent studies and compared with current risk estimation strategies.

Currently, the clinical significance of this advanced technology is monitoring cancer progression and distinguishing between mild and metastatic censorious lesions using FISH (Florescence in situ hybridization) probes on regions of recurrent copy number aberrations in several tumor types. It can be also used to reveal more regional copy number markers that can be used for cancer prediction. Finally, identifying and understanding the genes that are involved in cancer will help to design therapeutic drugs that target the dysfunction genes and/or avoid therapies that cause tumor resistance. However, one of the limitations of CGH analysis is due to the contamination of stromal cells that can affect the copy number detection. The routine application of CGH in clinical practice might be considerably limited by the issue of contaminating stromal cells, because defined thresholds of tumor content will a priori exclude a substantial fraction of samples from the analysis.

### 3. Gene Expression Profiling (GEP) of neuroblastoma

Microarray-based gene expression profiling enables the rapid, simultaneous measurement of gene expression on a genome-wide scale and consequently this technology represents an exciting development in genomic marker research that has been enthusiastically embraced by researchers worldwide.
Comparison between neuroblastoma subsets with good and poor prognosis, or based on neuronal functions has indicated diverse gene expression markers such as the genes whose mRNA expression is high in the favorable type of neuroblastoma include TRKA, CD44, pleiotrophin, N-cadherin, H-RAS, ECEL1, NLRR3, BMCC1, NEDL1 and ZNF423, and those in the unfavorable ones include TRKB, TERT, CDC10, NLRR1, HEN2 and LMO3 (Ohira and Nakagawara, 2010). There are evidences in literature that indicate the tyrosine kinase receptors as tractable therapeutic targets due to their cell surface localization and limited normal tissue distribution (Maris, 2005). A phase I trial of lestaurtinib, a multi-kinase inhibitor that strongly inhibits TrkB, has recently been completed in children with refractory neuroblastoma (Minturn et al., 2011). This study has identified a well-tolerated oral dose that results in effective TrkB inhibition, and its efficacy will be studied in future pediatric clinical trials. However, none of the above mentioned candidate genes alone could sufficiently explain the different biological behavior of neuroblastoma tumors, or could be considered as a prognostic marker in a clinical setting.

Global gene expression analysis can reveal significant information about the nature and state of cells, including disease progression, pharmacological response, and biological phenomena such as growth and development. These discoveries often point to sets of genes and gene pathways that are specifically associated with the state or change in state induced by disease. It also leads to the identification of gene sets associated with disease outcome and prognosis, as well as drug-induced responses. In many cases, only a small number of genes in the total genome (e.g., 25–500 genes) respond in a statistically significant way. The set can be further reduced by removing redundant genes. The resulting smaller subset becomes a true signature for diagnosis of the biological state and a predictor of therapeutic response. The identification of gene expression signatures is proving to be a powerful tool for disease diagnosis and drug discovery. For example, in oncology, microarray analysis is now being broadly applied toward the diagnosis and classification of a host of different cancers, including breast (van ’t Veer et al., 2002), kidney (Vasselli et al., 2003), prostate (Best et al., 2003), and childhood cancers (Khan et al., 2001). Microarray analysis has also been used to categorize chemotherapy response and potentially direct the course of chemotherapy, for example, in acute myeloid leukemia (Okutsu et al., 2002).

In the past few years, several reports have provided irrefutable, evidence that specific gene-expression patterns can predict the natural courses of neuroblastoma patients with high accuracy. In the 2004, the first study showed a gene-expression-based classifier able to predict the prognosis of patients efficiently by profiling 56 tumors using cDNA microarrays containing 42 578 cDNA clones (classification accuracy was 95% for 21 test samples) (Wei et al., 2004). However, this study shows important limitations that prevented a diagnostic application of the proposed classifier: the low number of cases analyzed and the lack of validation analyses in independent sets of neuroblastomas. Since this first report on gene-expression-based risk classification, there have been multiple attempts to define and validate an optimal mRNA prognostic expression signature in neuroblastoma.

