Published in final edited form as:
*Curr Opin Genet Dev*. 2008 February ; 18(1): 68–72. doi:10.1016/j.gde.2008.01.005.

# High-Content Analysis of Cancer Genome DNA Alterations

**Yan Y. Degenhardt**[1], **Richard Wooster**[1], **Richard W. McCombie**[2], **Robert Lucito**[2], and **Scott Powers**[2,3]

[1] Translational Medicine, GlaxoSmithKline, King of Prussia, PA, 19406, USA

[2] Cancer Genome Center, Cold Spring Harbor Laboratory, Woodbury, NY, 11797, USA

## Summary

New technologies as well as concerted brute-force approaches have increased the content (number of genes) that can be characterized for genomic DNA alterations. Recent advances include the detection of activating point mutations in key kinase genes (*BRAF*, *EGFR*, *PIK3CA*) in multiple cancer types; preliminary insight into the entire repertoire of genes that can be mutated in cancer; the discovery of new oncogenes by high-resolution profiling of DNA copy number alterations; and the bioinformatic-driven discovery of oncogenic gene fusions. High-content promoter methylation detection systems have been used to discover additional methylated genes and have provided evidence for a stem cell origin for certain tumors. Some of these advances have had significant impact on the development and clinical testing of new therapeutics.

## Introduction

There is considerable interest in being able to detect, catalog, and understand the entire set of DNA alterations in human cancers. The genetic model of cancer places primary importance on these changes, which result from recurring cycles of genomic alteration followed by clonal selection, and which predicts that the properties of a particular individual's tumor are dictated by the specific DNA alterations within that tumor (the cancer genotype). Support for this model comes from experiments and clinical results that demonstrate that reversing the effect of even a single DNA alteration - such as inhibiting the aberrant tyrosine kinase activity of the BCR-ABL gene fusion protein by the small-molecule inhibitor imatinib - can have dramatic effects on the ability of the tumor cell to proliferate or even survive [1]. However the cancer genotype is not the only determinant, with the intrinsic properties of the tumor-initiating cell and the tumor microenvironment also playing key roles in cancer progression [2,3].

The cancer genotype includes both hereditary predisposition genes and the genes that are somatically altered by a variety of mechanisms including mutation, amplification, deletion, and translocation. Epigenetic alterations caused by DNA methylation or modifications of histones can also comprise a substantial portion of the final set of cancer-relevant DNA alterations.

In the earlier part of this decade, the only genome-wide technology for profiling the cancer genome was comparative genomic hybridization (CGH) (Box 1). Now there are several means

3contributing author email addresses: yan.y.degenhardt@gsk.com, richard.f.wooster@gsk.com, lucito@cshl.edu, mccombie@cshl.edu, powers@cshl.edu.

to study DNA alterations on a genome-wide scale, including array CGH with very high genome resolution [4], new methylation detection systems [5], and last but not least, single-molecule sequencing technologies [6] (Box 1). In this review we focus on the recent biological discoveries that have been made with these new tools along with the large-scale efforts using older tools.

## Mutational Profiling (Resequencing)

Large-scale Sanger dideoxy sequencing of targeted regions by the Sanger Institute and the Johns Hopkins group has led to two basic discoveries that both have direct translational relevance. The first was the discovery that *BRAF* is frequently (40%) activated by point mutations in melanomas (Boxes 1,2) [7]. This is the second cancer characterized by frequent genomic activation of a kinase, with chronic myeloid leukemia (CML) and the *BCR-ABL* translocation being the first. The success of kinase inhibitor therapy in CML has created intense interest in testing Braf pathway inhibitors in melanoma patients. The second discovery was that the *PIK3CA* gene is frequently mutated in colorectal and other cancers (Box 2) [8]. If one includes mutations in all the pathway genes *(PTEN*, *PDK1*, *AKT2* and *PAK4*), the phosphatidylinositol–3-OH kinase pathway is mutationally activated in 40% of colorectal cancers [9]. These results have stimulated development and clinical testing of inhibitors of this pathway.

