Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

Integrated Systems and Technologies: Mathematical Oncology

Cancer Research

Selection of Personalized Patient Therapy through the Use of Knowledge-Based Computational Models That Identify Tumor-Driving Signal Transduction Pathways

Wim Verhaegh[1], Henk van Ooijen[1], Márcia A. Inda[1], Pantelis Hatzis[2], Rogier Versteeg[3], Marcel Smid[4], John Martens[4], John Foekens[4], Paul van de Wiel[1], Hans Clevers[2], and Anja van de Stolpe[1]

Abstract

Increasing knowledge about signal transduction pathways as drivers of cancer growth has elicited the development of "targeted drugs," which inhibit aberrant signaling pathways. They require a companion diagnostic test that identifies the tumor-driving pathway; however, currently available tests like estrogen receptor (ER) protein expression for hormonal treatment of breast cancer do not reliably predict therapy response, at least in part because they do not adequately assess functional pathway activity. We describe a novel approach to predict signaling pathway activity based on knowledge-based Bayesian computational models, which interpret quantitative transcriptome data as the functional output of an active signaling pathway, by using expression levels of transcriptional target genes. Following calibration on only a small number of cell lines or cohorts of patient data, they provide a reliable assessment of signaling pathway activity in tumors of different tissue origin. As proof of principle, models for the canonical Wnt and ER pathways are presented, including initial clinical validation on independent datasets from various cancer types. *Cancer Res; 74(11); 2936–45. ©2014 AACR.*

Major Findings

As expected, the Wnt pathway was predicted inactive in normal colon samples and active in 97% of tested colon adenomas and carcinomas (known to be Wnt driven) and in all tested medulloblastomas containing an activating β-catenin mutation. Furthermore, in primary liver cancer, Wnt activity was predicted in 56% of samples containing a β-catenin mutation, against 18% without documented mutations. In breast cancer, Wnt activity was predicted in 30% of basal-type breast cancers versus 7% of other subtypes, confirming expectations based on clinical research.

The estrogen receptor (ER) pathway model predicted inactivity in practically all tested samples, except for 39% of ER-positive breast cancer samples. Furthermore, ER pathway activity was associated with increased disease-free survival compared with patients in which the pathway was predicted inactive, even though the model was not trained for this purpose.

Clinical implementation of our models is expected to enable a more informed choice of therapy and improved prediction of targeted therapy response. Furthermore, it may help to focus search for tumor-driving genomic defects in whole-genome sequencing data to only those genes involved in the aberrantly active pathway(s).

Introduction

Knowledge on intracellular signal transduction pathways governing cancer cell behavior and controling cell division is rapidly increasing. This development has elicited a paradigm shift toward development of a whole new category of "targeted drugs," aiming to target the aberrant signaling pathway, which drives tumor growth in the individual patient with cancer (1). In contrast with conventional chemotherapy, targeted therapy requires a highly personalized approach to treatment choice, in principle based on predicting treatment response before administering the drug or drug combination of choice.

The anticipated increasing availability of targeted drugs (2) stresses the need for reliable companion diagnostics to predict therapy response, for which identification of the tumor-driving signaling pathway and the underlying defect that causes its aberrant activation is of high importance (3). Unfortunately, currently available tests often lack predictive value with respect to targeted therapy response. In general, these tests demonstrate (over-)expression of key proteins in signaling pathways of interest, e.g., estrogen receptor (ER) or HER2 in breast cancer, or identify DNA mutations (e.g., in the *PIK3CA* gene) or structural changes (like HER2 coding gene amplification) in genes encoding

Authors' Affiliations: [1]Molecular Diagnostics, Philips Research, Eindhoven; [2]Hubrecht Institute, Utrecht; [3]Human Genetics, AMC, Amsterdam; and [4]Medical Oncology, Erasmus MC, Rotterdam, the Netherlands

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Current address for P. Hatzis: BSRC Alexander Fleming, Vari, Greece.

Corresponding Author: Wim Verhaegh, Philips Research, High Tech Campus 11, 5656 AE Eindhoven, the Netherlands. Phone: 31-40-2747521: Fax: 31-40-2742944: E-mail: wim.verhaegh@philips.com

doi: 10.1158/0008-5472.CAN-13-2515

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

Quick Guide to Equations and Assumptions

Major assumptions

The key assumption in the Bayesian pathway model we present is that functional activity of a signaling pathway is determined by activity of its respective transcription complex, while the latter can be inferred from mRNA expression data of its transcriptional target genes.

Furthermore, we assumed that the oncogenic signaling pathway driving tumor growth is not transiently and dynamically activated, but long term or even irreversibly. Hence, the model was developed for interpretation of a static cellular condition, and complex dynamic pathway features were not incorporated.

Finally, to handle the noisy characteristics of mRNA expression data from microarrays, as well as additional unknown biologic factors, we modeled signaling pathways in a probabilistic manner using Bayesian networks.

Equations

The main equations related to our Bayesian models describe probabilistic relations between different types of nodes. Conditional probability tables describing the relation between transcription complex node TC and target gene nodes TG_i are given by Equation 1, and between each target gene node TG_i and its respective probeset nodes $PS_{i,j}$ by Equations 2a and 2b.

for proteins that directly influence signaling pathway activity. However, though associated with response to targeted drugs (hormonal therapy and trastuzumab in example cases ER and HER2, respectively), such tests do not provide conclusive information on the functional activity status of the associated signaling pathways (4).

Around 10 major oncogenic signaling pathways play a role in tumor growth and metastasis: ER and androgen receptor pathways, phosphoinositide 3-kinase (activated by multiple growth factor receptors, like HER2, EGFR), canonical Wnt, Notch, Hedgehog, TGFβ, NF-κB, VEGF, RAS/MAPK/ERK, and FGF signaling pathways (5); see also Supplementary Major Oncogenic Signaling Pathways. Whole-genome and transcriptome analysis methods, like DNA and RNA sequencing and microarray technologies, are in principle capable of producing all data necessary to extract information on signaling pathway activity from cancer tissue samples, on the premise that adequate software is available to interpret the highly complex data (6). This has proved to be a tremendous challenge, and most efforts are directed toward identification of genotypic changes, instead of using the transcriptome to provide information on the functional phenotype of the cancer cell—which is determined in concert by both genotype and microenvironment of cancer cells (7).

We describe the development and partial clinical validation of a new type of knowledge-based probabilistic computational modeling framework for oncogenic cell signaling pathways, which enables functional assessment of pathway activity in individual tissue samples based on quantitative transcriptome data as input. The models have been built using Affymetrix HG-U133Plus2.0 data, but can be calibrated to other quantitative mRNA data formats like RNA sequencing or other microarray types. First models of canonical Wnt and ER pathways are presented as proof of principle for prediction of pathway activity, therapy response, and prognosis in patients with cancer.

