On the molecular genetic etiology of osteosarcoma

Wilkens, Martje

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2006

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 5

Summary, General discussion and
Future perspectives
Chapter 5

Summary

Osteosarcoma is a malignant tumour of bone mostly arising in young children and adolescents. The incidence of osteosarcoma is 1.7 per million in individuals younger than 10 years of age and 8.2 per million in the age group between 10 to 19 years of age. In the Netherlands these incidences are very similar. The 5-year overall survival rate of 20% in 1970 changed to the current 55% to 70% with neoadjuvant chemotherapy and surgery. Today the limbsalvage can be achieved in over three quarter of the patients.

Osteosarcoma can be found as part of some rare inherited (cancer) syndromes, such as hereditary retinoblastoma, Li-Fraumeni syndrome, Paget disease of bone, Rothmund-Thomsom syndrome, RAPADALINO syndrome and Werner syndrome or as sporadic osteosarcoma. Sporadic osteosarcoma is far more frequent than the inherited forms of osteosarcoma and it is this sporadic form of osteosarcoma that is the focus of this research.

It is believed that osteosarcoma is caused by multiple genetic aberrations, however, the precise combination of mutated genes is yet unknown. Several genes known to be involved in the above-mentioned inherited syndromes in which osteosarcoma can occur, have been found mutated in sporadic cases. Furthermore, in several studies sporadic primary osteosarcoma tumours and cell lines derived from osteosarcoma have been cytogenetically characterised using different techniques. Most studies found that primary osteosarcoma tumours and osteosarcoma cell lines had losses in common at 2q31.1 3p12-p14, 4p16.2, 6q12, 6q21, 7q35, 10p15.1, 10q22-q23, 11q25, 13q12.2, 13q14.3, 17p13.1, 17q21, 18q12, 18q21-q22 and 20; and gains at 1p22-p31, 1p21-q24, 1q25-q31, 4p16, 8q21, 9q24, 12q13-q15, 14q24-qter, 16p13, 17p11-p12, 19p13 and 21q22. For two chromosomal regions, 1p22 and 8q23-q24, high-level amplification was detected in both primary osteosarcoma tumours and osteosarcoma cell lines. As mentioned, multiple (aberrant) genes appear to be involved in the development of sporadic osteosarcoma. Those that are known are reviewed in Chapter 1. To obtain more insight into the molecular genetic changes underlying the development of osteosarcoma, gene expression profiling of cultured osteoblasts, osteosarcoma cell lines and primary osteosarcoma tumour tissue has been performed as well as array-CGH analysis of the osteosarcoma tumour tissue.