Independently of these large-scale efforts, several laboratories discovered that mutations in *EGFR* occur frequently in lung cancer and predict a positive clinical response to EGFR inhibitors (Box 2) [10–12]. The frequency ranges from 10% in Caucasians to 40% in East Asian lung cancer patients [13]. Other *EGFR* pathway genes are mutated in lung cancer including *KRAS* (20%), and *ERBB2*, *BRAF*, and *PIK3CA* (each 2–4%) [14]. Pao and co-workers discovered that lung cancers that harbor mutations in *KRAS*, which encodes a downstream signaling component in the EGFR pathway, were resistant to treatment with EGFR inhibitors [15]. Another clear example of a genomic predictor of a negative response to a targeted therapeutic has recently been reported: oncogenic *PIK3CA* mutations or low *PTEN* expression in breast tumors are negative predictors for response to trastuzumab, the anti-Her2 monoclonal antibody used to treat breast cancer patients who harbor amplification of the *HER2/ERBB2* gene [16].

The first genome-wide resequencing effort, which initially involved the coding regions of 14,000 genes [17] and was later expanded to include all of the reference sequence genes [18], has triggered controversy. The major issue is the claim that the Johns Hopkins group has discovered approximately 200 new cancer genes. Other groups claim that their statistical analysis was flawed and that in particular the small sample size (~20 tumors) is too small to distinguish driver mutations from passenger mutations [19].

However, the Sanger Institute independently came to a similar conclusion -that the repertoire of cancer genes that are altered by mutations is substantially larger than previously thought (Box 2) - after sequencing a much larger set of tumors (~ 200) for 518 protein kinase genes [20]. The authors addressed biological significance in their study by employing statistical analyses that are used in genome evolution studies, the premise of which is that the higher the ratio of non-synonymous to synonymous mutations, the more likely there has been biological selection. The number of kinase genes that can be mutated into cancer driver genes was estimated to be within 52 to 149, which is two to five times larger than the set of known cancer-causing kinase genes (~ 25) [21]. Many of these mutations occurred in the same functionally important protein regions that are mutated in *BRAF*, providing additional evidence that they contribute to the cancer phenotype [20].

It is noteworthy that none of the 20,000 genes examined by the Johns Hopkins group are mutated at frequencies greater than 10% in colorectal or breast cancer besides those previously identified (*TP53*, *PIK3CA*, *APC*, *KRAS*, *SMAD4*, and *CDC4*) [17,18]. Assuming for the moment that there really is significant diversity in the set of genes that can be mutated to become driver genes in cancer, how will we move forward to develop target therapeutics? One possibility is to use pathway analysis to look for commonality. Along these lines, the set of genes rarely mutated in breast cancer are significantly enriched for components of the NF-kB pathway [18]. Thus NF-kB pathway inhibitors might be useful to treat a subclass of breast cancer patients identified as having mutations in NF-kB pathway genes, or, as we will see in the section that follows, amplification of the *IKBKE* gene.

## Copy Number Alteration Profiling

Array-based comparative genomic hybridization (aCGH) studies now number over 500 but for the most part have not addressed functional relevance of the altered genes. Two approaches have emerged as being particularly relevant to cancer biology. One has been to focus on DNA copy number alterations that are tumor-type specific. *MITF* was identified and subsequently validated as a melanoma-specific amplified oncogene based on its selective amplification in melanoma [22]. Similarly, three genes (*TTF1*, *NKX2-8*, and *PAX9*) that are within a 150 kb region on 14q13 were found to be determinants of a lung-cancer specific amplicon [23]. All three of these genes act synergistically to promote and maintain tumorigenicity [23]. Occasionally, the *TTF1* gene is amplified in the absence of *NKX2-9* and *PAX9* [24]. Both of these amplicons are relatively frequent oncogenetic events: *MITF* is amplified in 20% of metastatic melanoma [22], and the *TTF1/NKX2-8/PAX9* amplicon is the second most common amplicon in lung cancer (following *MYC*) and is amplified in up to 20% of lung cancers [23, 24] (Box 2).

Another productive approach that has emerged is to compare human DNA copy number alterations to those that occur during cancer progression in mouse models. The 11q22 amplicon, which is found in several tumor types including liver cancer, was found to be syntenic with an amplicon located at 9qA1 in murine model of liver cancer [25]. Two genes, *cIAP1* and *Yap*, were overexpressed in amplified tumors from both species and acted cooperatively to accelerate and maintain tumorigenesis [25]. Similarly, the *Nedd9* gene was found within a small amplicon in metastatic mouse melanomas and is amplified in human melanomas as part of a large gain of the p arm of chromosome 6 [26]. *Nedd9* encodes a signaling protein that enhances invasion and metastasis when overexpressed in melanocytes [26]. In the future, introduction of genomic instability into mouse models will enable increased utility of this cross-species approach [27].