Materials and Methods

Development of Bayesian models for signal transduction pathways

Our signal transduction pathway modeling approach is based on inferring pathway activity from the expression profile of its target genes using probabilistic Bayesian network inference. Bayesian networks were built using the Bayes Net Toolbox for MATLAB, as detailed in Supplementary Methods. The Bayesian network structure used as a basis for our modeling approach (Fig. 1) is a simplified model of the transcriptional program of a cellular signal transduction pathway, consisting of three types of nodes: (i) transcription complex TC (with states 'absent' and 'present'), (ii) target genes TG_i (with states 'down' and 'up'), and (iii) microarray probesets $PS_{i,j}$ (with states 'low' and 'high') corresponding to target genes. The model describes (i) how expression of target genes depends on transcription complex activation and (ii) how probeset intensities in turn depend on expression of the respective target genes. For the latter, probeset intensities are taken from frozen Robust Multi-array Analysis preprocessed Affymetrix HG-U133Plus2.0 microarrays, widely available from Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) and ArrayExpress (www.ebi.ac.uk/arrayexpress); an overview of used datasets is provided in Supplementary Table S3.

As our pathway models are a simplification of signaling pathway biology and as biologic measurements are typically noisy,

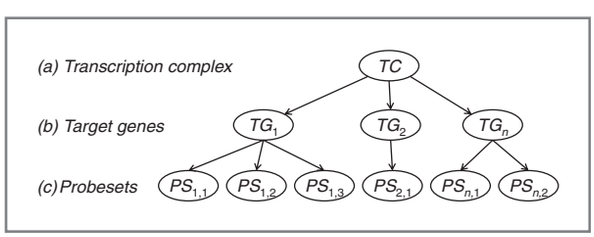

Figure 1. The structure of the Bayesian networks used to model the transcriptional program of signaling pathways.

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

we opted for a probabilistic approach, meaning that relationships (i) between transcription complex and target genes and (ii) between target genes and their respective probesets are described in probabilistic terms. Furthermore, we assumed that the oncogenic signaling pathway driving tumor growth is not transiently and dynamically activated, but long term or even irreversibly. Hence, the model was developed for interpretation of a static cellular condition, and complex dynamic pathway features were not incorporated.

Once the Bayesian network model has been built and calibrated for a particular signaling pathway (sketched below and further detailed in Supplementary Methods), the model can be used on microarray data of a new tumor sample by entering probeset measurements as observations in the bottom layer, and inferring backwards in the model the activity probability of the transcription complex. Hence, this latter probability is the primary read-out used to indicate pathway activity, which can be translated into odds of the pathway being active by taking the ratio of the probability of being active versus inactive (i.e., odds are given by $p/(1-p)$ if p is the predicted probability of being active).

Target gene selection

For optimal performance, the Bayesian network models should contain (only) direct target genes of the respective pathways. Unfortunately, pathway databases such as KEGG (www.genome.jp/kegg) and Biocarta (www.biocarta.com) are fairly incomplete and inconsistent on this aspect (8). Hence, we manually selected target genes based on extensive scientific evidence for each individual gene being a direct target gene of the respective transcription complex, including promotor region motif analysis, transcription factor–binding experiments, and differential expression analysis. For Wnt, extensive research at the Hubrecht Institute over the past decades (9–12) culminated in a list of 34 "*bona fide*" target genes listed in Supplementary Table S1. For ER, we extensively investigated available literature, as detailed further in the Supplementary ER Pathway Target Gene Selection, yielding the 27 genes in Supplementary Table S2. These numbers of genes are on the one hand low enough to give specific results, but large enough to get robust models.

Model calibration

The probabilistic relations in the Bayesian network models need to be made quantitative to allow for quantitative probabilistic reasoning. To improve generalization behavior across tissue types, we manually set parameters describing the probabilistic relationships (i) between transcription complex and target genes using the following table:

A: for upregulated target genes			B: for downregulated genes		
	TG_i = down	TG_i = up		TG_i = down	TG_i = up
TC = absent	0.95	0.05	TC = absent	0.45	0.55
TC = present	0.30	0.70	TC = present	0.95	0.05

(1)

Parameters describing relationships (ii) between target genes and their respective probesets were calibrated on experimental data. For the latter, we used microarray data either from cell line experiments with defined active and inactive pathway settings (Wnt and ER pathway) or from patient samples with known pathway activity status (Wnt pathway only). The resulting conditional probability tables are given by:

A: for upregulated target genes

	$PS_{i,j}$ = low	$PS_{i,j}$ = high
TG_i = down	$\frac{AL_{i,j}+1}{AL_{i,j}+AH_{i,j}+2}$	$\frac{AH_{i,j}+1}{AL_{i,j}+AH_{i,j}+2}$
TG_i = up	$\frac{PL_{i,j}+1}{PL_{i,j}+PH_{i,j}+2}$	$\frac{PH_{i,j}+1}{PL_{i,j}+PH_{i,j}+2}$

(2a)

B: for downregulated target genes

	$PS_{i,j}$ = low	$PS_{i,j}$ = high
TG_i = down	$\frac{PL_{i,j}+1}{PL_{i,j}+PH_{i,j}+2}$	$\frac{PH_{i,j}+1}{PL_{i,j}+PH_{i,j}+2}$
TG_i = up	$\frac{AL_{i,j}+1}{AL_{i,j}+AH_{i,j}+2}$	$\frac{AH_{i,j}+1}{AL_{i,j}+AH_{i,j}+2}$

(2b)

In these tables, the variables $AL_{i,j}$, $AH_{i,j}$, $PL_{i,j}$, and $PH_{i,j}$ indicate the number of calibration samples with an absent (A) or present (P) transcription complex that have a low (L) or high (H) probeset intensity, respectively. Dummy counts have been added to avoid extreme probabilities of 0 and 1. For more details on the model construction, we refer to the Supplementary Bayesian Network Construction.

Statistical tests

Statistical tests, generation of Kaplan–Meier curves, and other graphics were performed using R (13). Generally, one-sided tests were used because the expected sign of a relation is known.

Results

Initial validation of the Wnt pathway model

Two instances of the Wnt model were created, a first one for initial proof of concept using cell line data for calibration, and a second one using a larger calibration dataset with patient data, with the advantage of better reflecting the variation encountered across patient samples.

The first Wnt model was calibrated on data from 12 samples of a Wnt abrogation experiment on an LS174T colon cancer cell line (GSE18560; ref. 14), of which six have an active Wnt pathway and six an inactivated Wnt pathway. For initial proof of concept, this model was tested on a dataset with 32 normal colon tissue samples and 32 colon adenoma samples from patients (GSE8671; ref. 11). The Wnt pathway is thought to be active in colon adenoma and inactive in normal intestinal tissue (15), and Fig. 2A shows that this is almost perfectly predicted by our model. Although two of the 32 adenoma samples are predicted to have an inactive Wnt pathway, if the threshold is set at odds of 1:1, the difference with normal colon samples is highly distinctive. Note that reported odds get as large as a million to one, which is due to the model using 34 genes (83 probesets in total), so although one gene may give quite noisy information, combining 34 genes gives quite confident predictions.