Ohira and colleagues, in the 2005, published a prognostic gene-expression signature for neuroblastoma patients that was derived from a large cohort of 136 primary neuroblastomas diagnosed in Japan (Ohira et al., 2005). They utilized an algorithm calculating survival probabilities of each patient at 2 years or 5 years after diagnosis, which is indicated by a “posterior” value range from 0 to 1. The gene expression signature was able to predict
patient prognosis with high efficiency (90% accuracy, 96% sensitivity and 90% specificity). Of note, an intermediate-risk group (stage 3 or 4 without MYCN amplification) was classified with an accuracy of 86% better than the current clinical markers (age, stage and MYCN) that exhibited 64% accuracy. The authors also made a “mini-chip” carrying the top-ranked 200 genes for clinical use. It was tested on 50 independent tumor samples of all stages indicating high reproducibility as well as high efficiency. This ‘mini-chip’ is now under clinical validation in a larger cohort in Japan.

In the 2006, Asgharzadeh et al. have built a classifier comprising 55 genes using 102 patients with metastatic, MYCN non-amplified neuroblastoma, that was able to separate patients into two groups with statistically different event-free survival rates (Asgharzadeh et al., 2006). Regardless of MYCN status, all patients 18 months or older with metastatic neuroblastoma are currently considered high risk, and receive intensive therapies. The use of a molecular risk classification based on each tumor’s gene expression signature may be able to identify a subgroup with less aggressive disease that may not require intensive, high-risk therapy. However, the 55-gene signature determined by Asgharzadeh et al. has not been validated in independent data sets, no trial for the this sub-group of patients is prospectively assessing the use of the 55-gene signature, so its clinical application is still far.

Oberthuer et al., in the 2006, published a gene signature including 144-genes able to improve the risk estimation of current neuroblastoma trials (Oberthuer et al., 2006). They utilized a customized oligonucleotide microarray comprising 10,163 probes. The 144-gene signature obtained by analyzing 77 samples could classify 174 patients more accurately than risk stratification of current trials from Germany, Japan and the United States. To further evaluate the impact of this 144-gene expression-based classifier on the neuroblastoma risk classification, Oberthuer et al., in the 2010, determined its prognostic value in a validation cohort of 440 internationally collected neuroblastoma specimens, 125 of which were examined prospectively (Oberthuer et al., 2010). The results demonstrated that the 144-gene classifier improved the accuracy of discrimination of true low-risk patients who may need no chemotherapy at all from patient’s aggressive tumors who require intense cytotoxic treatment. The main limitation of these two studies is that a large number of genes has been included in the classifier and this might negatively influence its translation to the clinical routine as the costs to perform a gene expression analysis of 144 genes will be likely high. Even if the high number of genes analyzed can balance the outlying values of individual transcripts, and thus yield more stable classification results, the future directions for this kind of study will be construct small gene signatures with high accuracy and specificity of outcome prediction. This study strategy can more quickly facilitate the introduction of gene-based classifiers into the clinical routine use.

Another interesting gene signature was recently built (De Preter et al., 2010; Vermeulen et al., 2009). Based on an innovative data-mining strategy, the authors identified a 59-signature that was built using 30 training samples, tested on 313 test samples, and subsequently validated in a blind study on an independent set of 236 tumors. This gene signature can act as an independent risk predictor enabling the identification of patients with increased risk in the current European treatment group. This 59-gene classifier shows three important advantages compared to the previously published gene expression classifiers: the small amounts of starting material, the low number of genes, high cost-efficiency and speed of the quantification method (real-time quantitative polymerase chain reaction) and the possibility
of cross lab data comparison. This gene expression based classifier mostly responds to criteria for facing future studies to identify gene signatures that may be rapidly translated into the clinical management of neuroblastoma.

In one study published in the 2010, (Fardin et al., 2010) and co-workers investigated the prognostic potential of hypoxia induced genes in neuroblastoma tumors. Hypoxia, a condition of low oxygen tension occurring in poorly vascularized areas, has profound effects on tumor cell growth, susceptibility to apoptosis, and resistance to radio- and chemotherapy. A biology-driven approach was chosen to define the hypoxia signature. The authors obtained a 62 probsets neuroblastoma hypoxia signature (NB-hypo) by performing a systematic analysis of the transcriptome of neuroblastoma cell lines cultured under hypoxic or normoxic conditions. NB-hypo signature efficiency were tested on 88 neuroblastomas, it resulted be an independent prognostic factor for neuroblastoma. This study also supports the view that hypoxia is negatively correlated with tumors' outcome. However the performance of this signature has been assessed only in a modest set of patients and of course it needs to be tested in diverse and larger cohorts of neuroblastoma patients.

