# Distinct recurrent copy number changes between BRCA1-related and sporadic ovarian tumours, identified by aCGH

Karin Leunen[1*], Olivier Gevaert[2*], Anneleen Daemen[2], Vanessa Vanspauwen[3], Geneviève Michils[3], Eric Legius[3], Bart De Moor[2] and Ignace Vergote[1]

(1) Division of Gynecological Oncology, Department of Obstetrics and Gynecology, University Hospital Leuven, Katholieke Universiteit Leuven, Belgium
(2) Cancer bioinformatics, Department of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, Belgium
(3) Department of Human Genetics, University Hospital Leuven, Katholieke Universiteit Leuven, Belgium

* Both authors contributed equally

TODO:
- discussie herschrijven: vgl maken tussen gevonden CNA en onze CNA + verwijzen naar supplementary figures
- betere beschrijving figuur 3 in results
- getallen in tekst en in figuren controleren.
- ref Bueker 2000 toevoegen

Wat van **Genes, chromosomes and cancer** [IF 4.5] en ik denk dat we dan al op ons doos gaan krijgen voor dit artikel, maar bon ☺ wie nie waagt …

## Abstract

**Aim:** As BRCA related ovarian cancer seems to follow distinct pathways in their carcinogenesis such as described by standard Comparative Genomic Hybridization (CGH), we evaluated this at a much higher resolution with array CGH (aCGH).

**Materials and Methods:** A total of 13 ovarian tumor samples originating from 8 sporadic and 5 BRCA1 related ovarian cancers, were subjected to aCGH. After pre-processing, both groups of patients were modeled using a recurrent Hidden Markov Model to detect recurrent copy number alterations (CNAs). CNAs with a probability higher than 80% were called.

**Results:** We observed more CNAs in the BRCA1 group compared to the sporadic group: the sporadic group had the most gains whereas the BRCA1 group had the most losses. Overall we did not observe a difference in length of the CNAs. However, when focusing on the type of CNAs, no significant difference in length was seen for the gains, but there was a statistically significant difference between the lengths of the losses (median length 5.2 Mb vs. 0.2 Mb) respectively, P-value < 1.8e-8). When considering all differential CNAs, up to 6581 genes are affected. Focusing on the losses, 4775 genes were affected in the BRCA1 group, in contrast to 327 genes in the sporadic group.

**Discussion:** BRCA1 ovarian tumors present a different type, number and length of CNAs. A huge amount of the genome is lost in the BRCA1 group compared to the sporadic group. This indicates genomic instability in BRCA1-mutated tumors and confirms the known role of BRCA1 in homologous recombination repair of double stranded DNA breaks. Further analysis is required to investigate whether the genes in these CNA correspond to different biological pathways in BRCA1-mutated tumors.


**Comment [OG1]:** Gains/losses or breakpoints???

**Comment [OG2]:** Mischien hier andere afsluitende zin, Ik denk zo weinig mogelijk naar genen/pathways verwijzen want anders gaan we dat ook mogen in het artikel schrijven.

## Introduction

Array comparative genomic hybridization (aCGH) is gaining more interest in the last years and is now widely used in different fields of oncology (1). The main advantage of aCGH compared to traditional CGH is that this technology measures variations in DNA copy number at a genome wide scale at a much higher resolution.

The aim of this study was to look for differences in gains and losses in BRCA related ovarian tumors compared with sporadic ones. Many previous studies investigated copy number alterations (CNA) in sporadic and BRCA related ovarian tumors using traditional low resolution CGH, but comparison between both using aCGH in order to identify possible distinct pathways and to learn more about their oncogenesis is - to our knowledge - not performed yet.

Here we report the results of the first genome wide characterization of recurrent CNA in BRCA1-mutated and sporadic ovarian tumors. We focus on their type, number and length, as a basis in our search to demonstrate that distinct pathways are active in sporadic and hereditary ovarian cancer.

## Materials and methods

### Sample description

Data from 13 patients treated for ovarian cancer at the University Hospital of Leuven, Belgium (5 BRCA1 related - and 8 sporadic ovarian cancers) were collected for participation at this study. In order to obtain a group of patients as homogeneous as possible, only patients with similar clinical characteristics were retained: they were all stage III-IV, serous papillary poorly differentiated ovarian cancers. All tumor samples

were collected at the time of primary surgery and immediately rinsed and frozen. In aCGH, the number of malignant cells in each sample determines the sensitivity of the technique. Therefore, in order to reduce the sensitivity to noise and to limit stromal contamination which also decreases significantly the sensitivity for detecting copy changes, histological confirmation of high percentage of tumor load (70%) was requested to our pathologist. DNA was extracted from the fresh frozen samples after histological confirmation that a high percentage of tumor tissue was present in each biopsy. Each tumor was hybridized twice (dye-swap) against a common reference pool. The reference cell population was extracted from peripheral blood samples of each matched participating patient.

