BMC Cancer

REVIEW   Open Access

The breast cancer genome - a key for better oncology

Hans Kristian Moen Vollan[1,2,3,4] and Carlos Caldas[3,4,5]*

Abstract

Molecular classification has added important knowledge to breast cancer biology, but has yet to be implemented as a clinical standard. Full sequencing of breast cancer genomes could potentially refine classification and give a more complete picture of the mutational profile of cancer and thus aid therapy decisions. Future treatment guidelines must be based on the knowledge derived from histopathological sub-classification of tumors, but with added information from genomic signatures when properly clinically validated. The objective of this article is to give some background on molecular classification, the potential of next generation sequencing, and to outline how this information could be implemented in the clinic.

Molecular classification of breast cancer

The diversity of breast cancer has been acknowledged for decades, but recent technological advances in molecular biology have given detailed knowledge on how extensive this heterogeneity really is. Traditional classification based on morphology has given limited clinical value; mostly because the majority of breast carcinomas are classified as invasive ductal carcinomas, which show a highly variable response to therapy and outcome [1]. The first molecular sub-classification with a major impact on breast cancer research was proposed by Perou and colleagues where the tumors were subdivided according to their pattern of gene expression [2,3]. Five groups were identified and named Luminal A, Luminal B, Basal-like, Normal-like and the HER-2-enriched subgroups. These intrinsic subgroups have been shown to be different in terms of biology, survival and recurrence rate [3,4]. The molecular subgroups have been extended to also include a sixth subgroup which has been named the claudin- low group, based on its low expression level of tight junction genes (the claudin genes) [5]. Different methods for the assignment of individual tumors to its molecular subgroup is proposed; each based on the expression levels of different sets of genes [4,6,7]. The agreement between methods on how to classify individual tumors are not optimal and how to establish

more robust single sample predictors is actively debated [8-11].

Aneuploidy is the presence of an abnormal number of parts of or whole chromosomes and is one feature that clearly separates cancer cells from normal cells. This was proposed as being important in cancer nearly a century ago by Theodor Boveri [12]. With array-based comparative genomic hybridization (aCGH) a genome wide profile of the copy number alterations in the tumor can be obtained. These patterns are related to the molecular subtypes with distinct differences in the number of alterations between the subtypes [13-16]. These copy number alterations (CNAs) alter the dosage of genes and highly influence the level of expression [17,18]. This frequently affects the activity in oncogenes and tumor suppressor genes and in this way CNAs are important for the carcinogenic process. CNAs in tumors are a result of deregulated cell cycle control and of DNA maintenance and repair [19]. Different patterns of copy number alterations have been identified with distinct differences; simplex profiles are characterized by few alterations and complex genomic profiles have extensive changes [20]. Complex genomic rearrangements are areas with high-level amplifications and have prognostic value in breast cancer even when they do not harbor known oncogenes, suggesting that the phenotype of defect DNA-repair may be associated with more aggressive disease [20,21].

Alterations in the expression pattern are caused by changes at the genomic level and a robust classification

* Correspondence: carlos.caldas@cancer.org.uk
[3]Breast Cancer Functional Genomics, Cancer Research UK Cambridge Research Institute, Cambridge, UK
Full list of author information is available at the end of the article

of breast cancer for clinical use should probably take these more into account. Changes at the genomic level include point mutations, changes in copy number and epigenetic events. These are characteristics that enable and drive carcinogenesis together with tumor-promoted inflammation [22].

The era of sequencing of cancer genomes

We are now in the exciting era of full sequencing of cancer genomes. Paired-end sequencing is based on massive parallel sequencing of short stretches of nucleotides at each end of fragmented DNA [23]. The basis of paired-end sequencing technology is shown in Figure 1. Next generation sequencing gives additional information to cancer genomics at many levels, including point mutations, insertions, deletions, copy number and translocations depending on the level of the coverage [23]. The copy number alterations in breast cancer are well characterized by aCGH, but sequencing has given important insight into how alterations are structured given that information on translocations/rearrangements is added [24].