3.1 Gene expression profiles of non-coding RNA

MicroRNA (miRNA), a class of small non-coding RNA that regulates gene expression at a post-trascriptional level, has been studied in the context of cancer, and promising miRNA biomarkers have been identified for numerous major cancer types. miRNA are molecules comprising 22-25 nucleotide and function as negative regulators of gene expression. Literature data show that miRNAs can have a role in neuroblastoma and their expression correlates to prognosis, diagnosis and response to treatment. A recent whole genome search for MYCN target miRNA promoters differentially repressed under high MYCN conditions has indicated two miRs, miR-591 and miR-558, as potent tumor suppressors (Shohet et al., 2011). An other interesting study, conducted by Lin and co-workers, has analyzed 66 neuroblastoma tumors using a plausible neural networks to select a combination of 15 biomarkers that consist of 12 miRNAs' signature, expression levels of Dicer and Drosha (miRNA processing enzymes that are required for the maturation of miRNAs), and age at diagnosis, were able to segregate all patients into four distinct patterns that were highly predictive of clinical outcome (Lin et al., 2010).

4. Genome Wide Association Studies (GWAS) of neuroblastoma

GWAS is indicated as an unbiased approach to revealing the risk alleles for genetically complex non-Mendelian disorders. One important characteristic of this advance technology is rapidly scanning markers across the complete sets of DNA, or genomes, of many people to find genetic variations associated with a particular disease. Indeed, during the past few years, this methodological approach has identified several affordable associations between specific chromosomal loci and common complex human disease. So far, GAWS have identified more than 200 new common low-penetrance susceptibility loci for cancers (NHGRI GWAS Catalog; http://www.genome.gov/gwastudies/). The discovery of new genetic associations is particularly useful in determining the genetic risk profile for a specific disease and so to develop better strategies to detect, treat and prevent the disease.
Moreover, GWAS analysis is very useful in identifying new disease genes and so to understand the functional consequences of these genes.

Recently, by using a GWAS approach, we have identified multiple genetic risk factors associated with the neuroblastoma development and its clinical phenotypes (Table 1). Our data unequivocally demonstrate that multiple common DNA variations cooperate to increase the risk for neuroblastic malignant transformation. In the 2008, we identified the first predisposition locus by genotyping over 550,000 germline DNA SNPs from 1,752 neuroblastoma pantients and 4,171 controls of European descent (Maris et al., 2008). Three common SNPs located in intronic region of the predicted genes LINC00340 (chromosome 6p22) were strongly associated with the development of sporadic neuroblastoma. Of particular interest was the finding that not only were the three 6p22 SNPs associated with the likelihood of developing neuroblastoma, but patients who carried the 6p22 risk alleles were more likely to develop the clinically aggressive form of the disease and suffer tumor recurrence (Maris et al., 2008). We therefore performed a second genetic analysis, this time limiting the cases to those patients with high-risk neuroblastoma; we were able to identify a

<table>
<thead>
<tr>
<th>Gene</th>
<th>Most Associated genetic variant</th>
<th>Type</th>
<th>Location</th>
<th>Risk allele</th>
<th>Estimated odds ratio</th>
<th>P-value</th>
<th>Protein function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINC00340</td>
<td>rs6939340</td>
<td>SNP</td>
<td>6p22</td>
<td>A</td>
<td>1.37</td>
<td>9.33x10^{-15}</td>
<td>Long intergenic non-protein coding RNA 340</td>
<td>Maris et al., 2008</td>
</tr>
<tr>
<td>FLJ22536</td>
<td>rs6435862</td>
<td>SNP</td>
<td>2q35</td>
<td>G</td>
<td>1.68</td>
<td>8.65x10^{-18}</td>
<td>Binding partner of BRCA1; BRCA1–BARD1 heterodimer has tumour-suppressor functions; splice variant BARD1 has oncogenic functions</td>
<td>Capasso et al., 2009</td>
</tr>
<tr>
<td>BARD1</td>
<td>rs110419</td>
<td>SNP</td>
<td>11p15</td>
<td>A</td>
<td>1.34</td>
<td>5.2x10^{-18}</td>
<td>Cysteine-rich transcriptional regulator with two LIM zinc-binding domains</td>
<td>Wang et al., 2011</td>
</tr>
<tr>
<td>LMO1</td>
<td>rs1027702</td>
<td>SNP</td>
<td>1q23</td>
<td>T</td>
<td>2.01</td>
<td>3.38x10^{-06}</td>
<td>Member of the dual specificity protein phosphatase subfamily</td>
<td>Nguyen le et al., 2011</td>
</tr>
<tr>
<td>DUSP12</td>
<td>rs2619046</td>
<td>SNP</td>
<td>5q11</td>
<td>T</td>
<td>1.48</td>
<td>5.70x10^{-05}</td>
<td>DEAD box proteins, characterized by the conserved motif Asp-Glu-Ala-Asp (DEAD), are putative RNA helicases</td>
<td>Nguyen le et al., 2011</td>
</tr>
<tr>
<td>DDX4</td>
<td>rs10055201</td>
<td>SNP</td>
<td>5q11</td>
<td>A</td>
<td>1.49</td>
<td>3.85x10^{-05}</td>
<td>Type I cytokine receptor family</td>
<td>Nguyen le et al., 2011</td>
</tr>
<tr>
<td>IL31RA</td>
<td>rs11579261- rs3853524</td>
<td>CNV</td>
<td>1q21</td>
<td>-</td>
<td>2.49</td>
<td>2.97x10^{-17}</td>
<td>Multifunctional isoenzyme functional in the conversion of estrone to estradiol (E2), and elongation of long-chain fatty acids</td>
<td>Nguyen le et al., 2011</td>
</tr>
<tr>
<td>HSD17B12</td>
<td>rs11037575</td>
<td>SNP</td>
<td>11p11</td>
<td>C</td>
<td>1.67</td>
<td>1.07x10^{-06}</td>
<td>Novel member of the NBPF (neuroblastoma breakpoint family)</td>
<td>Diskin et al., 2009</td>
</tr>
</tbody>
</table>

