

Hidden Markov Model for Defining Genomic Changes in Lung Cancer Using Gene Expression Data

CHIANG-CHING HUANG,¹ JEREMY M.G. TAYLOR,² DAVID G. BEER,³ and SHARON L.R. KARDIA⁴

ABSTRACT

The study of gene expression patterns in relationship to chromosomal position, the “transcriptome map,” has become an area of active research and has revealed unexpected chromosomal regions within which gene expression levels are highly correlated. In cancer research, these regional changes in gene expression that may result from alterations at the chromosome level such as gene amplification or loss. To facilitate the search for such regions utilizing gene expression data, we have developed a hidden Markov model (HMM). Maximum penalized likelihood is used to estimate the parameters in the model. This method is applied to a lung cancer microarray experiment, including 86 human lung adenocarcinomas. Several regions identified through the HMM are consistent with known recurrent regions of amplification or deletion in this cancer. We further demonstrate the association of these abnormal expression regions with measures of disease status, such as tumor stage, differentiation, and survival. These findings suggest that genes in these regions may play a major role in the process of carcinogenesis of the lung. Our proposed method provides a valuable tool to accurately pinpoint regions of abnormal expression for further investigation.

INTRODUCTION

RECENT ADVANCES in DNA microarray technologies and the abundance of genomic information have provided unprecedented opportunities to decipher the underlying molecular mechanisms related to disease physiology. However, the exploding amount of gene expression data gathered poses great challenges to the scientific community and may provide misleading information without proper statistical interpretation. Consequently, the rich information provided in gene expression data and its broad biological implications will require new methods of analysis. These include employing sophisticated statistical and machine learning algorithms and intensive interaction with other genomic data sources to discover complex gene expression patterns and to correlate gene expression patterns with other biological processes. Recent examples of this include: combining of transcription factor binding and gene expression data to define transcriptional networks (Gao et al., 2004; Lee et al., 2002); utilizing DNA sequence, protein interaction, and microarray data to model genetic regulatory networks (Tamada et al., 2003; Nariai et al., 2004); and incorporating gene function data into microarray analysis for unknown gene functional inference (Zhou et al.,

¹Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois.
Departments of ²Biostatistics, ³Surgery, and ⁴Epidemiology, University of Michigan, Ann Arbor, Michigan.

2002; Cui et al., 2004). The advantage of these approaches over the analysis of gene expression data alone is that, they can not only reduce false positive findings, but also render biological significance to the statistical significance of the analysis result.

One endeavor in this area that has received increasing attention is the study of the “transcriptome map,” where the transcriptome map is the gene expression values as they relate to the position of the genes along the chromosome. These studies allow the visualization of global chromosomal expression patterns and have led to the discovery of many unexpected associations among genes (Cohen et al., 2000; Caron et al., 2001; Spellman and Rubin, 2002). The results of such analyses can also challenge the current notion of genetic regulation. For example, Cohen et al. (2000) used chromosome correlation maps from the *Saccharomyces cerevisiae* genome to demonstrate examples of adjacent pairs of genes with highly correlated expression patterns, in which the promoter of only one of the two genes contains an upstream activating sequence known to be associated with the expression pattern.

One application of the study of the transcriptome map is to identify chromosomal regions of genes demonstrating abnormal expression levels in cancer (i.e., regions in which all genes are either all highly under- or over-expressed). These genes have the potential for identifying areas where there has been gene amplification or deletion (Fujii et al., 2002; Bisognin et al., 2004). Such regions of gain and loss potentially harbor tumor oncogenes and suppressor genes. Hence, detection of these genetic events and the associated pattern of molecular architecture as evidenced by gene expression can provide a crucial step towards understanding genetic instability.

Current methods used to search for regions of increased or decreased expression rely mainly on averaging gene expression values from neighboring genes along the chromosome (Caron et al., 2001; Fujii et al., 2002; Pollack et al., 2002) or using scan statistics on a pre-defined length of gene sequence windows (Husing et al., 2003; Levin et al., 2005). These techniques could be effective at screening for possible regions of interest. However, they may lack precision and require ad-hoc choices, such as how many neighbors to include in this average. Another challenge is selecting a suitable threshold that categorizes a gene as having abnormal expression.

