Analysis of chromosomal copy number aberrations in gastrointestinal cancer

Haan, J.C.

2014

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

SUMMARY AND DISCUSSION

The overall aim of this thesis was to identify chromosomal copy number alterations that could serve as biomarkers for outcome and guide therapy selection in patients with gastrointestinal (GI) cancer, with an emphasis on colorectal cancer (CRC).

We used DNA copy number profiles to compare different types of GI cancers, to predict response to therapy and select potential drug targets. For data analysis the statistical computing language R was used. R-packages available at Bioconductor or developed in-house were used and new methods were programmed in R.

Gastrointestinal cancers harbor both large and focal chromosomal aberrations. The first arrayCGH platforms were only able to detect the large aberrations. These aberrations encompass hundreds of genes, which made it difficult to pinpoint actual cancer genes. Increased resolution has led to the discovery of focal chromosomal aberrations, defined as 3Mb or smaller in size. Inherent to their limited size, focal aberrations harbor one or very few genes only. In the first part of this thesis (**chapter 2**) we demonstrated that focal chromosomal copy number aberrations in stage II colon cancer are enriched in cancer genes that contribute to and drive the process of colorectal cancer development. Aberrations were determined in 38 formalin fixed paraffin embedded (FFPE) by 44K arrayCGH colon cancer samples using matched normal mucosa as a reference.

In total, we identified 81 focal chromosomal aberrations, harboring 177 candidate genes. Validation of focal aberrations and identification of candidate genes were done using publicly available copy number and mutation data of colorectal cancer, breast cancer, pancreatic cancer and glioblastomas. Enrichment analysis demonstrated a statistically significant overlap of the focal aberrations with previously identified focal amplifications in colorectal cancer, but not with focal amplifications of cancers from other sites. In contrast, focal deletions we detected, also showed significant overlap with focal deletions of cancer from other sites. Focal deletions therefore appear to be less tumor type-specific than focal amplifications. In addition, focal deletions are enriched for cancer census genes and genes frequently mutated in colorectal and other cancers.

As a technical validation we demonstrated that focal aberrations detected with the DNA from an FFPE sample was identical to focal aberrations detected with DNA isolated from a fresh sample. In addition, we also detected similar results on arrayCGH platforms of an independent manufacturer with a completely different probe design, i.e. the Nimblegen 135K array and 44K Agilent array platforms. Hence, we conclude that focal chromosomal aberrations can be detected with FFPE tumor tissue independent of the platform.

As biological validation we demonstrated gene dosage effects of the 177 candidate genes located on the identified focal aberrations. The mRNA expression of these genes is significantly correlated with DNA copy number status, supporting the relevance of focal aberrations.

In the second part of this thesis we applied (chapter 3-5) several approaches to investigate the use of DNA copy number aberrations as biomarkers for therapy selection in small bowel, gastric and colorectal cancer.

Small bowel cancer (SBC) has a low incidence and consequently hardly any clinical trials are performed to determine optimal drug therapy regimens. In practice therefore, either a colorectal or a gastric cancer treatment regimen is currently chosen. In **chapter 3** we demonstrated that small bowel adenocarcinoma copy number profiles are more similar to CRC than to gastric cancers (GC). The common genomic characteristics of CRC and SBC might be an indication of shared biology between these phenotypes, and based on the paradigm that genotype drives phenotype, including response to drug therapy, one may hypothesize patients with small bowel cancer may benefit more from colorectal cancer drug regimens than from gastric cancer drug regimens.

Chromosomal regions can be either lost (< 2 copies), gained (3-4 copies) or gained at high copy number level (> 4 copies) with proportional effects on the level of gene expression. New anti-cancer drugs increasingly are aimed at specific biological targets like overexpressed oncogenic proteins. Since gains of DNA at high level copy number is an important mechanism of overexpression, DNA copy number profiling may provide clues about potential relevant drug targets present in a given tumor. Therefore a genome-wide survey for gene high level copy number gains may serve as a screen for potential candidate drug targets relevant to the individual patient, e.g. ERBB2 in breast cancer.

In **chapter 4** we catalogue high level copy number gains of established and potential drug target genes in gastric cancer as a potential lead for treatment development and selection.

Aberrations were analyzed for 183 gastric adenocarcinomas by arrayCGH, performed on Bacterial Artificial Chromosomes (BAC). Because of the relatively low resolution of this platform, only large aberrations could be analyzed. For practical purposes high level copy number gains were defined as a minimum of 2 consecutive clones showing a log2-ratio of ≥1, matching at least 4 copies and were used for further identification of drug target genes.