Recently, *RAB25* on 1q22 and *IKBKE* on 1q32 were shown to be functional oncogenes in ovarian and breast cancers by both overexpression and RNAi studies [28,29]. Presumably, both *RAB25* and *IKBKE*, the latter encoding an IkappaB kinase that activates the NF-kB pathway, are likely driver genes for the gain of the q arm of chromosome 1, which is one of the most frequent genetic events in human cancer – occurring in 50% of breast, liver, lung, ovarian and other tumor types. As of today, there is no systematic approach to identifying the driver genes for such large chromosomal gains or losses. This is a major black box in our understanding of the oncogenetic events that underlie cancer progression.

In an innovative approach to finding the genetic underpinnings of a molecularly defined subclass of tumors, Chang and co-workers determined that the co-amplification and co-overexpression of *MYC* and a subunit of the COP9 signalosome (*COPS5*) underlies the poor-prognosis wound signature seen in a subset of primary breast cancers [30]. Co-expression of

both *MYC* and *COPS5* was necessary and sufficient to induce the wound repair signature in a normal mammary epithelial cell line.

Copy number profiling has also been used to look for molecular subclasses of common cancers. Profiling can distinguish three molecular subclasses of gliomas, one of which corresponds to a known clinical subtype but the other two are novel [31]. In breast cancer, where gene expression profiling has defined new molecular subclasses, copy number profiling has shown that within different gene-expression subclasses, tumors harboring amplicons show poorer prognosis than those that do not [32]. Multiple closely spaced amplicons on the same chromosome are highly correlated with aggressive disease and poor survival [33].

## Profiling Translocations and Gene Fusions

Gene fusions play key roles in the initial steps of hematological cancers and childhood sarcomas [34]. Over 300 genes are affected by gene fusion events in cancer and the majority of these were identified in hematological cancers. Although there are particular hematological cancers in which almost 100% of the tumors will have a specific gene fusion, e.g. *BCR-ABL* in CML or *PML-RARA* in acute promyelocytic leukemia, the frequency and patterns of gene fusions in hematological cancers is better typified by acute myeloid leukemia (AML). Only 20% of AML cases contain translocations, and these translocations are very diverse and constitute 109 different gene fusions [34]. Thus it is important keep these extremes in mind when considering the potential of discovering gene fusions in more common epithelial tumors.

In a truly amazing line of investigation, Chinnaiyan and colleagues analyzed the expression patterns of gene fusions in hematological cancers and determined an algorithm for finding candidate gene fusions using RNA expression data [35]. Among the top 10 candidate prostate cancer fusion genes were *ERG* and *ETV1* —transcription factor genes known to be involved in fusions in other tumor types. They found that these genes were fused to 5′ promoter of the highly androgen-responsive gene *TMPRSS2*. Up to 80% of primary prostate cancers harbor *TMPRSS2* fused to *ERG*, *ETV1*, or *ETV4* (Box 2). These fusions are absent in normal prostate, proliferative inflammatory prostate, and benign prostate hyperplasia; but are detected in 20% of early neoplastic prostate lesions [36].

In addition, there was a recent serendipitous discovery of a gene fusion involving a tyrosine kinase gene (*ALK*) affecting 7% of non-small cell lung cancers (7% in an Asian population – this frequency is likely to be different in other races). This fusion, *EML4-ALK*, was detected in a cDNA library from a lung cancer following a transformation screen in NIH-3T3 cells [37]. This and the prostate cancer results raise the question as to why these gene fusions have not been detected beforehand? The answer appears to be the much greater technical and analytic problems associated with cytogenetic analysis of epithelial tumors compared to that of hematological cancers [34]. Thus there may be many more epithelial cancer gene fusions that remain to be discovered, either by informatic approaches or by experimental approaches such as high-throughput sequencing of both 5′ and 3′ ends of transcripts.