In Chapter 2 the isolation and molecular characterisation of osteoblasts, the precursor cell of osteosarcoma, is described. Osteoblasts are cultured out of bone fragments, but inadequate pre-processing of these bone fragments may also result in
culturing fibroblasts. Since morphologically the cultured osteoblasts cannot be
discriminated from fibroblasts, we developed a test to discriminate both cell types
from each other. Gene expression profiles of 10 osteoblast and 11 skin fibroblast
cultures were compared using a 21,000-oligonucleotide micro-array.
This comparison resulted in many differentially expressed genes, 42 of which
showed a significantly different expression between both types of cultured cells.
These differentially expressed genes were specifically related to membrane
proteins/processes such as cell communication, cell adhesion and receptor activity.
For six of these genes expression differences defined by micro-array analysis were
confirmed by an independent method, real-time RT-PCR. VCAM1, KIAA1644,
FGFR2 and COL27A1 had a higher expression in osteoblasts, whereas IMP-3 and
MME were higher expressed in fibroblasts. Since for all six genes differences
between all pairs of individual cultures were fully consistent, we consider expression
analysis of these genes as a fast and efficient test to discriminate between cultured
osteoblasts and fibroblasts.
Chapter 3 describes differences in expression profiles of 5 primary osteoblast
cultures and 5 osteoblastic osteosarcoma cell lines by micro-array analysis. This
analysis resulted in 286 genes significantly differentially expressed. Functional
annotation of these genes revealed a lower expression in osteosarcoma cell lines of
genes encoding proteins associated with the cell membrane and extracellular matrix
and a higher expression of genes encoding nuclear proteins, such as cell cycle-
related and DNA repair-associated proteins. Furthermore, regional expression biases
in the genome of the osteosarcoma cell lines were found with an on average lower
expression of the MMP gene cluster at 11q22.2 and of the KAP gene cluster at
17q21.2. An on average higher expression was found for the HOXB cluster at
17q21.3. Such co-expression of genes points to either chromosomal losses in
tumours and tumour cell lines or to a common transcriptional regulation. To answer
the question whether the lower expression in osteosarcoma cell lines of genes
encoding membrane and extracellular matrix-associated proteins and the higher
expression of genes encoding nucleus-associated proteins should be considered as
cell line-specific or as tumour-specific, an additional comparison with gene
expression profiles of primary osteosarcoma tumours was needed.
Therefore, in Chapter 4, array-CGH analysis and gene expression profiling of three
osteoblastic osteosarcoma tumours and three fibroblastic osteosarcoma tumours
were performed. Gene expression profiles of all six primary osteosarcoma tumours
were compared to those of five osteosarcoma cell lines and five osteoblast cultures.
In total, 838 genes were significantly differentially expressed between any two of these groups of specimens. Functional analysis revealed that differences between primary tumour tissue on one hand and cultured osteoblasts on the other were mainly due to genes encoding proteins associated with inflammatory processes and angiogenesis. Genes encoding proteins involved in cell cycle regulation were expressed most prominently in the osteosarcoma cell lines. Regional repression of elevated expression in the genome for genes belonging to the MMP gene cluster at 11q22.2 was also identified in the osteosarcoma tumours. When genomic DNA from the six-osteosarcoma tumours was compared with normal human genomic DNA by array-CGH analysis, an under-representation of 11p was found only in the three osteoblastic tumours. This makes it more likely that this cluster of differentially expressed genes on 11q22 is due to a regional common transcriptional regulation.

By performing array-CGH we also found under-representation of 13q in five of the six tumours. Under-representation of 2p and 8p appeared to be specific for fibroblastic cases. Amplifications were detected in 13 different chromosomal regions, namely at 6p21, 6pter, 7p, 8q, 9p, 11q, 12p, 12q12, 12q13-q14, 12q14-q21.3, 16q22.1-22.2, 17p and 17q. Many of these genomic aberrations have been found before in osteosarcoma and other cancers. Gene expression analysis of these amplified regions in osteosarcoma indicate several genes with an increased expression. In these amplified regions unknown amplification target genes may be located, in addition to one already known gene, \textit{CDK4}, which has earlier been described as amplified in osteosarcoma.
General discussion

Osteosarcoma

Osteosarcoma comprises a heterogeneous group of tumours arising from the bone and either occurs as part of a hereditary (cancer) syndrome or sporadically. Sporadic osteosarcoma, the most frequent form is the focus of this thesis, more in particular the underlying developmental mechanism for this tumour type. It is possible that genes involved in the inherited (cancer) syndromes of which osteosarcoma is part of the clinical phenotype, also play a role in the development of sporadic osteosarcoma. In addition, a number of “common cancer genes” have been found to become altered in osteosarcoma tumours. Cytogenetic studies of primary sporadic osteosarcoma show that the karyotypes of these neoplasms are very complex and make it difficult to define specific chromosome regions associated with osteosarcoma development. Although many genes and chromosomal aberrations have been identified, the developmental process of this tumour is only partly understood. In an attempt to further unravel the underlying genetic aberration involved in osteosarcoma development, we have performed both gene expression profiling and array-CGH analysis of primary osteosarcoma tumours. Since osteosarcoma develops out of osteoblasts, we have compared the gene expression pattern of osteoblasts with those of primary osteosarcoma tumours and osteosarcoma cell lines.