*Official full name provided by HUGO Gene Nomenclature Committee (HGNC)
'Start point SNP and end point SNP of the CN
Odds Ratios and P-values as reported in the corresponding published research article
Abbrevations: SNP, single nucleotide polymorphism; copy number variation

Table 1. Neuroblastoma susceptibility genes and variants identified by GWAS
new association with common intronic and nonsynonymous SNPs at the BARD1 tumor-suppressor gene (Capasso et al., 2009). In this GWAS, we compared the allelic frequencies between more than 500 high-risk patients and 2043 controls, six SNPs at chromosome 2q35 located within introns 1, 3 and 4 of BARD1 reached the genome-wide significant level. Furthermore, by TaqMan based genotyping in 540 high-risk neuroblastoma cases and 1,142 controls, we demonstrated that also four nonsynonymous SNPs were genetic risk factors for the susceptibility to the development of high-risk neuroblastoma. BARD1 has been previously implicated in several cancers due to its association with BRCA1, a well-known breast cancer susceptibility gene. BARD1 heterodimerizes with BRCA1 and is thought to be necessary for the tumor suppressive function of BRCA1. Studies are ongoing to understand how sequence variations within BARD1 influence neuroblastoma tumorigenesis.

Most recently, increasing of the number of cases and controls of the original GWAS cohort allowed us to identify a further predisposition locus at 11p15, within LIM domain only 1 (LMO1) gene, a transcriptional regulator (using 1,627 neuroblastoma patients and 3,254 controls) (Wang et al., 2011). Two SNPs were found to be associated with neuroblastoma in general and with high-risk neuroblastoma. Additional genomic analyses of DNA from 701 primary neuroblastoma tumors demonstrated that a duplication of the LMO1, presents in 12% of tumors, resulted in increased expression of LMO1 in cell lines and tumors and was associated with more advanced disease. Further functional experiments indicated LMO1 as an oncogene involved in the pathogenesis of neuroblastoma.

To confirm the validity of these findings beyond the Caucasian population, an association test was performed by using the genotype data from 326 self-reported African-American neuroblastoma patients and 2500 African-American unaffected children. Devoto and co-workers showed the first preliminary results at 60th annual meeting of The American Society of Human Genetics in November 2010. Among the known loci, the one that was most consistently replicated was BARD1, with six SNPs showing p-value less than 0.05 (smallest p-value equal to 0.0004 for rs7587476). Two SNPs in the LINC00340 region had nominally significant p-value, and only one in LMO1. Overall these results support the hypothesis that the same genetic factors affect risk of neuroblastoma in both populations.