## Epigenetic Profiling

The importance of epigenetic alterations in cancer progression was shown years ago when methylation of the 5′ CpG-island of the *p16/CDK2NA* gene was proven to be responsible for its transcriptional silencing in 20%–40% of most common cancers [38,39]. Several cancer susceptibility genes, including *BRCA1* and *VHL*, which cause familial forms of breast and kidney cancer, respectively, are silenced by methylation in a significant percentage of sporadic forms of the respective tumor types. 15% of sporadic breast cancers harbor methylated *BRCA1* genes and their gene-expression profiles are identical to those of tumors from inherited

families where *BRCA1* is mutated; both are completely distinct from those of other breast-cancer types [40].

Several groups have been developing array-based methods for genome-wide detection of methylation or other epigenetic alterations such as histone modifications [5,41,42]. These emerging methods have not yet yielded the depth of biological insight that methods that are lower in gene content, but technologically more robust, have yielded. Measurement of ~200 CpG islands in a panel of ~300 colorectal cancers with a PCR-based technology has provided convincing evidence that a subset of colorectal cancers – approximately 15%-have significantly greater CpG island methylation than other colon cancers, and that this subset is characterized by a methylated mismatch repair gene *MLH1* and the sporadic form of microsatellite instability. Thus it appears that a methylator type of genomic instability proceeds the development of microsatellite instability (Box 2) [43]. This same platform was used to determine that stem cell Polycomb group targets are greater than 10-fold more likely to have cancer-specific promoter DNA hypermethylation than non-targets, supporting a stem cell origin of cancer in which reversible gene repression is replaced by permanent silencing [44].

## Conclusion

Within the next five years it is likely that we will be able to detect and catalog the entire set of DNA alterations in a given cancer, which will be a significant milestone in the history of cancer research. New sequencing technologies will make it possible to sequence hundreds of cancer genomes, and will also enable ultimate resolution and genome coverage for detecting RNA expression, copy number and epigenetic alterations, as well as gene fusions. What will clearly lag behind is an understanding of the functional importance of all the DNA alterations, which will take several additional decades. In the meantime, cancer genome profiling will have important applications in pinpointing new targets, discovering resistance mutations to existing therapies, and discovering both positive and negative genomic predictors of response to specific therapies.

**Box 1. Milestones in Cancer Genome DNA Profiling**

| | |
|---|---|
| 1998 – 1999 | Development of 1st generation arrays for profiling DNA copy number alterations |
| 2001 | Development of 1st generation arrays for profiling epigenetic alterations |
| 2002 | First major finding reported from large-scale resequencing projects |
| 2004 | Development of 1st generation single-molecule sequencers |
| 2005 | Bioinformatic approach developed that enables discovery of gene fusions |

**Box 2. Major Discoveries Resulting from Cancer Genome DNA Profiling**

- Frequent mutational activation of *EGFR*, *PIK3CA*, and *BRAF* [7,8,11]
- Frequent translocations and interstitial deletions causing gene fusions in prostate cancer [35]
- Strong evidence for a methylator phenotype in colorectal cancer [43]
- Discovery of frequent tumor-type specific amplicons in melanoma and lung [22, 23,24]

- Substantial diversity in the set of genes affected by putative driver mutations [17,20]