A total of 147 highly gained regions were identified of which 78 regions were found to contain a total of 167 genes that were annotated as drug target genes by Ingenuity Pathway Analysis (IPA). Thirty of these occurred in at least 2% of patients.

The drug target genes found to be gained included 5 targets of cytotoxic agents and 4 targeted agents that are currently being used clinically such as the microtubule function related genes (TUBB3, TUBG1 and TUBG2) and gain of the TOP2A coding gene. In addition, 4 gained genes are known drug targets for other indications, but these could be of interest for anticancer treatment as well. Sixteen other gained genes included molecular targets that are either under (pre)clinical evaluation or are new targets that could potentially be druggable.

Copy number aberrations may also indicate resistance for therapy such as mutations in KRAS to anti-EGFR therapy. To avoid unnecessary toxicity and healthcare costs, predictive markers of response to therapy are needed. Current systemic treatment for metastatic colorectal cancer is still largely based on shot gun approaches ("one-size-fits-all"), and most patients are treated with combination chemotherapy regimens containing (oral) 5-fluorouracil (5FU) formulations and oxaliplatin or irinotecan, with or without bevacizumab, while the EGFR-targeted monoclonal antibodies cetuximab and panitumumab are administered to patients with KRAS wild type tumors. In **Chapter 5** we documented the landscape of DNA copy number aberrations in primary tumors of a defined subset of colorectal cancer patients who developed metastatic disease and are amenable for systemic treatment. Furthermore we catalogue aberrations which correlate to treatment or might be markers for therapy selection.

We generated a high quality arrayCGH data set of clinically well annotated FFPE colorectal cancer samples from patients who participated in two phase III clinical trials, (CAIRO)[1] and (CAIRO2).[2] Both tumor DNA and normal DNA was isolated. To detect the DNA copy number alterations high resolution array 180K Agilent was used. After passing a number of inclusion criteria the final dataset comprised 349 high quality copy number profiles including 111, 111 and 134 samples of both CAIRO armA, CAIRO armB and CAIRO2 armA, respectively. To measure response to treatment, first line progression free survival (PFS) was used. Patients in CAIRO armA were treated with capecitabine monotherapy, patients in CAIRO armB with capecitabine plus irinotecan (CAPIRI) and CAIRO2 ArmA with capecitabine, oxaliplatin, and bevacizumab (CAPOX-B). Unsupervised hierarchical clustering of these profiles did not reveal an association of cluster membership with PFS for any of the three regimens. However, by supervised analysis 187 regions could be identified that significantly associated with PFS in at least one of the three arms (p-value < 0.05). For the patients treated with capecitabine monotherapy as first-line treatment, the most significant association with PFS was found for loss of a region at 5q12.1-q12.3. Loss of this region was associated with a significantly shorter PFS. For the patients treated with CAPIRI, the most significant association with PFS was found for loss of a region at 18q21.33-q22.3. Loss of this region was associated with a significantly longer PFS. For the patients treated with CAPOX-B, the most significant association was found with loss of a region at 5q34. Loss of this region was associated with a significantly shorter PFS.

To further narrow down on the gene level we determined gene dosage effects of the significant regions with p-values < 0.005. Twenty-four of the 187 significantly associated regions were associated to PFS with p-value < 0.005, containing 1744 genes in total. Since no expression data were available for the CAIRO and CAIRO2 series, matched DNA copy number and mRNA expression microarray data from The Cancer Genome Atlas (TCGA) for 141 colorectal cancer cases were used to identify genes that showed such a correlation. For 608 of these genes a positive correlation between DNA copy number and mRNA expression level was found.

In addition, as previously done in chapter 4, we catalogued high level copy number gains in this dataset of 349 primary colorectal cancer samples. In total 692 high level copy number gains were observed in total, 432 of which can be classified as amplifications (<3Mb).

A total of 5145 genes located on the highly gained regions and 1371 genes were gained at high level in more than 2% of the samples for 603 of which a positive correlation between DNA copy number and mRNA was found in with the TCGA dataset. On the high level copy number gains identified in this series of CRCs we identified 12 targets of FDA-approved drugs and 9 target genes for targeted anticancer agents under (pre) clinical evaluation or for FDA approved drugs with potential anticancer activity. Moreover we identified highly gained kinases, ligands and receptors of which 8 drug targets were associated with drugs, but also genes that have not yet been targeted.