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

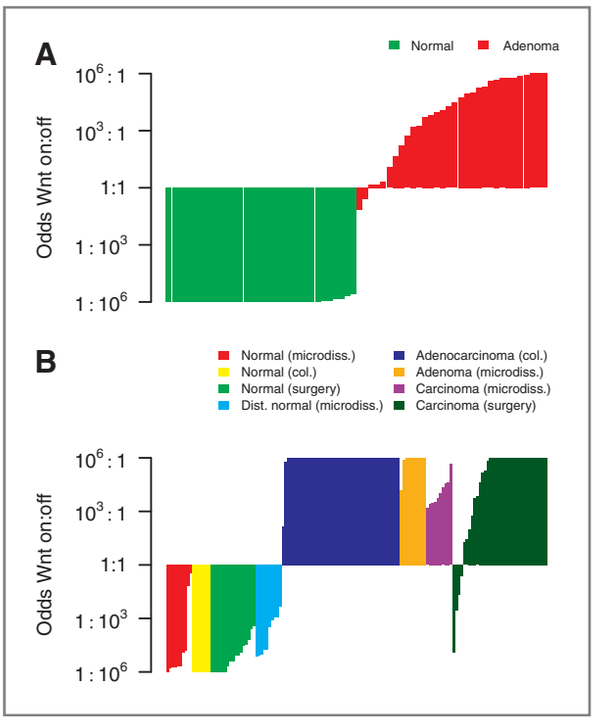

Figure 2. Results from Wnt model analysis of microarray data from samples from patients with colon adenoma and carcinoma. For each sample, a bar is plotted, indicating the odds of the Wnt pathway being active ("on") versus inactive ("off") on a logarithmic scale. At the top of the graph, the odds are 1 million to 1 that the pathway is active, at the lowest point they are 1 million to 1 that the pathway is inactive. A, results of Wnt model analysis on normal colon samples and colon adenoma samples (GEO dataset GSE8671; ref. 11), calibrated on colon cancer cell lines. B, results of Wnt model analysis of GEO dataset GSE20916 (16) using the Wnt model calibrated on the dataset of normal colon and colon adenoma (shown in A). Analyzed samples were colon tumors ($n = 101$) obtained by colonoscopy (dark blue), microdissected adenoma (orange), microdissected colon carcinoma (purple), and colon carcinoma obtained by surgery (dark green); corresponding control intestinal tissue samples ($n = 44$) consist of microdissected distal normal colon tissue (yellow and light blue) and normal colon tissue obtained by colonoscopy (red) or surgery (green).

Because the first calibration dataset has limited diversity, and ground truth information on Wnt pathway activity is in principle known for normal colon and adenoma, a second model was calibrated using the 32 normal colon samples and 32 colon adenoma samples from GSE8671 (11), and this model was used for the rest of the Wnt experiments reported in this article.

Wnt pathway in colon cancer

In general, in colon tumors, adenomatous polyposis coli (APC) tumor suppressor activity is absent due to loss of functional *APC* alleles, which is associated with active Wnt signaling, providing an excellent opportunity for clinical evaluation of the Wnt model (15). In dataset GSE20916 (Fig. 2B; ref. 16), all cancer and microdissected adenoma samples obtained through colonoscopy, as well as most (32 of 36) surgically resected colon carcinoma samples were predicted by our model to have an active Wnt pathway (97%, $n = 101$), whereas all normal colon tissue samples ($n = 44$) were predicted Wnt inactive. The four Wnt-inactive surgical colon cancer samples may be explained by cancer tissue heterogeneity and abnormal gene promoter methylation associated with more advanced cancer, or the surgery-associated sampling procedure may have resulted in mRNA degradation and unreliable microarray results (7, 17). Such factors may interfere with the expected Wnt target gene mRNA profile and reduce sensitivity of our current model. Indeed and illustratively, when our model was applied to the transcriptome of Wnt-active colon cancer cell line HCT116, in which specific Wnt target genes are methylated (18), the model predicted an inactive Wnt pathway (odds 8:1, data not shown).

Above results provide evidence that the model can identify active versus inactive Wnt pathway state in tumors arising from the colon epithelial cell type used to develop and calibrate the model.

Use of the Wnt model in tumors of other tissue origin

The aim of our models is to enable wide diagnostic usage across tumors of different cellular origins. Although direct target genes are transcribed by induced activity of one or more pathway-specific transcription factors binding to their respective gene response elements, indirect target genes are more likely to depend on additional cellular proteins for their transcription, increasing the likelihood of cell type–specific effects on expression regulation. For this reason, gene selection used to build the models focused on direct pathway target genes. To evaluate this premise of relative tissue-type independent functioning, we subsequently tested the model on other tumor types.

Liver cancer

In hepatocellular carcinoma (HCC) and hepatoblastoma, heterozygous somatic mutations (or deletions) in the third coding exon (codons 32, 33, 34, 37, 41, and 45) of the β-catenin (*CTNNB1*) gene have been frequently identified, resulting in substitution (or loss) of an amino acid, which may be associated with aberrant activation of the Wnt pathway (19–21).

Dataset GSE9843 (22) contains 91 liver cancer tissue samples, of which 27 contain a β-catenin gene mutation, 60 possess wild-type β-catenin, and four are unknown. Furthermore, 31 of the 91 samples scored positive for nuclear β-catenin staining, 55 negative, and five unknown. Interestingly, correlation between β-catenin mutation status and nuclear staining is not significant in this dataset (OR = 2.3; one-sided Fisher exact test, $P = 0.07$), illustrating the difficulty to get ground truth information on Wnt pathway activity in these samples, due to lack of a reliable test.

Chiang and colleagues (22) applied hierarchical clustering based on mRNA microarray data of the 91 samples of GSE9843, yielding five groups labeled "unannotated," "polysomy chr7," "inflammation," "proliferation," and "CTNNB1." Despite its label, the latter group of 24 samples only contains 16 of the 27 samples with a β-catenin mutation, and 14 of the 31 samples with a positive staining. The results of our Wnt pathway model on this dataset are shown in Fig. 3A. In the CTNNB1 group, 83% (20 of 24) of the samples are predicted to have an active Wnt pathway, versus 26% (6 of 23) in the proliferation group and 0% in the other three groups. In addition, although β-catenin mutation status and staining were not significantly correlated,