Stephens *et al*. described multiple rearrangement architectures after sequencing 9 breast cancer cell lines and 15 tumors [24]. Intrachromosomal rearrangements were found to be far more frequent than between chromosomes and the most common event was tandem duplications, but with a high degree of variation among tumors. They hypothesized that these extensive alterations are a consequence of a DNA repair defect that leads to a 'mutator phenotype' similar to what causes microsatellite instability in other cancers. Breakpoints tended to fall into areas with microhomology and non-template sequences. Fusion genes are hybrid genes formed from two separate genes (for example, by translocations), which can lead to functional proteins with oncogenic properties. These are important in leukemias and lymphomas, but the role of fusion genes in breast cancer is unclear [25]. Stephens *et al*. found enrichment for alterations within genes and 29 of these were predicted to generate in-frame gene fusions. Transcripts were found for 21 of these, but none of these were recurrent among cancers [24]. Sequencing of the cell line MCF-7 has revealed that breakpoints that are evenly dispersed over the genome tend to be in areas of low copy repeats while the more clustered breakpoints occur close to high-level amplified genes, pointing to different mechanisms for genomic instability [26]. Important point mutations are present already at an early stage, as has been shown in a comparative deep sequencing study of the genomes, and transcriptomes of a primary lobular tumor and its distant metastasis 9.5 years later [27].

The sequencing technology is now capable of sequencing genomes of single cells. As there are heterogeneity among cells of the tumor and infiltration of normal cells and inflammatory cells, picking the right cell to sequence may be challenging. Navin *et al*. sequenced 100 single cells from a polygenic tumor that revealed four distinct groups of genomes; the diploids and the pseudo-diploids (representing normal cells and immune cells), one hypo-diploid and two aneuploid groups [28]. Their analysis suggests that these represent three clonal expansions in the primary tumor as they share many common aberrations. A total of 52 cells from a second tumor and 48 cells from a paired liver metastasis were sequenced and the results indicated that a clonal expansion from a single aneuploid cell had formed the primary tumor and that one of these had metastasized to the liver forming the metastasis.

Deep sequencing of cancer genomes is a costly process and the amount of biological material needed has been a challenge, but technology is moving fast and both cost and tissue demands are continuously decreased. International consortia have formed to do large-scale analysis of cancer genomes at all different levels of large sets of tumors that will provide essential future information on the landscape of cancer genomes [29].
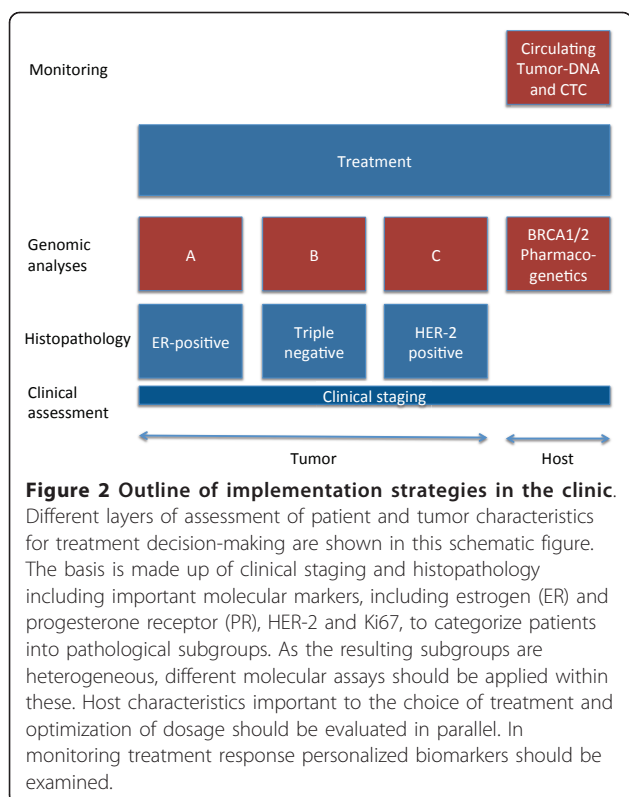
Implementation strategies in the clinic

Molecular classification has had limited implementation in standard clinical treatment guidelines [30,31]. There are two molecular signatures that are approved for clinical use in breast cancer; one microarray-based for fresh frozen tumor material (Mammaprint®, Agendia, Irvine, CA, USA) and one PCR-based for paraffin embedded tumor material (OncotypeDX®, Genomic Health, Inc., Redwood City, CA, USA) [32,33]. The evolving knowledge from molecular classification provides information about disrupted pathways in great detail as well as global changes in expression of genes and genomic alterations. At the same time it is important to acknowledge that existing data for treatment guidelines are based on traditional histopathology and some single molecular markers. To build treatment algorithms that integrate all existing knowledge is currently the challenge.