ArrayCGH and preprocessing

After extraction of DNA, tumor DNA was labeled by a random priming method (Bioprime® DNA labeling system; invitrogen, Cergy-Pontois, France) with cyanine-5. Reference DNA was similarly labeled with cyanine-3. After purification, co-precipitation with 50µl Cot-1 DNA (Invitrogen, Cergy-Pontoise, France), resuspension in hybridization buffer (20% formamide), denaturation at 75°C for 10min and prehybridization at 37°C for 60 min. were subsequently performed. Probes were co-hybridized on aCGH. The aCGH slide was previously preblocked with a buffer. The used aCGH slide was a 1Mb array CGH platform, version CGH-SANGER 3K 7 developed by the Flanders Institute for Biotechnology (VIB), Facility, Leuven, Belgium. After washing, arrays were scanned using a 4000B scanner (Axon, Union City, CA). Genepix Pro 6.0 software (Axon) was used for image analysis and data were further analyzed by Microsoft Excel. First, Cy5 and Cy3 fluorescence intensities were background corrected.

Then, the ratio of the Cy5 to Cy3 was calculated and normalized by the median Cy5-Cy3 ratio. Finally, these ratios were log transformed and the average of the replicate (dye-swap) experiments was used as input for subsequent modeling.

Recurrent Hidden Markov Model

To identify differential regions, the BRCA1 and sporadic samples were analyzed using Hidden Markov Modeling (HMM)(2). In brief, each sample is modeled by a HMM with three hidden states corresponding to Copy Number Loss (CNL), neutral and Copy Number Gain (CNG). A HMM allows to take into account the genomic neighborhood. Instead of modeling each sample separately with a HMM, we grouped samples according to there BRCA1 status with a recurrent HMM (RHMM). This allows identification of recurrent CNA found at the same location in multiple samples using a statistical model. The recurrent HMM delivers the probability of CNA across all samples belonging to the same group. A CNL or CNG was called when its probability of occurring was more than 80%.

Statistical analysis of CNA

Based on the results of the differential analysis, we compared the CNA of both groups according to their number, type (i.e. CNG or CNL) and length. We used the Wilcoxon rank sum test to assess statistical significance. All tests were two-sided and the significance threshold was 0.05.

**Results**

**Identification of CNA**

After pre-processing, we applied RHMM-modeling for both the BRCA1 and sporadic samples separately. This results in a separate set of CNA characterizing the BRCA1 and sporadic group. Figure 1 shows the probability of CNA for the BRCA1 and sporadic group on chromosome 1 (see supplementary file 1 for the CNA for all other chromosomes). The probability of copy number gains (CNG) and copy number losses (CNL) are indicated in green and red respectively. Next, CNG and CNL with a probability of 80% or more are shown on top of the raw probabilities. These CNA are "called" and will be used for further analysis.

**Statistical analysis of CNA from differential analysis**

Based on these called CNA, we investigated whether there is a difference in the number, the length and the type of CNA between both groups. The number of breakpoints was similar between the two groups. There were 178 breakpoints in the BRCA1 tumors compared to 164 breakpoints in the sporadic patients. Figure 2a shows the number of breakpoints for each chromosome and according to patient group. This indicates that the number of breakpoints is similar between BRCA1 and sporadic patients.

When focusing on the type of CNA, the sporadic group thus has the most CNG while the BRCA1 group had the most CNL. More specifically, the BRCA1 group is characterized by 98 CNA corresponding to 34 recurrent CNG and 64 CNL, compared to 86 CNA in the sporadic group corresponding to 40 CNG and 46 CNL.

Next, we focused on the length of the CNA. We did not observe an overall difference in length between the two groups. When focusing on the type of CNA however, there was a statistically significant difference between the lengths of the CNL. CNL in the BRCA1 group were typically longer. The median length was 5.2 Mb in the BRCA1 group vs. 0.2

Mb in the sporadic group (P-value < 1.8e-8, see Figure 2b). The CNG were also longer in the BRCA1 group; however this difference was not significant (median length of 5.3 Mb and 2.5 Mb for the BRCA1 and sporadic group, respectively, see Figure 2c).