## References


1. Varmus H. The new era in cancer research. Science 2006;312:1162–1165. [PubMed: 16728627]
2. Ince TA, Richardson AL, Bell GW, Saitoh M, Godar S, Karnoub AE, Iglehart JD, Weinberg RA. Transformation of different human breast epithelial cell types leads to distinct tumor phenotypes. Cancer Cell 2007;12:160–170. [PubMed: 17692807]
3. Karnoub AE, Dash AB, Vo AP, Sullivan A, Brooks MW, Bell GW, Richardson AL, Polyak K, Tubo R, Weinberg RA. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. Nature 2007;449:557–563. [PubMed: 17914389]
4. Stallings RL, Nair P, Maris JM, Catchpoole D, McDermott M, O'Meara A, Breatnach F. High-resolution analysis of chromosomal breakpoints and genomic instability identifies PTPRD as a candidate tumor suppressor gene in neuroblastoma. Cancer Res 2006;66:3673–3680. [PubMed: 16585193]
5. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, Doucet D, Thomas NJ, Wang Y, Vollmer E, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res 2006;16:383–393. [PubMed: 16449502]
6. Shendure J, Mitra RD, Varma C, Church GM. Advanced sequencing technologies: methods and goals. Nat Rev Genet 2004;5:335–344. [PubMed: 15143316]
7. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, et al. Mutations of the BRAF gene in human cancer. Nature 2002;417:949–954. [PubMed: 12068308]
8. Samuels Y, Velculescu VE. Oncogenic mutations of PIK3CA in human cancers. Cell Cycle 2004;3:1221–1224. [PubMed: 15467468]
9. Parsons DW, Wang TL, Samuels Y, Bardelli A, Cummins JM, DeLong L, Silliman N, Ptak J, Szabo S, Willson JK, et al. Colorectal cancer: mutations in a signalling pathway. Nature 2005;436:792. [PubMed: 16094359]
10. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–2139. [PubMed: 15118073]
11. Riely GJ, Politi KA, Miller VA, Pao W. Update on epidermal growth factor receptor mutations in non-small cell lung cancer. Clin Cancer Res 2006;12:7232–7241. [PubMed: 17189394]
12. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–1500. [PubMed: 15118125]
13. Thomas RK, Weir B, Meyerson M. Genomic approaches to lung cancer. Clin Cancer Res 2006;12:4384s–4391s. [PubMed: 16857815]
14. Marks JL, McLellan MD, Zakowski MF, Lash AE, Kasai Y, Broderick S, Sarkaria IS, Pham D, Singh B, Miner TL, et al. Mutational analysis of EGFR and related signaling pathway genes in lung Adenocarcinomas identifies a novel somatic kinase domain mutation in FGFR4. PLoS ONE 2007;2:e426. [PubMed: 17487277]
15•. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q, Ladanyi M, Zakowski MF, Heelan RT, Kris MG, Varmus HE. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. PLoS Med 2005;2:e17. [PubMed: 15696205]A clear demonstration in human tumors that gain-of-function mutations in a gene encoding a downstream signaling component of the EGFR pathway confers resistance to EGFR inhibitors
16. Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K, Linn SC, Gonzalez-Angulo AM, Stemke-Hale K, Hauptmann M, et al. A Functional Genetic Approach Identifies the PI3K Pathway as a Major Determinant of Trastuzumab Resistance in Breast Cancer. Cancer Cell 2007;12:395–402. [PubMed: 17936563]

17•. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, et al. The consensus coding sequences of human breast and colorectal cancers. Science 2006;314:268–274. [PubMed: 16959974]The results of the first cancer genome resequencing project that surveyed (most) all known genes. Although controversial, the finding that there are infrequent cancer-relevant mutations in a much larger set of genes than previously imagined in consistent with another report highlighted below

18. Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J, et al. The Genomic Landscapes of Human Breast and Colorectal Cancers. Science. 2007

19. Getz G, Hofling H, Mesirov JP, Golub TR, Meyerson M, Tibshirani R, Lander ES. Comment on "The consensus coding sequences of human breast and colorectal cancers". Science 2007;317:1500. [PubMed: 17872428]

20•. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, et al. Patterns of somatic mutation in human cancer genomes. Nature 2007;446:153–158. [PubMed: 17344846]The results of a major cancer genome resequencing project that surveyed all kinase genes in a relatively large sample set. Viewed together with the report highlighted above, this argues that there is larger set of genes than previously imagined that have oncogenic potential in human cancer

21. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR. A census of human cancer genes. Nat Rev Cancer 2004;4:177–183. [PubMed: 14993899]

22. Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S, Beroukhim R, Milner DA, Granter SR, Du J, et al. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. Nature 2005;436:117–122. [PubMed: 16001072]

23•. Kendall J, Liu Q, Bakleh A, Krasnitz A, Nguyen KC, Lakshmi B, Gerald WL, Powers S, Mu D. Oncogenic cooperation and coamplification of developmental transcription factor genes in lung cancer. Proc Natl Acad Sci U S A. 2007Perhaps the most dramatic demonstration of cooperating oncogenes within a single amplicon (the 14q13 amplicon in lung cancer), this paper highlights the increasing importance of performing functional analysis of cancer-relevant genes in parallel rather than individually