We believe that the baseline will still be traditional histopathology combined with clinical staging, but with a second layer of molecular classification with subtype specific prognostic and predictive tests (Figure 2). The heterogeneity of breast cancer makes it likely that different tests should be considered in the different clinical settings. Prognostic tests like MammaPrint or Oncotype DX must be validated for such subgroups of patients and their use must be limited to groups where their prognostic power is validated. Such validation in clinically relevant groups of patients is crucial. Many prognostic signatures are published but inadequate validation makes clinical use futile [34].
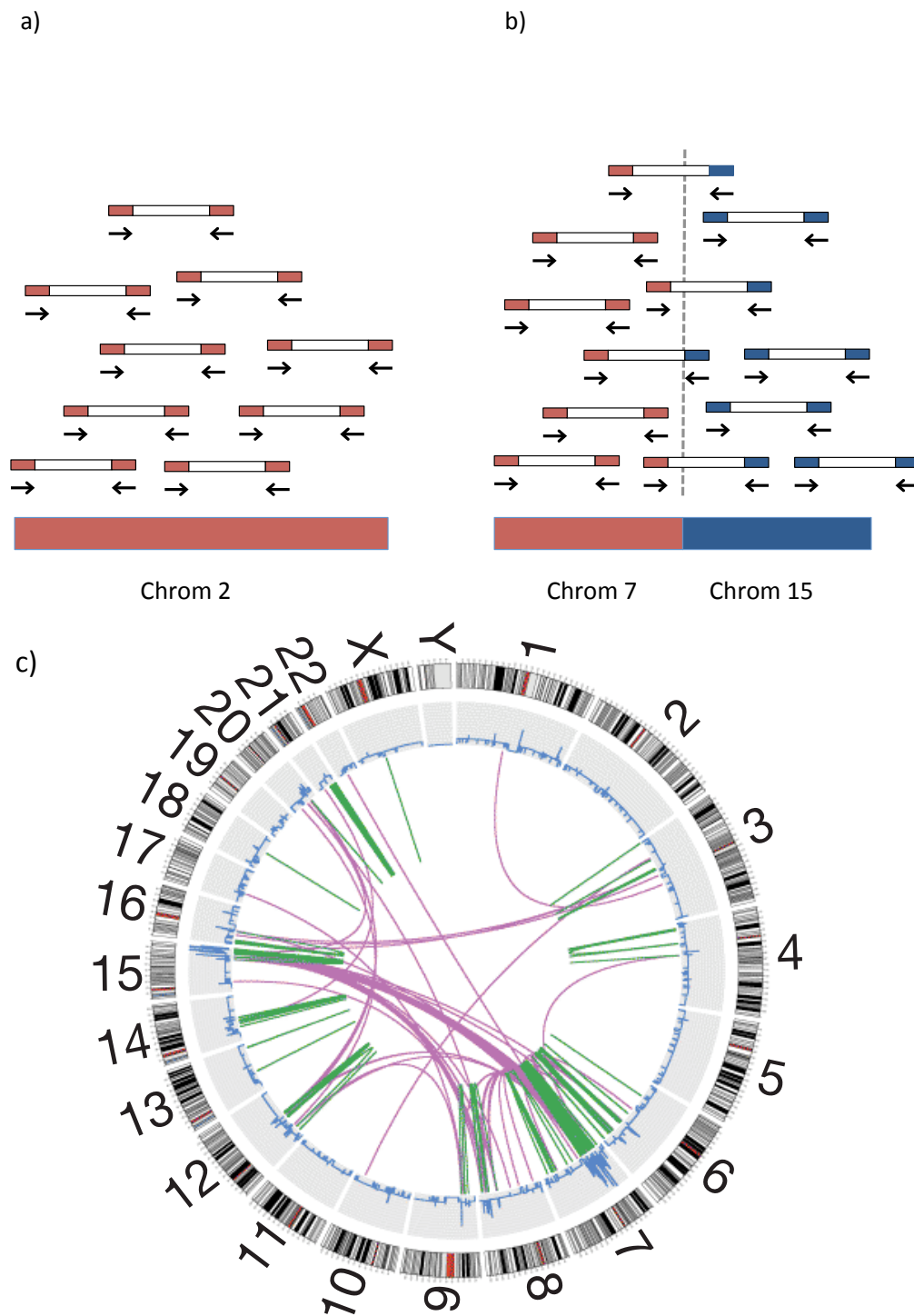
Figure 1 The basis of translocation mapping from paired-end sequencing. **(a)** Paired end sequencing is based on sequencing a short sequence of nucleotides of each end of fragmented and amplified genomic DNA. Reads without the desired length are filtered out. All reads are aligned to a reference genome. The average number of reads per genomic locus is called the coverage of the genome of the sequenced sample. A high coverage (20× to 40×) is needed for detection of point mutations while a much lower coverage is required for other analysis such as copy number and mapping of translocations. The number of reads that map to a locus can be regarded as a function of the number of copies of that locus. As reads can be binned across windows the coverage does not need to be high for such analyses. **(b)** When a part of a chromosome is fused to a part of another chromosome the read from this region will have a sequence in one end that maps to one chromosome and the other end maps to another. When this pattern is consistent over several reads the translocations can be precisely mapped. Intrachromosomal rearrangements are mapped the same way. **(c)** A circos plot of a breast cancer genome. The chromosomes are arranged as a circle from chromosome 1 to the sex chromosomes X and Y. The outer part of the circle shows the chromosomes with cytoband information. The blue line represents the copy number at the given loci. The lines in the middle represent translocations. The inter-chromosomal translocations are in purple and the intra-chromosomal translocations are shown in green. Part (c) is modified from Russnes *et al.* [21].