An other published article that utilized a GWAS approach to detect allelic risk variants suggested again that neuroblastoma is both phenotypically and genetically heterogeneous. Indeed, an analysis restricted to low-risk neuroblastoma cases identified predisposing loci for less aggressive disease within four additional genes: DUSP12 at 1q23.3, DDX4 and IL21RA both at 5q11.2 and HSD17B12 at 11p11.2 (Nguyen le et al., 2011). Interestingly, the same risk alleles resulted to be not associated with high-risk neuroblastoma. Furthermore, a single SNP replication analysis in an Italian cohort of 115 low-risk cases and 680 controls showed that the three most significant SNPs (rs1027702, rs2619046, rs11037575) in the three loci that contain DUSP12, DDX4/IL31RA, and HSD17B12, respectively were confirmed to be genetic risk factors for the susceptibility to the development of low-risk neuroblastoma. Together, these results suggest that genetic initiating events may predispose not only to neuroblastoma, but to clinically relevant sub-phenotypes as well.

We are currently adding new cases and controls to our data sets for both United States population and Italian population. Moreover, we are utilizing alternative strategies for finding those genetic risk loci hidden among signals discarded by multiple testing
corrections needed in the analysis of GWAS data. Together these two approaches will allow us to identify new genetic variants associated with neuroblastoma.

The next few years will see the completion and publication of additional genome-wide scans that involve higher resolution tools and substantially larger DNA collections than have been used before. Even when performed with exquisite resolution, however, association and epidemiology studies are limited in their power to prove causation. Thus, effective functional studies will be required that connect genetic variation with disease pathophysiology. Furthermore, while GWAS are a valuable tool for assessment of the effects of common polymorphisms, the aggregate role of low-frequency rare functional gene variants in neuroblastoma has not been properly evaluated. Uncommon variants with relatively large effects might account for part of the unexplained heritability in neuroblastoma. Moreover, some of these rare variants might reside in genes previously identified in GWAS. Rare genetic variation is not well represented in current databases or commercially available fixed arrays, and can be only ascertained by targeted deep sequencing of samples from affected individuals (Nejentsev et al., 2009).

4.1 Copy Number Variations (CNVs) based GWAS

In addition to SNP genotypes, CNVs (deletion or duplication of segments of germline DNA) represent a significant source of genetic diversity that may influence disease susceptibility. An analysis of copy number at 550,000 SNPs in germline DNA from 1,441 neuroblastoma cases and 4,160 controls identified a common deletion polymorphism spanning less than 145-Kb at 1q21.1 associated with neuroblastoma, no duplications reached genome-wide significance (Diskin et al., 2009). A novel member of the NBPF (“neuroblastoma breakpoint family”) gene family mapping within the CNV was cloned and sequenced. Expression of this transcript, termed NBPF23, was found to be significantly correlated with the underlying CNV genotype in neuroblastoma tumors and cell lines, further supporting the biological relevance of the CNV association (Diskin et al., 2009). The highest levels of NBPF23 expression were seen in fetal brain and sympathetic ganglia and expression was directly correlated with CNV status in neuroblastoma cells, implicating NBPF23 in the molecular pathogenesis of neuroblastoma. Further CNVs-based GWAS analyses are ongoing and it is likely that other CNVs affect the neuroblastoma development.

The results from GWAS are clearly telling us that common, heritable genetic variants contribute to the predisposition for sporadic neuroblastoma (Figure 1). We now can state that the neuroblastoma belongs to the group of complex genetic diseases, which are common disorders characterized by modest disease-risk heritability and multifaceted interactions between genetic and environmental factors. The full list of disease genes (susceptibility and modifiers) and environmental triggers in neuroblastoma remains incomplete (Figure 1). Moreover these studies are giving an important contribution in understanding the complex genetic architecture of neuroblastoma and in providing valuable insight into the underlying tumor biology. Identifying the genes and pathways involved in predisposing individuals to neuroblastoma may also lead us to the development of novel treatments. On the other hand, we can not state that there is an immediate clinical utility of knowing an individual genotype at these disease variants. However it is plausible to prefigure that the discovery of many others single disease risk variants, disease risk gene-gene and gene-environment interactions together with the development of new statistical
methods (able to construct reliable disease predictor) will make feasible the clinical application of these GWAS findings in the relatively near future. Indeed, predictive modeling might be substantially created if more-penetrant variants are identified, and transcriptional signatures and accurate environmental exposures are taken into account. The translational potential of such modeling might extend to redefining disease classification, prognosticating progression, and predicting treatment responses and/or adverse effects.