Osteoblast profiling

We have isolated osteoblasts out of bone fragments. Morphologically osteoblasts on culture resemble fibroblasts, another cell type that may grow out of inadequately pre-processed bone fragments, in culture. To discriminate between these cell types, we performed gene expression profiling on cultured fibroblasts and on our short term cultured osteoblasts. We observed clear differences between the cultures, allowing us to characterise osteoblasts by their gene expression pattern. In our list of genes differentially expressed between osteoblasts and fibroblasts, known osteoblast-specific markers such as BMP2, BMP4, OCA, ALP and OPN were missing. Presumably, this is partly due to the short-term culture of the cells, as osteoblasts after long-term culturing. Osteocalcin (OCA) and alkaline phosphatase (ALP) are both late stage markers of osteoblastic differentiation, expressed in mature osteoblasts.
rather than in developing osteoblasts (Bennett et al., 2001). Furthermore, osteopontin (OPN) is excluded from our gene list of differentially expressed genes because of high individual variation, most likely due to low signal intensities. Based on our list of differentially expressed genes we propose that VCAM1, KIAA1644, FGFR2 and COL27A1 are “specifically” higher expressed in short-term cultured osteoblastic cells than fibroblasts and that these genes can be used in differentiating between cell types by real-time RT-PCR.

**Transcriptional profiling**

Transcriptional profiling is a high throughput method for defining global gene expression patterns that likely specify specific cell types (Chapter 2), specific tumour cell lines (Chapter 3) or specific primary tumours (Chapter 4). As we describe in Chapters 2, 3 and 4, we were able to find specific osteoblast, osteosarcoma cell line and primary osteosarcoma tumour profiles. Analysis of the differentially expressed genes between any individual comparisons of the three sample classes resulted in 838 genes co-ordinately discriminating between the classes as visualised in Chapter 4, Fig. 1. Functional analysis showed that in primary osteosarcoma gene ontology categories related to immunoglobulin binding and angiogenesis were significantly differentially expressed in comparison to primary osteoblasts. The fact that the gene ontology category immunoglobulin binding is significantly differentially expressed can be caused by admixture of the tumour tissue with normal tissue. In the osteosarcoma tumours a lower expression of the MMP matrix metalloproteinase gene cluster at 11q22.2-22.3 was observed, indicating a reduced remodelling of the extracellular matrix compared to cultured primary osteoblasts. The comparison of osteosarcoma cell lines and primary osteosarcoma tumours showed significant differences for the gene ontology categories related to transcription and the transcriptional machinery. Investigation of the functional differences between osteosarcoma cell lines and osteoblasts resulted in an elevated cell cycling and DNA replication activity and a diminished extracellular matrix constitution for the cell lines. All these genes ontology categories are represented in Chapter 4, Table 1. Although clear differences are observed between different sample classes, gene expression profiling results in a high degree of variation, attributable to various sources of variation in the total process of transcriptional profiling. Therefore, it is advised to standardise the experiments as much as possible or to avoid common sources of variation by using randomised experimental designs (Dobbin et al., 2005; Bammler et al., 2005).
RNA isolation

Standardisation starts with a verified method for total RNA isolation, throughout the studies described in this thesis the RNeasy method (Qiagen), also recommended by Affymetrix. Integrity of the isolated RNA is extremely important and needs to be assayed and approved before performing transcriptional profiling. We have analysed all RNA isolations by denaturing gel electrophoresis before starting the amplification and labelling procedures. The mRNA amplification protocol was applied in all our studies since short-term cultured osteoblasts as well as osteosarcoma tumour samples yield only small amounts of RNA. The efficiency of the mRNA amplifications was highly consistent across all samples used and resulted in on average 30 micrograms of amplified RNA (aRNA) starting from 2 micrograms of total RNA, resulting in a 1500-fold increase in gene specific transcripts. Labelling of the samples was performed using aminoallyl-modified UTP incorporation during amplification of mRNA and subsequent chemical coupling to Cyanine-3 and Cyanine-5 fluorophores to avoid dye incorporation biases during reverse transcription (’t Hoen et al., 2004).