24•. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R, Lin WM, Province MA, Kraja A, Johnson LA, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature 2007;450:893–898. [PubMed: 17982442]The most comprehensive and extensive copy number alteration survey of human cancers to-date, this paper also describes the 14q13 amplicon present in a large percentage of lung cancers

25. Zender L, Spector MS, Xue W, Flemming P, Cordon-Cardo C, Silke J, Fan ST, Luk JM, Wigler M, Hannon GJ, et al. Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. Cell 2006;125:1253–1267. [PubMed: 16814713]

26. Kim M, Gans JD, Nogueira C, Wang A, Paik JH, Feng B, Brennan C, Hahn WC, Cordon-Cardo C, Wagner SN, et al. Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. Cell 2006;125:1269–1281. [PubMed: 16814714]

27. Maser RS, Choudhury B, Campbell PJ, Feng B, Wong KK, Protopopov A, O'Neil J, Gutierrez A, Ivanova E, Perna I, et al. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. Nature 2007;447:966–971. [PubMed: 17515920]

28. Cheng KW, Lahad JP, Kuo WL, Lapuk A, Yamada K, Auersperg N, Liu J, Smith-McCune K, Lu KH, Fishman D, et al. The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers. Nat Med 2004;10:1251–1256. [PubMed: 15502842]

29. Boehm JS, Zhao JJ, Yao J, Kim SY, Firestein R, Dunn IF, Sjostrom SK, Garraway LA, Weremowicz S, Richardson AL, et al. Integrative genomic approaches identify IKBKE as a breast cancer oncogene. Cell 2007;129:1065–1079. [PubMed: 17574021]

30. Adler AS, Lin M, Horlings H, Nuyten DS, van de Vijver MJ, Chang HY. Genetic regulators of large-scale transcriptional signatures in cancer. Nat Genet 2006;38:421–430. [PubMed: 16518402]

31. Maher EA, Brennan C, Wen PY, Durso L, Ligon KL, Richardson A, Khatry D, Feng B, Sinha R, Louis DN, et al. Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. Cancer Res 2006;66:11502–11513. [PubMed: 17114236]

32. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, Lapuk A, Neve RM, Qian Z, Ryder T, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006;10:529–541. [PubMed: 17157792]

33. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V, et al. Novel patterns of genome rearrangement and their association with survival in breast cancer. Genome Res 2006;16:1465–1479. [PubMed: 17142309]

34. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer 2007;7:233–245. [PubMed: 17361217]

35••. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–648. [PubMed: 16254181]This paper represents a milestone in cancer genomics as it represents a truly creative and important development in the application of statistics and informatics to derive biologically important information - in this case a very frequent genetic alteration occuring in prostate cancers - from large cancer genome datasets

36. Perner S, Mosquera JM, Demichelis F, Hofer MD, Paris PL, Simko J, Collins C, Bismar TA, Chinnaiyan AM, De Marzo AM, et al. TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. Am J Surg Pathol 2007;31:882–888. [PubMed: 17527075]

37. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, Fujiwara S, Watanabe H, Kurashina K, Hatanaka H, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007;448:561–566. [PubMed: 17625570]

38. Herman JG, Merlo A, Mao L, Lapidus RG, Issa JP, Davidson NE, Sidransky D, Baylin SB. Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. Cancer Res 1995;55:4525–4530. [PubMed: 7553621]

39. Merlo A, Herman JG, Mao L, Lee DJ, Gabrielson E, Burger PC, Baylin SB, Sidransky D. 5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. Nat Med 1995;1:686–692. [PubMed: 7585152]

40. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet 2002;3:415–428. [PubMed: 12042769]

41. Callinan PA, Feinberg AP. The emerging science of epigenomics. Hum Mol Genet 2006;15(Spec No 1):R95–101. [PubMed: 16651376]

42. Wu J, Smith LT, Plass C, Huang TH. ChIP-chip comes of age for genome-wide functional analysis. Cancer Res 2006;66:6899–6902. [PubMed: 16849531]

43. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787–793. [PubMed: 16804544]

44•. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, et al. Epigenetic stem cell signature in cancer. Nat Genet 2007;9:157–158. [PubMed: 17200673]This study shows that Genes that are reversibly repressed in embryonic stem cells by Polycomb proteins are over-represented among genes that are permanently silenced in colon cancer; this link lends support to the theory that the tumor initiating cells of colon cancer are stem cells