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

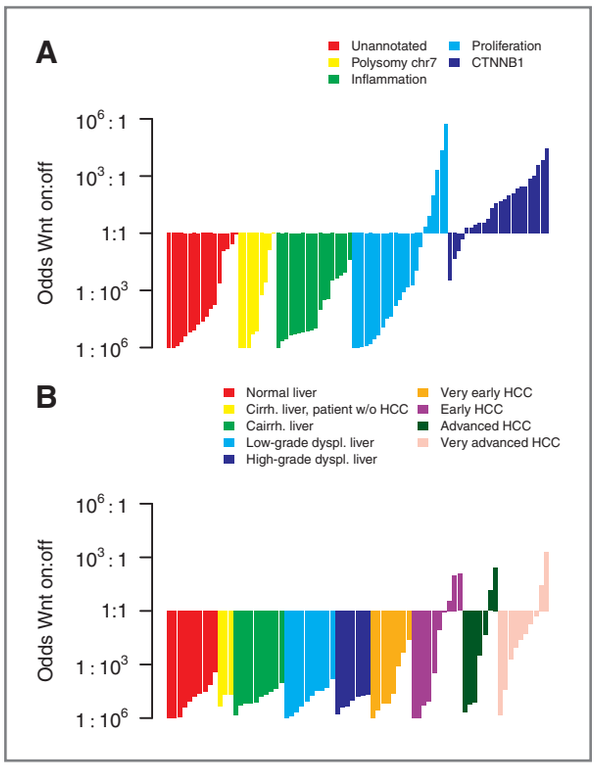

Figure 3. Results from Wnt model analysis of tissue samples from primary liver cancer. A, patients with hepatocellular carcinoma ($n = 91$, dataset GSE9843; ref. 22) labeled "unannotated" (red), "polysomy chr 17" (yellow), "inflammation" (green), "proliferation" (light blue), and "CTNNB1" (dark blue). B, patients ($n = 69$, dataset GSE6764, β-catenin mutational status unknown; ref. 23) with normal liver ($n = 10$, red), cirrhotic liver tissue ($n = 13$, yellow and green), low-grade ($n = 10$, light blue) and high-grade ($n = 7$, dark blue) dysplastic liver tissue, and hepatocellular carcinoma ($n = 8$ very early HCC, orange; $n = 10$ early HCC, purple; $n = 7$ advanced HCC, dark green; $n = 10$ very advanced HCC, pink). The *y*-axis shows the odds of the Wnt pathway being active ("on") versus inactive ("off") on a logarithmic scale. Each tissue sample result is represented by a bar.

we found a significant correlation of Wnt pathway activity with both β-catenin mutation status and β-catenin staining (OR = 5.4 and 6.1; respectively; and one-sided Fisher exact test, $P = 6.9e{-}4$ and $3.4e{-}4$, respectively; detail in Supplementary Detailed Liver Cancer Results on GSE9843).

Analysis of data from a second dataset of hepatocellular carcinoma samples (GSE6764; ref. 23) with unknown mutational status yielded an increased incidence of Wnt activity in patients with relatively more malignant tumors, as seven of the 27 early HCC, advanced HCC, and very advanced HCC samples were predicted to have an active Wnt pathway, compared with none in the very early HCC and nonmalignant sample groups (Fig. 3B, one-sided Fisher exact test, $P = 4.5e{-}4$).

Medulloblastoma

In medulloblastoma, a subset of tumors is known to possess an activating mutation in the β-catenin gene (24). We applied our model in a blinded manner to the medulloblastoma dataset from Kool and colleagues ($n = 62$, GSE10327; ref. 24), and successfully identified all samples with a Wnt pathway–activating β-catenin mutation (Fig. 4). In another medulloblastoma dataset ($n = 40$, GSE12992; ref. 25), the Wnt model also correctly identified the four samples with a driving β-catenin mutation against 36 without (data not shown). These two datasets show perfect performance of the model in this tumor type with 100% sensitivity and 100% specificity.

Breast cancer

In patients with breast cancer, direct Wnt-activating gene mutations are generally not present, except for a few rare metaplastic breast cancer cases (26). Rather, activation of the Wnt pathway has been indicated by circumstantial evidence in patients with triple-negative or basal-type breast cancer, most likely induced by interaction between the cancer cell and its microenvironment (27).

Two datasets were used for analysis of Wnt pathway activity within breast cancer subtypes: GSE12276 (28) and GSE21653 (29); see Fig. 5A and B. Using mRNA data, patient samples in these datasets were subtyped according to Desmedt and colleagues (30) and Perou and colleagues (31), respectively. Despite the difference in subtyping approach, Wnt pathway activity was very comparable, with an active Wnt pathway in 30% (21 of 65 and 21 of 75, respectively) of basal-type cancer samples, versus only 7% (15 of 139 and 9 of 191, respectively) in other breast cancer subtype samples (one-sided Fisher exact test, $P = 2.7e{-}4$ and $4.8e{-}7$, respectively).

Taken together, the above results provide initial evidence that the Wnt model performs well on a variety of tumors without requiring additional training steps on the different cell types of origin.

Initial validation of the ER pathway model

The ER pathway model was calibrated on data from eight samples of the breast cancer cell line MCF7, of which four were deprived from estrogen and four were stimulated with 25 nmol/L E2 (GSE8597; ref. 32), yielding an inactive and an active ER pathway, respectively. Estradiol concentrations are typically around 0.5 nmol/L in normal breast tissue, but elevated in breast cancer to around 2 nmol/L (33). As a result, the 25

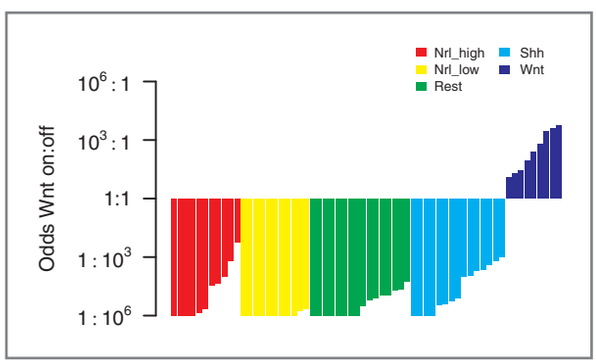

Figure 4. Results from Wnt pathway model analysis of samples from patients with medulloblastoma ($n = 62$, GSE10327; ref. 24), ordered as: samples expressing retinal differentiation genes, either high (red) or low (yellow), samples with a mutation in *SHH* (light blue) or *CTNNB1* (dark blue), and rest (green). The *y*-axis shows the odds of the Wnt pathway being active (on) versus inactive (off) on a logarithmic scale. Each tissue sample result is represented by a bar.