Figure 2 Outline of implementation strategies in the clinic. Different layers of assessment of patient and tumor characteristics for treatment decision-making are shown in this schematic figure. The basis is made up of clinical staging and histopathology including important molecular markers, including estrogen (ER) and progesterone receptor (PR), HER-2 and Ki67, to categorize patients into pathological subgroups. As the resulting subgroups are heterogeneous, different molecular assays should be applied within these. Host characteristics important to the choice of treatment and optimization of dosage should be evaluated in parallel. In monitoring treatment response personalized biomarkers should be examined.

At present, patient genotype information is not included in treatment of breast cancer. We indicate in Figure 2 that genotype testing in the future should be included parallel to assessment of the tumor. Germline variation in genes involved in drug metabolism may guide the choice of drugs as well as dosage monitoring, as the influence of CYP2D6 variants on Tamoxifen metabolism [35]. Germline mutations leading to deficient proteins (like BRCA1/2) increase the risk of breast cancer, but can also be exploited in therapy. Cells with deficient BRCA have impaired homologous recombination (HR) and are dependent on alternative DNA repair mechanisms. Inhibition of poly ADP ribose polymerase (PARP) leads to the accumulation of multiple DNA double strand breaks and without efficient repair mechanisms the cell dies [36,37]. Such a synthetic lethality approach is a promising therapeutic strategy.

The highly individualized information provided from deep sequencing has the potential to find individualized biomarkers for treatment and disease monitoring [38,39]. Deep sequencing of single cells will give detailed information about the clonal landscape in tumors [28]. It is likely that clonal diversity affects the response to chemotherapy [40]. Targeted therapy approaches have a great potential in oncology, but resistance to the agents is a clinical problem. In colorectal cancer, it has been shown that treatment with Cetuximab, an inhibitor of

EGFR, is ineffective in the presence of an activating mutation of k-ras, a downstream protein in the EGFR signaling pathway [41]. This mechanism of drug resistance is likely to be present for other agents as well.