In the previous chapter we identified potential markers for therapy selection and potential drug targets. These markers and the analysis of other copy number profiles used in part 2 of this thesis were based on the primary tumor. However the majority of colorectal cancer deaths are due to metastasis. For that reason in part 3 of this thesis we studied in 2 chapters (6 and 7) DNA copy number profiles in association with metastasis. Twenty percent of CRC patients have metastatic disease at time of diagnosis and up to 50% of all CRC patients develop metastasis.[3] The liver is the predominant metastatic site in approximately 80% of metastatic CRC patients. In 40 to 50% of these patients, also extrahepatic metastases are present.[4] Other metastatic sites include lung, the central nervous system, adrenal glands, ovaries, skeleton and skin.[5] Studying copy number profiles of metastatic preference of the primary tumor may be of prognostic value for predicting metastatic site and recurrence of CRC patients. In **Chapter 6** we demonstrated that chromosome 20p11 gains were associated with hepatic-specific metastasis in colorectal cancer patients.

For the data analysis, DNA copy number profiles generated for the study of chapter 5 were re-analyzed. Patients were selected according to the site of the metastases, i.e. hepatic or extrahepatic. Patients with a combination of hepatic and extrahepatic metastases, locally advanced disease and of whom the metastatic site was unknown were excluded. Furthermore, the CAIRO series was used as a test set and the CAIRO2 series as a validation set. The test set included 85 and 54 primary tumors of CRC patients with hepatic and extrahepatic metastases, respectively. The validation set included 45 and 35 primary tumors with hepatic and extrahepatic metastases, respectively.

Although the frequencies of DNA copy number aberrations were highly similar for patients with hepatic and extrahepatic metastases, supervised analysis revealed a few differences between hepatic and extrahepatic metastases. Patients with hepatic metastases had significantly more frequently gains at 20p11 and significantly fewer losses at 5q12 compared to patients with extrahepatic metastases. In the validation set we confirmed the association with 20p11 gain in patients with hepatic versus extrahepatic metastases, but differences in copy number aberrations at 5q12 could not be validated.

To narrow down on the genes with a potential role in hepatic-specific CRC metastasis, we determined gene dosage effects of the genes located at 20p11 in the publicly available TCGA dataset as previously used as validation in chapter 5. This approach revealed 12 genes that were significantly differently expressed between samples with or without a gain of 20p11. Of these 12 differentially expressed genes, C20orf3 showed the strongest correlation between copy number status and RNA expression. Subsequently the protein expression of C20orf3 was tested by using tissue microarrays of both CAIRO and CAIRO2 patients for immunohistochemical (IHC) stainings. We found that also protein expression of C20orf3 was significantly correlated with the copy number data. Moreover we performed a test to compare protein expression between patients with hepatic versus extrahepatic metastases. For this purpose 581 patients, of which 325 had hepatic metastases and 283 extra-hepatic metastases were available. We could confirm a significantly higher percentage of C20orf3 staining in the primary tumor of patients with hepatic metastases compared to patients with extrahepatic metastases. We thus identified a potential marker that could be targeted to prevent metastasis to the liver. It is also important to understand differences between primaries and metastases in order to understand molecular events leading to metastasis. Current clinical practice is to use archived material of the primary tumor for therapy selection, even though therapies are targeted against metastases. In this thesis, copy number profiles used for therapy selection and prediction of response to therapy were based on the primary tumors. It is assumed that the metastatic potential is predestined early in the development of the primary tumor.

In **Chapter7** we demonstrated that patterns of DNA copy number aberrations between primaries and metastases were highly similar for all patients and additional copy number aberrations in colorectal metastases were rare and typically non-recurrent.

For this study we analyzed DNA copy number profiles of 62 primary tumors and 68 metastases of the same patients. For six patients metastases were available from two sites. Unsupervised clustering revealed that most primaries are much more similar to their metastases than to other primary tumors. For only six patients, metastases and the corresponding primary tumors did not completely join pair wise in the cluster dendrogram.

By pairwise comparison we discovered at three sites in two patients a co-amplification of two amplifications in the metastases, which were not detected in the primary tumor. The amplifications were located at 6q21 and 8q24.21, the latter encompassing the MYC oncogene. To confirm that the co-amplification is metastasis specific we analyzed 349 primary colorectal tumors of the CAIRO studies and 193 primary colorectal tumors of the TCGA dataset. We detected an amplification of MYC alone, once in a primary colorectal tumor of the CAIRO datasets and three times in the TCGA dataset. The 6q21 amplification was only detected once in the CAIRO datasets, but not observed in the TCGA dataset. In none of these 542 primary tumors a co-amplification of MYC and chromosomal locus 6q21 was observed.