Next, Figure 3 shows the number and median length of the CNG and CNL per chromosome. When focusing on the median length of the CNL, Figure 3d confirms the above mentioned results since the CNL are longer in BRCA1-mutated tumors in almost all chromosomes.

In total 730 Mb or ~22% of the human genome are aberrated (257 Mb gained and 473 Mb lost) in BRCA1 tumors compared to 475 Mb or ~15% in sporadic tumors (384 Mb gained and 91 Mb lost). When considering all differential CNA, up to 6581 genes are affected. Focusing on the losses, 4775 genes were affected in the BRCA1 group, in contrast to 327 genes in the sporadic group.

Comment [OG3]: Total nr of base pairs zie ensembl: 3 253 Mb

**Discussion**

Our results confirm that CNA are distinct in BRCA and sporadic ovarian cancer. This was already demonstrated by others but at lower resolution seen performed using metaphase CGH (3-6). (Bueker 2000) To our knowledge, this is the first study which profiled CNA at a much higher resolution to study BRCA1 mutations in ovarian cancer.

Our results show that losses are more frequent in the BRCA1 group. It is important to state that different probability levels in the range of 50% to 90% for calling CNA all resulted in a significantly higher length of CNL in BRCA1 tumors compared to sporadic tumors. Our results thus are not depended on the threshold for calling CNA. This suggests that related down-regulated genes are more likely to be tumor suppressor genes,

Comment [OG4]: Ik begrijp dit stuk niet

potentially not yet known. Moreover, taking into account the central role of BRCA1 in DNA damage response via multiple cellular and biochemical functions such as DNA repair, transcription, cell cycle control and ubiquitination, this gene seems to use numerous pathways to ensure the genomic stability. Additionally, figure 2 shows that the losses are lager in the BRCA1 group (5.2Mb versus 0.2Mb) suggesting that large parts of the genome are simply absent in the hereditary group and a huge amount of genetic material is lost (in total 473 Mb lost). This contributes to the important genetic instability of these hereditary tumors and has not been described before.

Previous reports using metaphase CGH described a number of these CNA and there localization on the genome. For example, gains in 8q and 3q were often demonstrated but not qualifying as specific for the sporadic group (3). However differential higher proportion of gains at 11q22, 13q22 and 17q24-25 were described in the hereditary group (3). Also deletions at 15q11-15, 15q2-25, 8q21 ter, 22q13 and 12q24 appear to be specific to hereditary ovarian cancer. Tapper(4) described amplification of 2q24-32 as the only statistically difference between BRCA and sporadic ovarian cancer.

In our series, losses are more clearly present than gains in contradiction with the CGH findings of Zweemer et al. Moreover, the lengths of CNA reveals that BRCA related ovarian cancers presents significantly more breakpoints (gains and losses) in chromosome 5,7, 8, 10 and 18 whereas chromosome 2, 4 and 23 harbor a larger number of breakpoints in the sporadic group. Restriction to the CNLs, highlight chr 5, 7 and 13 for the BRCA1 group whereas most losses in the sporadic group are located on chr 2, 4 and 6.

*It will now be of special interest to investigate these CNA and their location [??? Location zit toch in figuur 1 and supplementary figures] on the genome. As this will result in an enormous amount of genes, statistical analysis of these genes and their co-occurrence will permit signature construction, identifying patterns of pathway deregulation in these different tumours. This can be clinically relevant when these data can be associated with disease outcomes and possibly predict sensitivity to therapeutic agents(7).*

**Conclusion:**

We describe important differences in type, number and length of CNA as a reflection of different carcinogenesis between both sporadic en BRCA1- related ovarian cancer.

*Further localization of these CNA on the genome is ongoing and will learn us more about the involved genes and several gene-sets which contribute to the carcinogenesis of these sporadic and BRCA related ovarian tumours.*

Further analysis is required to investigate whether the genes in these CNA correspond to different biological pathways in BRCA1-mutated tumors. (ipv van vorige)

**Comment [OG5]:** Mischien hier andere afsluitende zin, Ik denk zo weinig mogelijk naar genen/pathways verwijzen want anders gaan we dat ook mogen in het artikel schrijven.