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

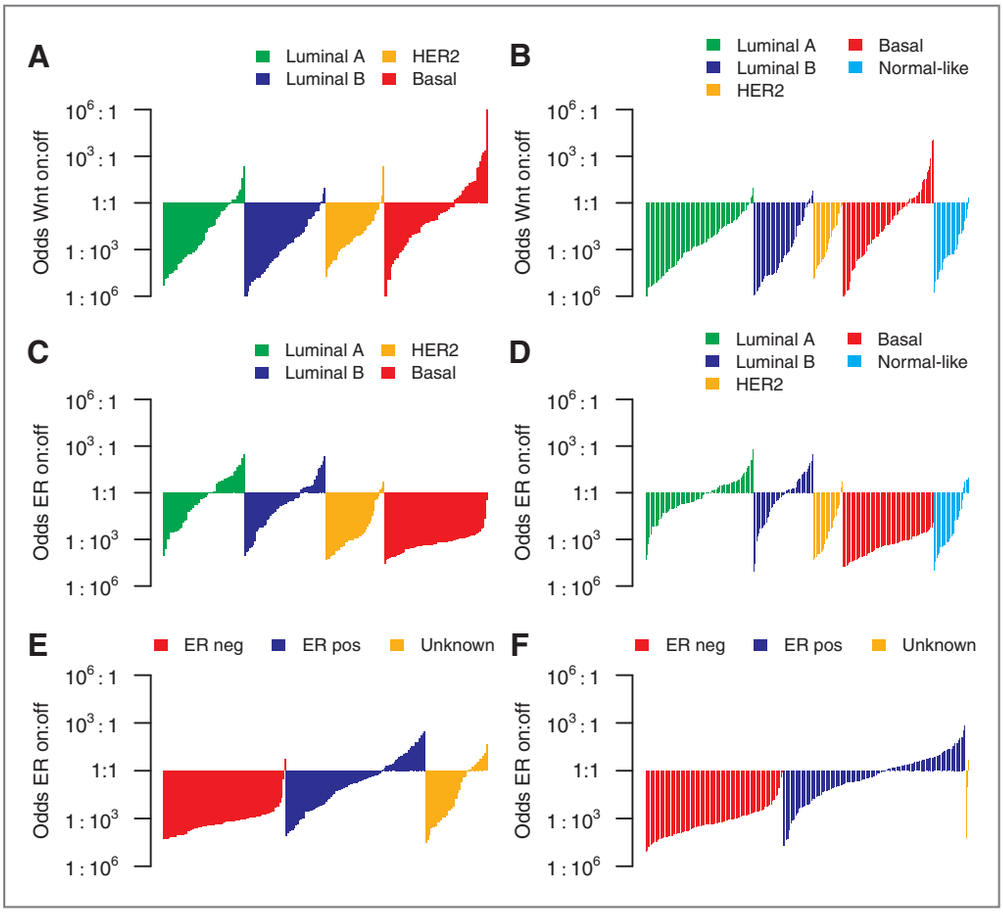

Figure 5. Results from Wnt (A and B) and ER (C–F) pathway model analysis on mRNA microarray data from samples of patients with breast cancer. The *y*-axis shows the odds of the respective pathway being active ("on") versus inactive ("off") on a logarithmic scale. Each tissue sample result is represented by a bar. A and C, odds of Wnt (A) and ER (C) pathway activity in a dataset with breast cancer samples ($n = 204$, GSE12276; ref. 28) subtyped according to the module approach from Desmedt and colleagues (30) as luminal A (green), luminal B (dark blue), HER2 (orange), and basal (red). All patients within this dataset suffered a relapse (median time to recurrence, 21 months; range, 0–115 months). The ordering of samples within each group is different in the two graphs; only five of the 204 samples have both an active Wnt and ER pathway. B and D, odds of Wnt (B) and ER (D) pathway activity in a dataset with breast cancer samples ($n = 266$, GSE21653; ref. 29) subtyped according to Perou's subtyping scheme (31), as luminal A (green), luminal B (dark blue), HER2 (orange), basal (red), and normal-like (light blue). The ordering of samples within each group is different in the two graphs; only four of the 266 samples have both an active Wnt and ER pathway. E and F, results from the ER pathway model analysis on mRNA microarray data from samples of patients with breast cancer from datasets GSE12276 (E, $n = 204$; ref. 28) and GSE21653 (F, $n = 266$; ref. 29). Samples are grouped here according to ER IHC status as ER negative (red), ER positive (blue), or unknown (orange).

nmol/L model may be slightly less sensitive than desired, but still useful for a first analysis.

An initial validation of the ER pathway model was performed on MCF7 breast cancer cell lines with and without a knockdown of the gene encoding for ER (from datasets GSE10890 and GSE37820, only published at GEO). All knockdown samples ($n = 8$) were predicted to have an inactive ER pathway, whereas all other MCF7 samples ($n = 28$) were predicted to have an active ER pathway (data not shown). In addition, in the cancer cell line encyclopedia (GSE36133; ref. 34), in all 861 cancer cell lines other than breast cancer, the ER pathway was predicted to be inactive (data not shown), indicating a very high specificity of the current model for ER pathway activity in breast cancer. Furthermore, running the pathway model on replicate experiments from datasets E-MTAB-37 (35) and GSE23593 (36) showed good reproducibility of predictions (Supplementary Table S4 and Supplementary Fig. S1).

Next, we applied the ER pathway model to the same two datasets of patients with breast cancer used above: GSE12276 (28) and GSE21653 (29); Fig. 5C and D, respectively. The figures show an active ER pathway in 41% (38 of 102 and 61 of 138, respectively) of luminal-type patients, versus only 4% (3 of 102 and 7 of 128, respectively) in other breast cancer subtype samples (one-sided Fisher exact test, $P = 1.3e{-}10$ and $3.7e{-}14$, respectively). The five HER2-type patients with predicted active ER pathway had scored positive for ER and/or progesterone receptor (PR) by immunohistochemistry (IHC) staining. All basal-type breast cancer samples were predicted to have an inactive ER pathway.

If patient samples are grouped according to ER IHC status (Fig. 5E and F), we observed in datasets GSE12276 and GSE21653 an active ER pathway in 39% (27 of 88 and 67 of 150, respectively) of ER-positive tumors, versus practically none (1 of 77 and 0 of 113, respectively) of the ER-negative tumors.

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

ER pathway activity and tamoxifen sensitivity

To link ER pathway activity to tamoxifen sensitivity, we analyzed dataset GSE21618 (37), containing samples from tamoxifen-sensitive and -resistant MCF7 breast cancer cell lines (all ER positive) treated with estradiol. In particular, we took the samples that had first been deprived of estrogen, and next been stimulated for up to 48 hours. Figure 6 shows the resulting probability of ER pathway activity as a function of stimulation time. Clearly, tamoxifen-sensitive cell lines quickly respond to estrogen stimulation, with probabilities steeply increasing toward 1, whereas tamoxifen-resistant cell lines respond to a lesser extent.

Initial assessment of prognostic value of the ER pathway model

Although the pathway models have been developed and trained to assess pathway activity to predict therapy response, we also tested to what extent they can have prognostic value. To this end, survival time analysis was performed on a dataset of 164 patients with ER-positive breast cancer that all received (only) adjuvant tamoxifen treatment for 5 years (Fig. 7; combined datasets GSE6532 & GSE9195; refs. 38, 39). The analysis was restricted to the first 5 years only, as tamoxifen treatment was limited to 5 years. As expected, patients with an active ER pathway have a better survival prognosis on tamoxifen treatment than patients for which the ER pathway is predicted inactive (one-sided log-rank test $P = 0.034$).

These results indicate that the ER pathway model may also have clinical utility in assessing prognosis in individual patients with breast cancer, even though it has not been developed for this purpose.

Discussion

With a few exceptions, e.g., HER2 and ER protein staining in breast cancer, most companion diagnostic assays to predict therapy response focus on identification of a tumor-specific genetic defect, associated with activation of a specific oncogenic signaling pathway (4, 40, 41). If the test result is positive, the associated signaling pathway is assumed to be active and driving tumor growth. Although a recent study by MD Anderson provides proof of principle that mutation-based identification of the tumor-driving signaling pathway improves therapy choice, it also illustrates that such a DNA-based companion diagnostic approach is unlikely to provide the complete answer, as in this study therapy response increased from 5% to only 27% (41). Indeed, it is clear that signaling pathway activation status (the functional "phenotype" of the cell) is determined not only by errors in the cancer cell (epi-)genome, but to a large extent by interactions between the cancer cell and its microenvironment (1, 32, 42). To assess the phenotype of the cell, in addition to the genotype, we presented a method to interpret cancer tissue transcriptome data as direct quantitative "output" of active signal transduction pathway(s), using knowledge-based Bayesian models. The pathway "output" is represented by transcribed pathway-specific target genes, which need to be known to create the models. So far, the data input for the models is from Affymetrix HG-U133Plus2.0 microarrays, but an important advantage of this type of knowledge-based models is that they can be easily calibrated to other input modalities, such as other array types, RNA sequencing, or dedicated multiplex PCR assays. We provided proof of principle that the models when trained on a limited number of cell line samples, or if available patient samples, already perform very well and robustly identify active oncogenic signal transduction pathway(s) in individual tissue samples obtained from a variety of malignancies.