The expression arrays

All transcriptional profiling experiments described in this thesis have been performed using in-house printed array slides containing 21,329 gene specific 70-mer oligonucleotides from the human oligonucleotide library version 2.1 (Operon Biotechnologies). Moreover, 4000 control spots were included to account for a-specific binding, grid positioning and spike-based normalisation. Using the January 2005 build of Entrez Gene or Locus Link about 1000 genes are represented by at least two independent oligonucleotides on the array, allowing for internal control of the gene expression data as described in Chapter 3.

Experimental design

Another important issue is the experimental design of micro-array experiments (Simon et al., 2002), in particular for dual colour array experiments, but also for single colour arrays/DNA chips. Several designs can be implemented, based on criteria like comparison with future experiments, experiments involving large numbers of samples, multicentre experiments, etc. In this thesis, two different design types have
been used, the reference design (Chapter 2) and a randomised design (Chapters 3 and 4). The type of experimental design was changed from reference to a randomised design when data analysis showed that the assumed competitive hybridisation on dual colour arrays was absent, resulting in two independent readings of different samples on the same array and thus allowing for randomised designs instead of direct comparisons on the same array like the reference design (‘t Hoen et al., 2004). An advantage of this finding is that reference samples can be omitted and, using dye swap experiments, more observations for the same sample are generated, thereby, reducing variation of the signal intensity data. Furthermore, micro-array based signal intensity data for each sample on a dual colour array is treated as a single colour reading and analysed in that way.

Data analysis

Data analysis of micro-array based transcriptional profiling data is probably the most important issue. First, raw data has to be pre-processed before the statistical analysis can be performed. For example, spots with very low intensity signals for both samples, small area spots, spots with aberrant signal distributions, etc have to be omitted from the data. Rather than global methods of normalisation of the raw data, we applied a print-tip loess normalisation to account for technical variation like biases in micro-array fluorescence scanning and spatial effects on the slide due to printing or hybridisation. Subsequently, the normalised and filtered data was statistically analysed for biologically relevant genes using standard F-tests for pairwise significance of each gene. However, depending on the threshold for significance, the list of significant genes also includes a number of genes that are false positives. For example, pairwise tests for 10,000 genes with a significance threshold of 1% (p<0.01) will, by chance, result in 100 false positives. Implementing a more stringent significance threshold results in a lower number of false positive genes in the ultimate gene-list, however, the number of false negative genes, the ones that are truly different among the classes but not selected, will increase. Multivariate permutation is one of the methods to control the number of false positive genes in the gene list (Reiner et al., 2003). Here the phenotype classes to which each sample belongs are rearranged, while no class labels are lost, added or changed. In the transformed data set the significance of each gene is estimated again. This process of rearrangement of phenotype classes is performed randomly for many times and for each rearrangement of the data the number of genes significant for a given threshold
(p<0.001) are recorded. For each significance threshold the number of significant genes in the rearranged data are determined and form a data dependent distribution. Obviously, genes significant for random rearranged data are false positives. Thus, for each significance threshold applied in the true data, the proportion of false positives can be estimated with a preset confidence, e.g. 90%.

In our analyses we have with 90% confidence no more than 10% false positives within the list of significantly differentially expressed genes. For experiments with relatively few samples the number of permutations of the data is rather limited and hence the number of false positives in the lists with significant genes may still be underestimated. Therefore, validation of our obtained micro-array data by real-time RT-PCR is essential.

**Validation**

In **Chapter 2** the genes that discriminate between cultured primary osteoblasts and fibroblasts based on transcriptional profiling were confirmed by real-time RT-PCR on the individual samples. Several genes that discriminated cultured primary osteoblasts from cultured osteosarcoma cell lines were confirmed as well (**Chapter 3**). Moreover, the correlation of the obtained ratios of different oligonucleotide spots representing the same gene also confirmed the discriminatory capacity of the micro-arrays and indicates that multiple independent observations per gene will improve the confidence of results. This principle is taken even further by analysing not individual genes, but gene ontology categories containing many genes at once described in **Chapter 4**.