![Genetic model of neuroblastoma](https://www.intechopen.com)

**Fig. 1.** Genetic model of neuroblastoma. Four scenarios in which gene-gene and gene-environmental interactions play a crucial role might characterize the model of genetic susceptibility to neuroblastoma. (A) Both genetic and environmental components do not exceed so that there is not malignant transformation in patients without ALK or PHOX2B mutations. (B) There is an excess of the environmental component, which induces the malignant transformation in patients without ALK or PHOX2B mutations. (C) Multiple common DNA variations (excess of the genetic component) in a large number of genes can cooperate to reach the threshold malignant transformation in patients without ALK or PHOX2B mutations. (D) A mutation in the ALK or PHOX2B gene results in a single, highly penetrant risk allele that interacting with environmental factors allows a malignant transformation.
5. Genome-wide linkage analysis of familial neuroblastoma

An other important genome-wide approach, utilizing high density SNP-based microarray technology, made possible the identification of three separate missense mutations within coding exons of the anaplastic lymphoma kinase (ALK) gene which were responsible for familial neuroblastoma (Mosse et al., 2008). While almost 99% of neuroblastoma occurs sporadically the remaining 1-2% of neuroblastoma has a familial history (Maris, 2010). This genome-wide approach was utilized in an attempt to identify linkage signals that might identify predisposition loci for the majority of hereditary neuroblastoma pedigrees, which do not harbor germline PHOX2B mutations (explaining only a small subset of cases). Six thousands SNPs were genotyped in 176 individuals from 20 families with hereditary neuroblastoma. ALK gene which maps to 2p23 is a receptor protein-tyrosine kinase which functions as an oncogene in many human cancers, most notably through translocations resulting in constitutive activation of the ALK kinase domain as seen in anaplastic large cell lymphomas (Morris et al., 1994), inflammatory myofibroblastic tumors (Griffin et al., 1999), squamous cell carcinomas (Jazii et al., 2006), and non-small cell lung cancers (Soda et al., 2007). The ALK and PHOX2B mutations are responsible for approximately 90% of hereditary neuroblastoma. Based on these data a genetic testing could be created for facilitating the identification of unaffected siblings who carry highly penetrant germline mutations and would justify screening in an attempt to ensure early detection of neuroblastoma. By using an exome high-throughput sequencing, researchers are now attempting to find the heritable genetic etiology in the remaining rare families with evidence of hereditary neuroblastoma that do not harbor ALK or PHOX2B germline mutations.

Of note, somatic ALK mutations or amplifications were also identified in 6-9% of sporadic cases (Chen et al., 2008; George et al., 2008; Janoueix-Lerosey et al., 2008). According to the recent success of small molecule tyrosine kinase inhibitors in a certain subset of cancers, such as gefitinib in non-small-cell lung carcinoma with epidermal growth factor receptor (EGFR) mutations, a similar therapeutic approach based on inhibition of ALK-mediated signaling will be expected to target oncogenic ALK mutations in neuroblastoma (Li and Morris, 2008).

In a recent study a novel activating mutation in ALK (F1174) in a neuroblastoma patient in the course of analysis of genomic DNA from patient biopsy samples has been identified (Martinsson et al., 2011). The appearance of this novel F1174S ALK mutant correlated with the development of aggressive disease at the patient level with emergence of therapy resistance and fatal outcome. This study points out an important issue regarding the possibility of the occurrence of genomic mutation events during the disease outcome. So, the initial screening in the first tumor biopsy of a patient is not sufficient and further molecular analyses of the ALK locus, in particular in tumor progression and/or tumor relapse, is warranted for better understanding of the treatment of neuroblastoma patients.

6. Next Generation Sequencing (NGS) of neuroblastoma

Over the past four years, there has been a fundamental shift away from the application of automated Sanger sequencing for genome analysis. The automated Sanger method is now
considered as a ‘first-generation’ technology, and newer methods are referred to as next generation sequencing. This sequencing method that has emerged since 2005 parallelizes the sequencing process and produces millions of typically short sequence reads (50–400 bases) from amplified DNA clones. In general, the field of cancer genomics has been impacted most profoundly by the application of next generation sequencing technology, which has tremendously accelerated the pace of discovery while dramatically reducing the cost of data production. Hence, there has been a rapid progression from targeted gene re-sequencing using PCR and Sanger sequencing to either targeted, whole genome, or whole transcriptome sequencing using these massively parallel sequencing platforms, coupled with the requisite bioinformatics-based approaches to analyze the data. There are particular challenges for the detection and diagnosis of cancer genome alterations. For example, some genomic alterations in cancer are prevalent at a low frequency in clinical samples, often owing to substantial admixture with non-malignant cells. Second-generation sequencing can solve such problems (Thomas et al., 2006). Furthermore, these new sequencing methods make it feasible to discover novel chromosomal rearrangements (Campbell et al., 2008) and microbial infections (MacConaill and Meyerson, 2008) and to resolve copy number alterations at very high resolution (Campbell et al., 2008; Chiang et al., 2009).