Deep sequencing of cancer genomes makes it possible to have full mutational information on the important pathways, and methods to characterize the gene sets of mutations are being developed [42,43]. For several of the important carcinogenic pathways several inhibitors exist and more will come. The prospect is, therefore, for better prognostication, prediction and targeted therapy as the main result of full characterization of cancer genomes.

Conclusion
Results from next generation sequencing have the potential for revolutionizing the understanding of malignant disease. The challenge remains in the integration of new results with existing knowledge based on histopathological stratification of breast cancer.

Abbreviations
aCGH: array comparative genomic hybridization; CNA: copy number alterations; HR: homologous recombination; PARP: poly ADP ribose polymerase; PCR: polymerase chain reaction.

Author details
[1]Department of Genetics, Institute for Cancer Research and Department of Breast and Endocrine Surgery, Division of Surgery and Cancer, Oslo University Hospital Radiumhospitalet, 0310 Oslo, Norway. [2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway. [3]Breast Cancer Functional Genomics, Cancer Research UK Cambridge Research Institute, Cambridge, UK. [4]Department of Oncology, University of Cambridge, Li Ka-Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. [5]Cambridge Breast Unit, Addenbrooke's Hospital and Cambridge National Institute for Health Research Biomedical Research Centre, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK.

Authors' contributions
HKMV and CC wrote the paper. Both authors have read and approved the final manuscript

Competing interests
CC is a Section Editor for *BMC Cancer*. The authors declare that they have no competing interests.

References
1. Weigelt B, Reis-Filho JS: **Histological and molecular types of breast cancer: is there a unifying taxonomy?** *Nat Rev Clin Oncol* 2009, **6**:718-730.
2. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale A-L, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.
3. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lønning P, Børresen-Dale A-L: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci USA* 2001, **98**:10869-10874.
4. Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE,

Brown PO, Børresen-Dale A-L, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci USA* 2003, **100**:8418-8423.

5. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: **Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.** *Breast Cancer Res* 2010, **12**:R68.

6. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J Clin Oncol* 2009, **27**:1160-1167.

7. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, Perou CM: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC Genomics* 2006, **7**:96.

8. Weigelt B, Mackay A, A'hern R, Natrajan R, Tan DSP, Dowsett M, Ashworth A, Reis-Filho JS: **Breast cancer molecular profiling with single sample predictors: a retrospective analysis.** *Lancet Oncol* 2010, **11**:339-349.

9. Perou CM, Parker JS, Prat A, Ellis MJ, Bernard PS: **Clinical implementation of the intrinsic subtypes of breast cancer.** *Lancet Oncol* 2010, **11**:718-719, author reply 720-721.

10. Sørlie T, Borgan E, Myhre S, Vollan HKM, Russnes H, Zhao X, Nilsen G, Lingjaerde OC, Børresen-Dale A-L, Rødland E: **The importance of gene-centring microarray data.** *Lancet Oncol* 2010, **11**:719-720, author reply 720-721.

11. Dunning MJ, Curtis C, Barbosa-Morais NL, Caldas C, Tavaré S, Lynch AG: **The importance of platform annotation in interpreting microarray data.** *Lancet Oncol* 2010, **11**:717.

12. Holland AJ, Cleveland DW: **Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis.** *Nat Rev Mol Cell Biol* 2009, **10**:478-487.

13. Bergamaschi A, Kim YH, Wang P, Sørlie T, Hernandez-Boussard T, Lonning PE, Tibshirani R, Børresen-Dale A-L, Pollack JR: **Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer.** *Genes Chromosomes Cancer* 2006, **45**:1033-1040.

14. Chin K, Devries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo W-L, Lapuk A, Neve RM, Qian Z, Ryder T, Chen F, Feiler H, Tokuyasu T, Kingsley C, Dairkee S, Meng Z, Chew K, Pinkel D, Jain A, Ljung BM, Esserman L, Albertson DG, Waldman FM, Gray JW: **Genomic and transcriptional aberrations linked to breast cancer pathophysiologies.** *Cancer Cell* 2006, **10**:529-541.