For the Wnt pathway, our pathway model analysis results for adenoma, colon carcinoma, and medulloblastoma were in full concordance with existing evidence on Wnt pathway activation in these tumor types (15, 24, 42).

For other tumors such as primary liver and breast cancer, no easy "ground truth" with respect to activity status of the Wnt

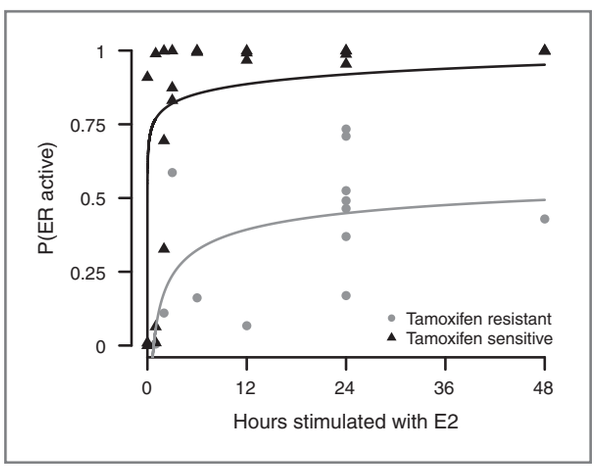

Figure 6. Prediction of ER pathway activity by the ER model in tamoxifen-sensitive and tamoxifen-resistant MCF7 cell lines over time. The cell lines have first been deprived of estrogen and next stimulated for up to 48 hours with estrogen. The vertical axis shows the predicted probability that the ER pathway is active. Microarray measurements were taken from dataset GSE21618 (37). Individual samples are plotted as points; the drawn lines are trend lines.

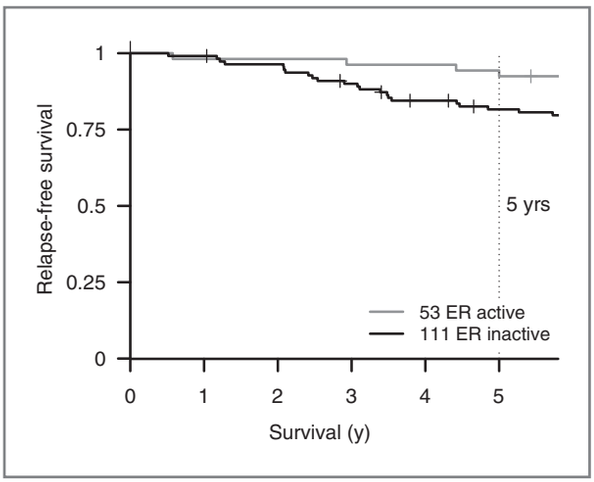

Figure 7. Kaplan–Meier curves showing prognostic value of ER pathway activity in patients with ER-positive breast cancer from datasets GSE6532 and GSE9195, all treated with adjuvant tamoxifen for 5 years ($n = 164$; refs. 38, 39). The gray line represents patients in which the model predicted an active ER pathway (odds > 1:1); the black line represents cases in which the model predicted an inactive ER pathway (odds $\leq$ 1:1).

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

pathway is available, but this pathway is likely to play a role in at least a number of cases (42–45). In agreement, our model identified an active Wnt pathway in the majority of liver cancer samples with, and a minority without, a β-catenin mutation. Identification of an active Wnt pathway in the absence of a β-catenin mutation may be due to the presence of other pathway-activating mutations in genes like *APC, AXIN1*, and *AXIN2* (46), or by paracrine Wnt activation (47, 48). Indeed, a significant correlation was found between model-predicted Wnt pathway activity and staining of nuclear β-catenin, presumably the active form of the Wnt pathway transcription factor. On the other hand, the presence of a β-catenin mutation does not necessarily mean that the pathway is activated, and in the cases in which the model did not detect Wnt activation despite a β-catenin mutation, another pathway may have taken the lead in tumor growth, for example induced by the microenvironment.

In breast cancer, Wnt activity was detected by the model in around one third of triple-negative or basal-type samples, which agrees with available evidence on a role for Wnt activity in this breast cancer subtype (43). In only very rare cases of breast cancer, a potentially Wnt-activating gene mutation has been found, suggesting that Wnt activity is most likely induced by paracrine interactions between cancer cells and their microenvironment. With a number of Wnt-targeting drugs in the pipeline of pharmaceutical companies, development of a reliable test to identify Wnt pathway activity in this cancer subtype with highly unfavorable prognosis is considered high priority (44, 49, 50), as β-catenin staining is not reliable enough to indicate Wnt pathway activity (50). Analysis of transcriptome data by our model is expected to provide information on Wnt activity in breast cancer, but final validation will require a clinical trial with an appropriate Wnt-inhibiting drug.

Other approaches have been directed toward assessing the phenotype of cancer by analyzing its transcriptome. To deduce pathway activity from tissue transcriptome data, most pathway analysis approaches use pathway information from databases such as KEGG (www.genome.jp/kegg) and Biocarta (www.biocarta.com) that mainly list genes encoding signaling proteins. Furthermore, they invariably extrapolate quantitative mRNA levels to levels of corresponding signaling proteins, followed by a search for a role of the transcript-encoded protein in a signaling pathway, which is subsequently defined as an active pathway (42, 44, 51). This approach is intrinsically flawed for pathway activation analysis because induction of an mRNA transcript coding for a component of a signaling pathway is not reliably correlated to the actual translated protein level and, even less, to the activation status of the encoded signaling protein, which requires additional post-translational protein modifications.

As expected, our approach leads to a different interpretation result of mRNA profiling data than conventional pathway analysis approaches. For example, we have run gene set enrichment analysis (52), using the accompanying curated canonical pathways, to identify the pathways differentially activated between the normal colon and colon adenoma samples from dataset GSE8671 (11), but none of the Wnt-related pathways was identified with a significant *P* value after correction for multiple testing (see Supplementary Table S5 in Supplementary Methods). Furthermore, another common pathway analysis approach, as presented by Skrzypczak and colleagues (16), in which first a list of differentially expressed genes is determined and next pathway sets are analyzed for overrepresentation of this gene list, did not identify the Wnt pathway as significantly different between normal colon and colon neoplasms—while our results were highly convincing about Wnt pathway activation status.

From the results on tamoxifen-sensitive and -resistant cell line data, it is inferred that the ER model can successfully detect ER activity in breast cancer cells. Being positive for the ER as measured by IHC or microarray analysis seems to be a necessary but not sufficient condition for ER pathway activity as assessed by the model. Because the model was trained on data from cell line experiments performed with a relatively high dose of estradiol (25 nmol/L), it cannot be excluded that the current model lacks to some extent in sensitivity to detect all samples with an active ER pathway. However, the finding that only a subgroup of ER-positive patients seemed to have an active ER pathway agrees with the common clinical observation that a number of ER-positive patients are primary resistant to hormonal therapy. Moreover, the model identified ER-positive patients treated with hormonal adjuvant therapy but with an inactive ER pathway as having a worse prognosis. This is in agreement with the concept that patients with the ER pathway driving tumor growth are more likely to benefit from hormonal adjuvant therapy.