So far, only a few studies have utilized this advanced methodology to search genetic alterations responsible of the neuroblastoma initiation. To identify potential drug targets against recurrent neuroblastoma, Morozova and colleagues used next generation RNA sequencing and/or human exon arrays to profile the transcriptomes of 11 tumor-initiating cells from six neuroblastoma patients, revealing genes that are highly expressed in the tumor-initiating cells compared with normal neural crest-like cells and unrelated cancer tissues (Morozova et al., 2010). This analysis revealed 30 targets with an available inhibitor, six of which have never been implicated in neuroblastoma. Further validation analyses were performed on one of the six novel drug targets, AURKB. Interestingly, a selective AURKB inhibitor, AZD1152, was cytotoxic to neuroblastoma tumor-initiating cells used in the study but not to normal pediatric neural crest-like precursor cells. This is the first report of AURKB inhibitors as potential therapeutics for neuroblastoma. However, AURKB inhibitors are currently in clinical trials so it is reasonable to forecast a rapid translation of this observation to neuroblastoma therapy.

Another recently published study shows that miRNA-based classifiers can be used to stratify neuroblastoma patients according to clinical course (Schulte et al., 2010). The small RNA transcriptomes of five favorable and five unfavorable neuroblastomas were analyzed using SOLiD next-generation sequencing, generating a total of more than 188,000,000 reads. Favorable and unfavorable neuroblastomas were distinguishable by hierarchical clustering of miRNA patterns. Expression of single miRNAs also differed significantly between the two groups. The authors then identified 13 candidates for novel miRNAs of which three were further validated in 70 primary neuroblastomas using RT-qPCR.

However, in literature, there is no published research article showing the use of next generation sequencing to search for somatically acquired DNA mutations in sporadic
neuroblastoma. The most significant impact of next generation sequencing on cancer genomics has been the ability to re-sequence, analyze and compare the matched tumor and normal genomes of a single patient. Many studies have identified several mutations responsible of the development of diverse cancers (Bass et al., 2011; Puente et al.; Wei et al., 2011). About neuroblastoma, a large collaborative effort, the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project (http://target.cancer.gov), is leveraging genomic profiling techniques to help identify the molecular mechanisms that drive oncogenesis and to discover new therapeutic targets. The first results have been shown at 102nd Annual Meeting of American Association for Cancer Research in April 2011. So far, the exome sequencing of 81 tumor/normal pairs from high-risk neuroblastoma patients has been completed. Pugh and colleagues reported that several genes resulted significantly mutated including ALK (6 tumors) and PTPN11 (3 tumors). Interestingly, there were 27 other genes with candidate mutation frequencies in the 4-10% range, but no single gene was identified with a mutation frequency greater than 10%. Morozova and co-authors instead presented a study about a whole genome shotgun sequencing of six stage 4 MYCN-non-amplified and four stage 4 MYCN-amplified neuroblastomas cases and matched peripheral blood, as well as whole transcriptome sequencing of the corresponding tumor RNA. The analysis of tumor and normal genomes revealed an average of 1664 candidate somatic mutations per case, the majority of which were single nucleotide variants and small indels falling within introns or intergenic sequence. Moreover, an average of 10% of candidate somatic mutations in coding sequence was expressed in the transcriptome representing candidate oncogenic events. Of note, two novel somatic gene fusions, between TRIM37 on chromosome 17 and RNF121 on chromosome 11, and between LSAMP and STAG1 on chromosome 3, previously uncharacterized in neuroblastoma, were identified.

It is likely that the biggest impact of second-generation sequencing of cancer genomes will be in cancer diagnostics. The major challenge will be to make biological sense of the mountains of genomic data. This will require computational, biological and clinical analyses of the genome data. These developments might potentiate accurate genome-based diagnosis for patients with neuroblastoma. Moreover, this advanced technology will rapidly identify therapeutic targets that will have implications in developing new and effective treatments for this devastating disease.