The amplifications at DNA level were confirmed by fluorescence in situ hybridization (FISH) analysis showing amplifications of MYC and chromosome 6q21 in the metastases. In addition by FISH analysis we did not observe subclones in the primary tumor with an amplification of MYC and/or chromosome 6q21, confirming the hypothesis that the metastatic potential is predestined early in the development of the primary tumor. Moreover we demonstrated that the co-amplification did not result from a translocation, since no co-localization of the amplified chromosomal regions was observed in the FISH analysis.

GENERAL CONCLUSION

In this thesis we analyzed chromosomal copy number aberrations in gastrointestinal cancer.

In a small series of stage II colon cancer patients we identified losses of 5q34 and gain of 13q22.1 as independent prognostic factors of survival. These genomic alterations may aid in selecting patients for adjuvant therapy. We furthermore show that microsatellite stable small bowel adenomas are more similar to colorectal than to gastric cancer. These molecular similarities provide additional support for treating microsatellite stable small bowel cancers according to a colorectal cancer regimen. A catalogue of high level copy number gains in primary gastric cancer revealed gene loci that may be of therapeutic interest, including known substrates for systemic therapies used in advanced gastric cancer, as well as new targets for treatment that are of interest for evaluation of antitumor activity in tumors carrying these high level copy number gains. Moreover we provided a detailed genomic landscape of metastatic colorectal cancer from actual clinical specimens. In a large cohort of metastatic CRC patients we found loss of 5q12.1-q12 to be related to progression free survival in patients treated with capecitabine monotherapy, a loss of 18q21.33-q22.3 to be related to progression free survival in patients treated with capecitabine combined with irinotecan as first-line treatment and a loss of 5q34 is related to progression free survival in patients treated with capecitabine combined with oxaliplatin and bevacizumab. The identified candidate biomarkers may have the potential to predict response to therapy, when further validated. The gene C20orf3 mapping at 20p11 is associated with hepatic-specific metastasis in patients with CRC. This gene is a candidate biomarker for liver metastases and may be of clinical value in early-stage CRC. The biomarkers identified have the potential to predict response to therapy and treatment outcome. However we have not yet been able to validate these findings, mainly because suitable validation series were not available. The challenge ahead is to validate these associations in independent series in order to arrive to a solid basis for these biomarkers to be used in clinical decision making. Moreover, a combination of copy number, methylation and mutation information might add power to identify genetic predictive markers. With the arrival of next generation sequencing and its applicability to DNA isolated from FFPE archival material, this has now come within reach.

Finally, in a series of matched primary and metastatic tumors we observed that additional DNA copy number aberrations in metastases are rare and rather than recurrent they are specific to individual patients. These observations are consistent with the hypothesis that the full genomic program that determines the biological and clinical phenotype of a tumor is already present at the time when a primary cancer arises. This observation is in line with those for KRAS mutations, that are similarly present in the primary tumor as well as in the metastasis.[6,7]

REFERENCES

1. Koopman M, Antonini NF, Douma J *et al*. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. Lancet 2007;370(9582):135-42.
2. Tol J, Koopman M, Cats A *et al*. Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. N Engl J Med 2009;360(6):563-72.
3. Kindler HL, Shulman KL. Metastatic colorectal cancer. Curr Treat Options Oncol 2001;2(6):459-71.
4. Sanoff HK, Sargent DJ, Campbell ME *et al*. Five-year data and prognostic factor analysis of oxaliplatin and irinotecan combinations for advanced colorectal cancer: N9741. J Clin Oncol 2008;26(35):5721-7.
5. Hermanek P, Jr., Wiebelt H, Riedl S *et al*. [Long-term results of surgical therapy of colon cancer. Results of the Colorectal Cancer Study Group]. Chirurg 1994;65(4):287-97.
6. Knijn N, Mekenkamp LJ, Klomp M *et al*. KRAS mutation analysis: a comparison between primary tumours and matched liver metastases in 305 colorectal cancer patients. Br J Cancer 2011;104(6):1020-6.
7. Diaz LA, Jr., Williams RT, Wu J *et al*. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature 2012;486(7404):537-40.