With respect to the prediction of hormonal therapy response in patients with breast cancer, Symmans and colleagues (53) have described an mRNA profile of 165 ER-related genes, called the sensitive to endocrine therapy (SET) index. In contrast with the ER target genes in our model, genes underlying the SET index were not selected based on evidence for them being target genes of the ER transcription factor, but on increased expression levels found in ER-positive patients. The SET index was shown to identify patients with node-negative disease that have a good prognosis when treated with adjuvant hormonal therapy, as well as patients treated with neoadjuvant chemotherapy who have a high risk at relapse when subsequently treated with adjuvant hormonal treatment. Its predictive value may be partly contributed to the incorporation of six of the ER target genes used in our model.

We hypothesize that in ER-positive patients with an inactive ER pathway, other pathways such as Wnt or Hedgehog, associated with more aggressive behavior and worse outcome, may actually have been driving tumor growth (1, 54, 55). This is in agreement with the reported decline in SET index with advancing pathologic cancer stage, despite ER positivity, suggesting decreasing tumor dependency on an active ER pathway (53). Second, a high gene expression grade index, developed in a similar way as the SET index, also identifies subpopulations of ER-positive breast cancer with unfavorable prognosis (56). Taken together, these results strongly suggest that conventional quantification of nuclear ER IHC staining is not sufficiently specific in detecting functional ER pathway activity. According to pathologist guidelines, ER activity in a breast cancer sample is inferred from the presence of positive ER staining, with a minimum of 1% of ER-positive tumor nuclei as a threshold level. Such staining assays have an estimated 20% error rate due to

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

multiple factors that influence reliability, among them nonstandardized staining procedures and subjective interpretation (42, 50, 57). In contrast, microarray assays are quite standardized, and the model-based analysis uses a number of ER target genes interpreted in a weighted manner to calculate a probability of pathway activation, instead of ER as single variable as is the case in IHC testing. These conditions are promising for a more robust prediction of ER pathway activity.

Clinical utility of pathway models

The fact that the Bayesian network models are knowledge- and not data mining-based has several advantages. First, the models seemed to be well applicable to data analysis of multiple unrelated tumor types and this property allows use of the models for diagnostics purpose in cancers of different cell types of origin. Nevertheless, it is expected that the models can be further optimized with respect to sensitivity and specificity by including target genes that are specifically expressed in the tissue type of origin of the tumor. Such adaptation could, for example, entail adjusting the conditional probabilities and/or adding new nodes for novel target genes. Another advantage is the relatively easy translation of the model to another data format, such as Illumina DASL microarrays, RNA sequencing, or PCR-based testing. Finally, the remarkable reproducibility of the pathway model results across multiple (mostly public) datasets of a specific tumor type, generated at completely different hospitals and locations, demonstrates the robustness of this modeling approach for individual patient diagnostic use. The pathway model series will be extended to include all major oncogenic signaling pathways, ultimately providing a multipathway analysis suite for identification of both the major active pathway and potential underlying resistance pathways, applicable to tissue samples from a number of tumor types. Upon further clinical validation, currently under way, the expected main clinical utility of the described pathway models lies in therapy response prediction and monitoring in neoadjuvant and metastatic settings.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: W. Verhaegh, A. van de Stolpe
Development of methodology: W. Verhaegh, H. van Ooijen, M.A. Inda, P. Hatzis, A. van de Stolpe
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R. Versteeg, J. Martens, J. Foekens, H. Clevers
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): W. Verhaegh, H. van Ooijen, M.A. Inda, M. Smid
Writing, review, and or revision of the manuscript: W. Verhaegh, H. van Ooijen, M.A. Inda, P. Hatzis, M. Smid, J. Martens, J. Foekens, P. van de Wiel, A. van de Stolpe
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R. Versteeg, J. Foekens
Study supervision: W. Verhaegh, P. van de Wiel, A. van de Stolpe

Grant Support

This work is supported by NanoNextNL, a micro and nanotechnology consortium of the Government of the Netherlands and 130 partners (www.nanonext.nl).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 20, 2013; revised February 20, 2014; accepted March 18, 2014; published OnlineFirst April 2, 2014.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646–74.
2. Pharmaceutical Research and Manufacturers of America. Medicines in development for cancer: more than 900 medicines and vaccines in clinical testing offer new hope in the fight against cancer, PhRMA, Washington, DC, 2012 [monograph on the Internet]. Available from: http://www.phrma.org/sites/default/files/pdf/phrmamedicinesindevelopmentcancer2012.pdf.
3. Hamburg MA., Collins FS. The path to personalized medicine. N Engl J Med 2010;363:301–4.
4. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 2012;483:570–5.
5. DeVita VT, Lawrence TS, Rosenberg SA. Cancer: principles and practice of oncology. Philadelphia: Wolters Kluwer; 2011.
6. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.
7. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell 2012; 21:309–22.
8. Shmelkov E, Tang Z, Aifantis I, Statnikov A. Assessing quality and completeness of human transcriptional regulatory pathways on a genome-wide scale. Biol Direct 2011;6:15.
9. Van de Wetering M, Sancho E, Verweij C, De Lau W, Oving I, Hurlstone A, et al. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell 2002;111:241–50.
10. Van der Flier LG, Sabates-Bellver J, Oving I, Haegebarth A, De Palo M, Anti M, et al. The intestinal Wnt/TCF signature. Gastroenterology 2007;132:628–32.
11. Sabates-Bellver J, Van der Flier LG, de Palo M, Cattaneo E, Maake C, Rehrauer H, et al. Transcriptome profile of human colorectal adenomas. Mol Cancer Res 2007;5:1263–75.
12. Hatzis P, van der Flier LG, van Driel MA, Guryev V, Nielsen F, Denissov S, et al. Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. Mol Cell Biol 2008;28:2732–44.
13. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2008.
14. Mokry M, Hatzis P, Schuijers J, Lansu N, Ruzius FP, Clevers H, et al. Integrated genome-wide analysis of transcription factor occupancy, RNA polymerase II binding and steady-state RNA levels identify differentially regulated functional gene classes. Nucleic Acids Res 2012;40:148–58.
15. Barker N, Ridgway RA, Van Es JH, Van de Wetering M, Begthel H, Van den Born M, et al. Crypt stem cells as the cells-of-origin of intestinal cancer. Nature 2009;457:608–11.
16. Skrzypczak M, Goryca K, Rubel T, Paziewska A, Mikula M, Jarosz D, et al. Modeling oncogenic signaling in colon tumors by multidirectional analyses of microarray data directed for maximization of analytical reliability. PLoS ONE 2010;5. pii:e13091.
17. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, et al. Tumour evolution inferred by single-cell sequencing. Nature 2011;472: 90–4.
18. De Sousa E, Melo F, Colak S, Buikhuisen J, Koster J, Cameron K, De Jong JH, et al. Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. Cell Stem Cell 2011;9:476–85.
19. Koch A, Denkhaus D, Albrecht S, Leuschner I, von Schweinitz D, Pietsch T. Childhood hepatoblastomas frequently carry a mutated