**Figures:**

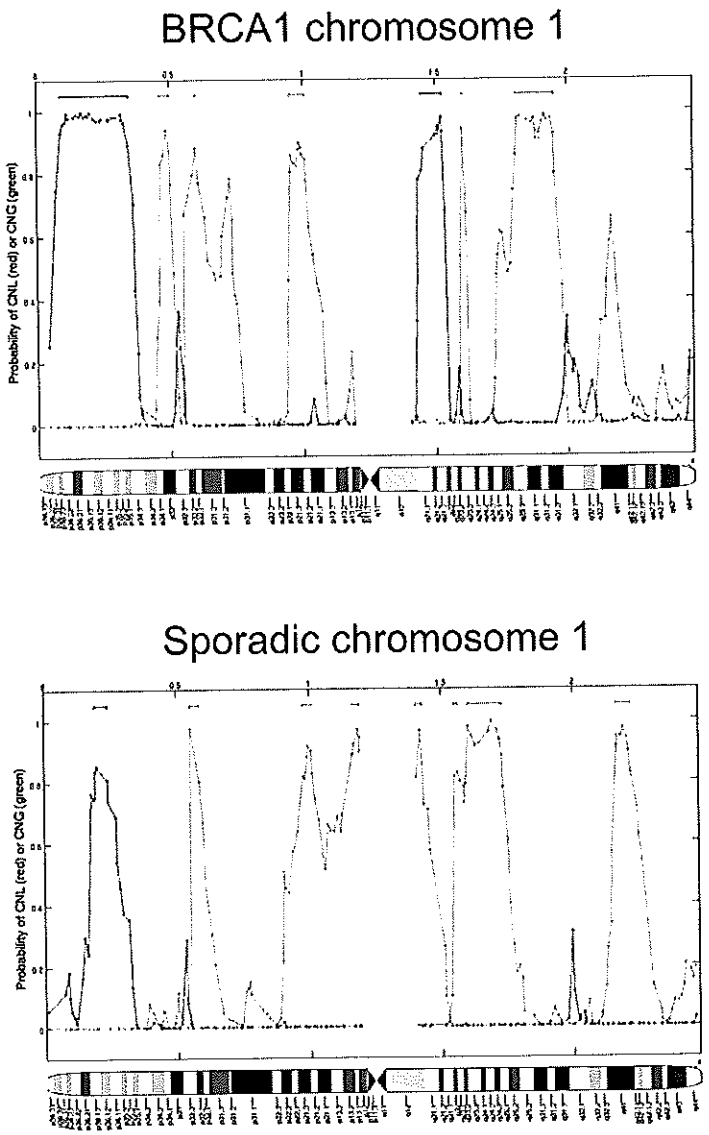



Figure 1: Probabilities for Copy Number Loss (CNL) in red and Copy Number Gains (CNG) in green after Recurrent Hidden Markov modeling (RHMM). Top panel: CNL and CNG on chromosome 1 for the BRCA1 group. Bottom panel: CNL and CNG on chromosome 1 for the sporadic group. CNG and CNL with a probability of 80% or more are shown on top of the raw probabilities and were further analyzed.

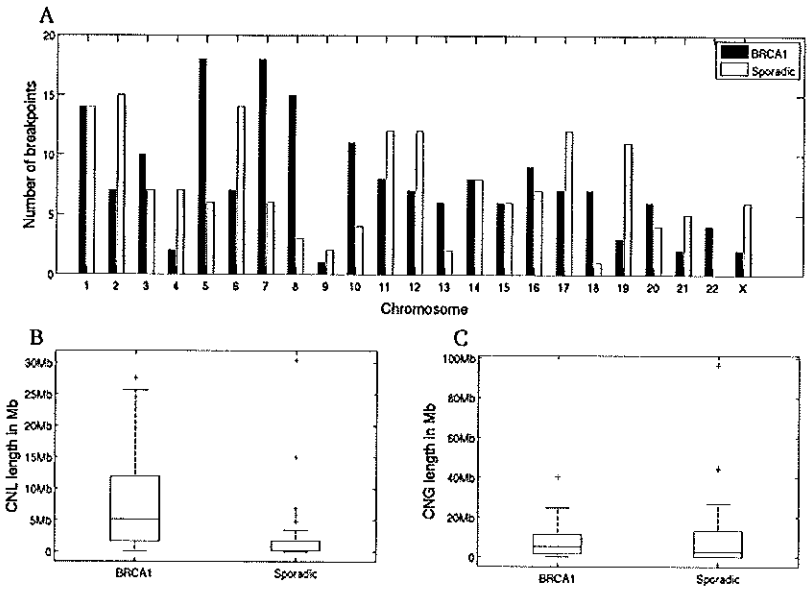



Figure 2: Visualization of the number of breakpoints per chromosome and the length of all copy number losses (CNL) and gains (CNG) separately for the BRCA1 and sporadic patient groups. A: number of breakpoints per chromosome, B: boxplot of length of CNL, C: boxplot of length of CNG.