Because of small differences in analysis settings, a comparison of the same samples can result in different lists of significantly differentially expressed genes. In **Chapter 3** the comparison of primary osteoblasts and osteosarcoma cell lines results in a top 10 list of differentially expressed genes (Table 2) whereas the same comparison in **Chapter 4** results in different top 10 lists (Table 3). Nevertheless, 7 genes of each of the top 10 lists in Table 2 of **Chapter 3** can be found in each of the top 20 lists in Table 3 of **Chapter 4**. The difference between both analyses is that the data has been pre-processed differently in **Chapter 3**. Spots with very low signal intensity were excluded, whereas in **Chapter 4** a lower threshold for low signal intensities was applied. Moreover, the threshold for excluding genes from the analysis based on the percentage of missing values was 25% in **Chapter 3** and 50% in **Chapter 4**, where in the latter case the number of samples was also larger. To retain only the most
relevant genes for each phenotype comparison, genes with more than 1 missing value within each phenotype class were afterwards manually removed from the lists of significant genes described in Chapter 4.

**Genomic aberration in primary osteosarcoma**

As described in Chapter 4 we analysed six primary osteosarcoma tumours, including three osteoblastic tumours and three fibroblastic tumours, by whole genome array-CGH. We observed many copy number changes. Overall, copy number losses were observed more often than copy number gains. A number of the observed alterations were recurrent such as a loss of 13q, in five of the six tumours, 6q, 9q and 20p in four of the six tumours and 1q, 2q, 3p and 11p in three of the six tumours. Furthermore, the loss of 11p was only observed in osteoblastic osteosarcoma. High-level amplifications were detected for 7p and 12p in one or more osteoblastic tumours and for 12q and 16q in fibroblastic osteosarcoma. High-level amplification or gain was observed for 8q and 9p in both types of tumours.

Combining the gene expression profiling data with the aberrant chromosomal aberrations showed that the following regions were amplified in our set of primary osteosarcoma tumours 6p21, 11q22.2, 12q13-q14, 12q14-q21.3, 16q22.1-q22.2, 17p11.2-p13.2 and 17q25-qter. In these regions various genes could be identified of which the expression is most likely affected by regulation of copy numbers. In the region 12q13 several genes, including CDK4, were higher expressed in the osteosarcoma tumours as was already described by Wunder et al. (1999).

**Future perspectives**

We have identified gene expression profiles specific for cultured primary osteoblasts, osteosarcoma cell lines and primary osteosarcoma tumour tissue. Moreover, with array-CGH we provided corresponding genomic profiles of these osteosarcoma tumours. These analyses are based on restricted numbers of cultured cells or whole tumour tissue samples. From these studies two main challenges for future experiments were distilled. First, the strong admixture effect of inflammatory cells in osteosarcoma tumour tissue (chapter 4) needs to be addressed. Tumour tissue heterogeneity of the samples increases the variability of gene expression data
considerably. Isolation of true cancerous osteosarcoma cells with laser capture microdissection techniques would prevent this admixture effect and hence reduce variability of gene expression data. Secondly and equally important, the number of samples for each phenotype class need to be increased in order to more comprehensively identify genes that are involved in osteosarcoma and to better discriminate between subtypes of osteosarcoma. Challenges for micro-array experiments in general will be reduction of the variability of gene expression and to a lesser extent array-CGH data. In almost all cases a single oligonucleotide spot represents a single gene, transcript or exon on the entire array. Small variations in the fluorescent signal intensity will therefore add to the total variability across multiple arrays. Multiple spots of the same gene-specific oligonucleotide on each array would surely reduce the inherent variability of the fluorescent signal intensities within and across arrays and improve the value of the experiments.

References