7. Whole genome searches to open the door to the personalized medicine

Recent advances in genomics are likely to change the molecular characterization of cancer and provide a path for the personalized treatment of patients with cancer. The generation of comprehensive catalogs of genetic alterations could inform strategies for effective prevention and early detection, as well as guiding the development of therapeutics in the appropriate patient subpopulation. Genome searches also hold great potential to inform the prognosis and guide evidence-based management of early stage diseases, which comprise an increasing proportion of cancer diagnoses. It is likely that whole genome searches will drive the transition from a morphology-based to a genetics-based taxonomy of cancer, point of care decisions will become increasingly customized to the unique genomic and proteomic features of a patient’s tumor. These molecular changes might be evaluated in circulating
tumor cells or in tumor biopsies or in body fluids (plasma, lymph and ascites) using specific diagnostic tests that can predict the probability of clinical benefit for all treatment options.

Genome searches have provided an important contribution in the risk stratification of neuroblastoma. To date, the International Neuroblastoma Risk Group (INRG) Task Force has assessed several potential prognostic factors in 8800 patients. These factors included stage, age, lactate dehydrogenase, as well as certain genetic markers such as MYCN amplification, chromosome 1p, 17q, and 11q aberration, and ploidy. After appropriate statistical analyses, four risk groups were established and the current consensus scheme includes stage, age, histologic category, grade of tumor differentiation, MYCN status, 11q aberration, and tumor cell ploidy as final criteria (Cohn et al., 2009). According to the INRG Task Force, these four risk groups were designed to stratify children into more homogenous pretreatment groups. Despite these efforts, more needs to be done to decrease the death rates of children with neuroblastoma and the upcoming whole genome searches might play a main role in both improving risk stratification and developing more appropriate treatments. Indeed, genome-wide analyses have already identified diverse therapeutic targets and biological agents are currently tested. They include small molecule inhibitors of genetic pathways implicated in tumor growth, such as inhibitor of the phosphoinositide 3-kinase/mammalian target of rapamycin pathway; insulin-like growth factor 1 receptor; ALK; Aurora kinase A and B; tyrosine receptor kinase, and histone deacetylase inhibitors. It is plausible that by the inclusion of biological end points in large cooperative studies, further insight will be gained about the underlying pathways in the tumorigenesis of neuroblastoma, which will lead to new therapies to improve outcome.

8. Future directions

Since the publication of the Human Genome Project data, medical practice and research have entered the genome era. From a predominant focus on single genes and disease-related mutations has emerged the technology to assay the genome in its entirety and the first means to interpret findings in a comprehensive manner. Several genome-wide approaches have been used to study neuroblastoma and several information about molecular mechanisms underlying the neuroblastic malignant transformation, and genomic or genetic abnormalities associated with neuroblastoma and its clinical phenotypes have been obtained. This allows us to understand how to ameliorate therapeutic treatments of neuroblastoma and how to optimize risk group classification. However, despite extensive research, relatively few genomic markers have been implemented into routine clinical use and the survival rate for neuroblastoma high-risk patients is still low (~40%). Integrative genomic approaches incorporating data from high-throughput genomic technologies are needed, on the one hand, to facilitate our understanding of neuroblastoma pathogenesis and consequently to accelerate the development of more effective therapies, on the other hand, to improve risk estimation of children with neuroblastoma (Figure 2). To give strong push for the achievement of these two objectives, inter and intra-disciplinary collaborations of biologists, chemists, clinicians, computer scientists, and mathematicians are strongly warranted. Furthermore, sharing and availability on public depositories of scientific results from neuroblastoma whole genome studies and in general from cancer studies are very important start points to get the main aim of research work: increasing the percentage of children who beat this devastating disease.
Fig. 2. Integrated genome-wide approaches in neuroblastoma. The integration of diverse genome-wide-based analyses will provide a large amount of biological and genetic information that presumably will determinate the identification of individual predisposition to the clinical sub-groups of neuroblastoma, the refinement of risk stratification and development of new and effective cures. Overall the integrated application of these approaches will lead us to improve the survival of children with high-risk neuroblastoma.

9. References


colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. Nat Genet 43, 964-968.


Contemporary Pediatrics
Edited by Dr. Öner Özdemir

Hard cover, 434 pages
Publisher InTech
Published online 21, March, 2012
Published in print edition March, 2012

Book Contemporary Pediatrics with its 17 chapters will help get us and patients enlightened with the new developments on the contemporary pediatric issues. In this book volume, beyond classical themes, a different approach was made to current pediatric issues and topics. This volume, as understood from its title, describes nutritional infant health and some interesting topics from pediatric subspecialties such as cardiology, hemato-oncology and infectious diseases.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