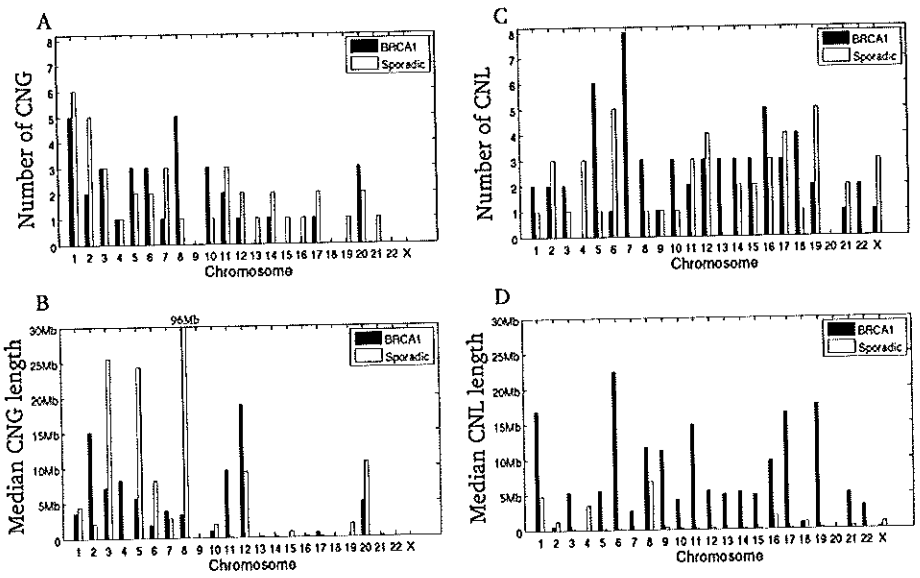



Figure 3: Visualization of the number and length of recurrent copy number gains (CNG) and losses (CNL) per chromosome separately for the BRCA1 and sporadic patient groups. A: Number of CNG, B: median length of CNG, C: Number of CNL, D: median length of CNL. The 8q gain of 96Mb in sporadic patients is not completely visualized such that y-axis length is similar in panels B and D.

## References


1 Davies,J., Wilson,I. and Lam,W. Array CGH technologies and their applications to cancer genomes, Chromosome Research, *13:* 237-248, 2005.

2 Shah,S.P., Lam,W.L., Ng,R.T. and Murphy,K.P. Modeling recurrent DNA copy number alterations in array CGH data, Bioinformatics, *23:*2007.

3 Zweemer,R.P., Ryan,A., Snijders,A.M., Hermsen,M.A., Meijer,G.A., Beller,U., Menko,F.H., Jacobs,I.J., Baak,J.P., Verheijen,R.H., Kenemans,P. and van Diest,P.J. Comparative genomic hybridization of microdissected familial ovarian carcinoma: two deleted regions on chromosome 15q not previously identified in sporadic ovarian carcinoma, Laboratory investigation; a journal of technical methods and pathology, *81:* 1363-1370, 2001.

4 Tapper,J., Sarantaus,L., Vahteristo,P., Nevanlinna,H., Hemmer,S., SeppÃ¤lÃ¤,M., Knuutila,S. and Butzow,R. Genetic changes in inherited and sporadic ovarian carcinomas by comparative genomic hybridization: extensive similarity except for a difference at chromosome 2q24-q32, Cancer research, *58:* 2715-2719, 1998.

5 Patael-Karasik,Y., Daniely,M., Gotlieb,W.H., Ben-Baruch,G., Schiby,J., Barakai,G., Goldman,B., Aviram,A. and Friedman,E. Comparative genomic hybridization in inherited and sporadic ovarian tumors in Israel, Cancer Genet Cytogenet, *121:* 26-32, 2000.

6 Israeli,O., Gotlieb,W.H., Friedman,E., Goldman,B., Ben-Baruch,G., viram-Goldring,A. and Rienstein,S. Familial vs sporadic ovarian tumors: characteristic genomic alterations analyzed by CGH, Gynecol Oncol, *90:* 629-636, 2003.

7 Bild,A., Yao,G., Chang,J., Wang,Q., Potti,A., Chasse,D., Joshi,M.B., Harpole,D., Lancaster,J., Berchuck,A., Olson,J., Marks,J., Dressman,H., West,M. and Nevins,J. Oncogenic pathway signatures in human cancers as a guide to targeted therapies, Nature, 2005.