15. Chin SF, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP, Costa JL, Pinder SE, van de Wiel MA, Green AR, Ellis IO, Porter PL, Tavaré S, Brenton JD, Ylstra B, Caldas C: **High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer.** *Genome Biol* 2007, **8**:R215.

16. Chin S-F, Wang Y, Thorne NP, Teschendorff AE, Pinder SE, Vias M, Naderi A, Roberts I, Barbosa-Morais NL, Garcia MJ, Iyer NG, Kranjac T, Robertson JFR, Aparicio S, Tavaré S, Ellis I, Brenton JD, Caldas C: **Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers.** *Oncogene* 2007, **26**:1959-1970.

17. Chari R, Coe BP, Vucic EA, Lockwood WW, Lam WL: **An integrative multi-dimensional genetic and epigenetic strategy to identify aberrant genes and pathways in cancer.** *BMC Syst Biol* 2010, **4**:67.

18. Jain AN, Chin K, Børresen-Dale A-L, Erikstein BK, Eynstein Lonning P, Kaaresen R, Gray JW: **Quantitative analysis of chromosomal CGH in human breast tumors associates copy number abnormalities with p53 status and patient survival.** *Proc Natl Acad Sci USA* 2001, **98**:7952-7957.

19. Aguilera A, Gómez-González B: **Genome instability: a mechanistic view of its causes and consequences.** *Nat Rev Genet* 2008, **9**:204-217.

20. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V, Yoon S, Wigler M, Ye K, Børresen-Dale A-L, Naume B, Schlicting E, Norton L, Hägerström T, Skoog L, Auer G, Månér S, Lundin P, Zetterberg A: **Novel patterns of genome rearrangement and their association with survival in breast cancer.** *Genome Res* 2006, **16**:1465-1479.

21. Russnes HG, Vollan HKM, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, Sørlie T, Borgen E, Rye IH, Langerød A, Chin S-F, Teschendorff AE,

Stephens PJ, Månér S, Schlichting E, Baumbusch LO, Kåresen R, Stratton MP, Wigler M, Caldas C, Zetterberg A, Hicks J, Børresen-Dale A-L: **Genomic architecture characterizes tumor progression paths and fate in breast cancer patients.** *Sci Transl Med* 2010, **2**:38ra47.

22. Hanahan D, Weinberg RA: **Hallmarks of cancer: the next generation.** *Cell* 2011, **144**:646-674.

23. Campbell PJ, Stephens PJ, Pleasance ED, O'Meara S, Li H, Santarius T, Stebbings LA, Leroy C, Edkins S, Hardy C, Teague JW, Menzies A, Goodhead I, Turner DJ, Clee CM, Quail MA, Cox A, Brown C, Durbin R, Hurles ME, Edwards PAW, Bignell GR, Stratton MR, Futreal PA: **Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing.** *Nat Genet* 2008, **40**:722-729.

24. Stephens PJ, McBride DJ, Lin M-L, Varela I, Pleasance ED, Simpson JT, Stebbings LA, Leroy C, Edkins S, Mudie LJ, Greenman CD, Jia M, Latimer C, Teague JW, Lau KW, Burton J, Quail MA, Swerdlow H, Churcher C, Natrajan R, Sieuwerts AM, Martens JWM, Silver DP, Langerød A, Russnes HEG, Foekens JA, Reis-Filho JS, van' t Veer L, Richardson AL, Børresen-Dale A-L, Campbell PJ, Futreal PA, Stratton MR: **Complex landscapes of somatic rearrangement in human breast cancer genomes.** *Nature* 2009, **462**:1005-1010.

25. Mitelman F, Johansson B, Mertens F: **The impact of translocations and gene fusions on cancer causation.** *Nat Rev Cancer* 2007, **7**:233-245.

26. Hampton OA, Den Hollander P, Miller CA, Delgado DA, Li J, Coarfa C, Harris RA, Richards S, Scherer SE, Muzny DM, Gibbs RA, Lee AV, Milosavljevic A: **A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome.** *Genome Res* 2009, **19**:167-177.

27. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, Delaney A, Gelmon K, Guliany R, Senz J, Steidl C, Holt RA, Jones S, Sun M, Leung G, Moore R, Severson T, Taylor GA, Teschendorff AE, Tse K, Turashvili G, Varhol R, Warren RL, Watson P, Zhao Y, Caldas C, Huntsman D, Hirst M, Marra MA, Aparicio S: **Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution.** *Nature* 2009, **461**:809-813.

28. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, Muthuswamy L, Krasnitz A, McCombie WR, Hicks J, Wigler M: **Tumour evolution inferred by single-cell sequencing.** *Nature* 2011, **472**:90-94.

29. International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, Bernabé RR, Bhan MK, Calvo F, Eerola I, Gerhard DS, Guttmacher A, Guyer M, Hemsley FM, Jennings JL, Kerr D, Klatt P, Kolar P, Kusada J, Lane DP, Laplace F, Youyong L, Nettekoven G, Ozenberger B, Peterson J, Rao TS, Remacle J, Schafer AJ, Shibata T, Stratton MR, *et al*: **International network of cancer genome projects.** *Nature* 2010, **464**:993-998.

30. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, Somerfield MR, Hayes DF, Bast RC, American Society of Clinical Oncology: **American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer.** *J Clin Oncol* 2007, **25**:5287-5312.

31. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ, Panel members: **Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009.** *Ann Oncol* 2009, **20**:1319-1329.

32. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.** *N Engl J Med* 2004, **351**:2817-2826.

33. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**:530-536.

34. Koscielny S: **Why most gene expression signatures of tumors have not been useful in the clinic.** *Sci Transl Med* 2010, **2**:14ps2.

35. Irvin WJ Jr, Walko CM, Weck KE, Ibrahim JG, Chiu WK, Dees EC, Moore SG, Olajide OA, Graham ML, Canale ST, Raab RE, Corso SW, Peppercorn JM, Anderson SM, Friedman KJ, Ogburn ET, Desta Z, Flockhart DA, McLeod HL, Evans JP, Carey LA: **Genotype-guided tamoxifen dosing increases active metabolite exposure in women with reduced CYP2D6 metabolism: a multicenter study.** *J Clin Oncol* 2011, **29**:3232-3239.

36. Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, Santarosa M, Dillon KJ, Hickson I, Knights C, Martin NMB, Jackson SP,

Smith GCM, Ashworth A: **Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.** *Nature* 2005, **434**:917-921.

37. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, Kyle S, Meuth M, Curtin NJ, Helleday T: **Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.** *Nature* 2005, **434**:913-917.

38. Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, Antipova A, Lee C, McKernan K, De La Vega FM, Kinzler KW, Vogelstein B, Diaz LA Jr, Velculescu VE: **Development of personalized tumor biomarkers using massively parallel sequencing.** *Sci Transl Med* 2010, **2**:20ra14.

39. McBride DJ, Orpana AK, Sotiriou C, Joensuu H, Stephens PJ, Mudie LJ, Hämäläinen E, Stebbings LA, Andersson LC, Flanagan AM, Durbecq V, Ignatiadis M, Kallioniemi O, Heckman CA, Alitalo K, Edgren H, Futreal PA, Stratton MR, Campbell PJ: **Use of cancer-specific genomic rearrangements to quantify disease burden in plasma from patients with solid tumors.** *Genes Chromosomes Cancer* 2010, **49**:1062-1069.

40. Gerlinger M, Swanton C: **How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine.** *Br J Cancer* 2010, **103**:1139-1143.

41. Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, Simes RJ, Chalchal H, Shapiro JD, Robitaille S, Price TJ, Shepherd L, Au H-J, Langer C, Moore MJ, Zalcberg JR: **K-ras mutations and benefit from cetuximab in advanced colorectal cancer.** *N Engl J Med* 2008, **359**:1757-1765.

42. Swanton C, Caldas C: **Molecular classification of solid tumours: towards pathway-driven therapeutics.** *Br J Cancer* 2009, **100**:1517-1522.

43. Boca SM, Kinzler KW, Velculescu VE, Vogelstein B, Parmigiani G: **Patient-oriented gene set analysis for cancer mutation data.** *Genome Biol* 2010, **11**:R112.