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

degradation targeting box of the beta-catenin gene. Cancer Res 1999;59:269–73.
20. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 2012;44:694–8.
21. Provost E, McCabe A, Stern J, Lizardi I, D'Aquila TG, Rimm DL. Functional correlates of mutation of the Asp32 and Gly34 residues of beta-catenin. Oncogene 2005;24:2667–76.
22. Chiang DY, Villanueva A, Hoshida Y, Peix J, Newell P, Minguez B, et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res 2008;68:6779–88.
23. Wurmbach E, Chen YB, Khitrov G, Zhang W, Roayaie S, Schwartz M, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 2007;45:938–47.
24. Kool M, Koster J, Bunt J, Hasselt NE, Lakeman A, Van Sluis P, et al. Integrated genomics identifies five medulloblastoma subtypes with distinct genetic profiles, pathway signatures and clinicopathological features. PLoS ONE 2008;3:e3088.
25. Fattet S, Haberler C, Legoix P, Varlet P, Lellouch-Tubiana A, Lair S, et al. Beta-catenin status in paediatric medulloblastomas: correlation of immunohistochemical expression with mutational status, genetic profiles, and clinical characteristics. J Pathol 2009;218:86–94.
26. Hayes MJ, Thomas D, Emmons A, Giordano TJ, Kleer CG. Genetic changes of Wnt pathway genes are common events in metaplastic carcinomas of the breast. Clin Cancer Res 2008;14:4038–44.
27. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 2011;121:2750–67.
28. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, et al. Genes that mediate breast cancer metastasis to the brain. Nature 2009;459:1005–9.
29. Sabatier R, Finetti P, Cervera N, Lambaudie E, Esterni B, Mamessier E, et al. A gene expression signature identifies two prognostic subgroups of basal breast cancer. Breast Cancer Res Treat 2011;126:407–20.
30. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clin Cancer Res 2008;14:5158–65.
31. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 2006;7:96.
32. Bourdeau V, Deschênes J, Laperrière D, Aid M, White JH, Mader S. Mechanisms of primary and secondary estrogen target gene regulation in breast cancer cells. Nucleic Acids Res 2008;36:76–93.
33. Van Landeghem AA, Poortman J, Nabuurs M, Thijssen JH. Endogenous concentration and subcellular distribution of androgens in normal and malignant human breast tissue. Cancer Res 1985:45:2907–12.
34. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603–7.
35. Krupp M, Itzel T, Maass T, Hildebrandt A, Falle PR, Teufel A. CellLineNavigator: a workbench for cancer cell line analysis. Nucleic Acids Res 2013;41:D942–8.
36. Barry WT, Kernagis DN, Dressman HK, Griffis RJ, Hunter JD, Olson JA, et al. Intratumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. J Clin Oncol 2010;28:2198–206.
37. Oyama M, Nagashima T, Suzuki T, Kozuka-Hata H, Yumoto N, Shiraishi Y, et al. Integrated quantitative analysis of the phosphoproteome and transcriptome in tamoxifen-resistant breast cancer. J Biol Chem 2011;286:818–29.
38. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol 2007;25:1239–46.
39. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomics 2008;9:239.
40. Papadopoulos N, Kinzler KW, Vogelstein B. The role of companion diagnostics in the development and use of mutation-targeted cancer therapies. Nat Biotechnol 2006;24:985–95.
41. Tsimberidou AM, Iskander NG, Hong DS, Wheler JJ, Falchook GS, Fu S, et al. Personalized medicine in a phase I clinical trials program: the MD Anderson Cancer Center initiative. Clin Cancer Res 2012;18:6373–83.
42. White BD, Chien AJ, Dawson DW. Dysregulation of Wnt/β-catenin signaling in gastrointestinal cancers. Gastroenterology 2012;142:219–32.
43. Geyer FC, Lacroix-Triki M, Savage K, Arnedos M, Lambros MB, MacKay A, et al. β-Catenin pathway activation in breast cancer is associated with triple-negative phenotype but not with CTNNB1 mutation. Mod Pathol 2011;24:209–31.
44. King TD, Suto MJ, Li Y.The Wnt/β-catenin signaling pathway: a potential therapeutic target in the treatment of triple negative breast cancer. J Cell Biochem 2012;113:13–8.
45. Schade B, Lesurf R, Sanguin-Gendreau V, Bui T, Deblois G, O'Toole SA, et al. β-catenin signaling is a critical event in ErbB2-mediated mammary tumor progression. Cancer Res 2013;73:4474–87.
46. Taniguchi K, Roberts LR, Aderca IN, Dong X, Qian C, Murphy LM, et al. Mutational spectrum of beta-catenin, AXIN1, and AXIN2 in hepatocellular carcinomas and hepatoblastomas. Oncogene 2002;21:4863–71.
47. Malanchi I, Santamaria-Martínez A, Susanto E, Peng H, Lehr HA, Delaloye JF, et al. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature 2011;481:85–9.
48. Green JL, La J, Yum KW, Desai P, Rodewald LW, Zhang X, et al. Paracrine Wnt signaling both promotes and inhibits human breast tumor growth. Proc Natl Acad Sci U S A 2013;110:6991–6.
49. O'Toole SA, Beith JM, Millar EK, West R, McLean A, Cazet A, et al. Therapeutic targets in triple negative breast cancer. J Clin Pathol 2013;66:530–42.
50. Khalil S, Tan GA, Giri DD, Zhou XK, Howe LR. Activation status of Wnt/β-catenin signaling in normal and neoplastic breast tissues: relationship to HER2/neu expression in human and mouse. PLoS ONE 2012;7:e33421.
51. Incassati A, Chandramouli A, Eelkema R, Cowin P. Key signaling nodes in mammary gland development and cancer: β-catenin. Breast Cancer Res 2010;12:213.
52. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
53. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. J Clin Oncol 2010;28:4111–9.
54. Rieger ME, Sims AH, Coats ER, Clarke RB, Briegel KJ. The embryonic transcription cofactor LBH is a direct target of the Wnt signaling pathway in epithelial development and in aggressive basal subtype breast cancers. Mol Cell Biol 2010;30:4267–79.
55. Hui M, Cazet A, Nair R, Watkins DN, O'Toole SA, Swarbrick A. The Hedgehog signalling pathway in breast development, carcinogenesis and cancer therapy. Breast Cancer Res 2013;15:203.
56. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 2006;98:262–72.
57. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). Arch Pathol Lab Med 2010;134:e48–72.

Published OnlineFirst April 2, 2014; DOI: 10.1158/0008-5472.CAN-13-2515

Cancer Research
The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

AACR American Association for Cancer Research

Selection of Personalized Patient Therapy through the Use of Knowledge-Based Computational Models That Identify Tumor-Driving Signal Transduction Pathways

Wim Verhaegh, Henk van Ooijen, Márcia A. Inda, et al.

Cancer Res 2014;74:2936-2945. Published OnlineFirst April 2, 2014.