To address these issues, we developed a hidden Markov model (HMM) to identify regions of genes with abnormal expression levels from microarray gene expression data. A HMM is a type of stochastic model that has been used successfully in a variety of scientific applications (Rabiner, 1989; Koski, 2002). Such a model can be useful when there is an ordered sequence of entities and a measurement is taken on each entity. In our case the ordered entities are the genes along the chromosome and the measurements are the expression values. The model is particularly useful when there is thought to be a finite number of possible states each entity could be in, and there is a correlation between neighboring entities. The model has a probabilistic structure, which enhances the ability to make statistical inference. In particular the model allows the probability of different expression status for each gene can be calculated, and hence regions of abnormal expression identified.

METHODS

The genes along each chromosome are ordered in the sequence 1, 2, 3, 4, etc. The observed gene expression data is denoted by $\{x_1, x_2, x_3, \dots\}$. It is assumed that each gene can be in one of a finite number of “hidden” states, and that the distribution of possible expression values for a gene is determined by its “hidden” state. The purpose of a HMM is, through a probabilistic structure, to infer the “hidden” state sequence $\{S_1, S_2, S_3, \dots\}$ from the corresponding observed data $\{x_1, x_2, x_3, \dots\}$. In our problem of the identification of regions of over- or under-expression, each gene can be in one of 5 possible states, defined as follows:

$$S_t = \begin{cases} 1, & \text{region of underexpression} \\ 2, & \text{singleton of underexpression} \\ 3, & \text{normal expression} \\ 4, & \text{region of overexpression} \\ 5, & \text{singleton of overexpression} \end{cases}$$

Regions 1 and 4 are the ones of primary interest to us, as they might represent genomic changes however as some abnormal expression genes may not reside in those regions we also include states 2 and 5. State 3 represents genes that are expressing at a level typical of that seen in normal samples. In the over and under expressed regions, expression values from neighboring genes are likely to be highly correlated. This is incorporated in a HMM by considering the transition probabilities between neighboring genes.

Let $P_{ij} = P(S_{t+1} = j | S_t = i)$ denote the transition probability of gene $t + 1$ being in the state j given gene t is in state i . Here, gene t indicates the t th gene in the ordered position from the direction of the p arm to the q arm in a chromosome. The distribution of the gene expression value for a gene is determined by its hidden state. We assume a parametric form for the density of the gene expression level for each state and denote these densities by $f(x|\theta_1), f(x|\theta_2), f(x|\theta_3), f(x|\theta_4)$, and $f(x|\theta_5)$, where x is the observed data and $\theta_j = (\mu_j, \sigma_j)$ are the parameters in each density function. Fitting a HMM requires estimation of the parameters P_{ij} and $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$, from which one can infer the probability that the gene is in a particular state. The fit also provides limiting probabilities π_j 's for the proportion of genes in each of the 5 states.

Order reversibility

We chose to order the genes from the p arm to the q arm. For the purpose of analyzing gene expression data we could equally well have reversed the order of the genes along the chromosome. Thus it is appropriate in this context to impose an order-reversible constraint on the transition probabilities. In other words, suppose the expression state sequence from p arm to q arm in a chromosome is a stationary Markov chain having transition probabilities P_{ij} 's and limiting probabilities π_j 's, then the reversed state sequence is also a Markov chain with transition probabilities given by $Q_{ij} = \pi_j P_{ji}/\pi_i$ (Ross, 1993). It is well known that if the state space is finite (as in our case) and $P_{ij} > 0$ for all i, j , the necessary and sufficient condition for order reversibility is

$$P_{ij}P_{jk}P_{ki} = P_{ik}P_{kj}P_{ji} \quad \forall i, j, k \quad (1)$$

Using simple combinatorial mathematics, it can be shown that 6 non-redundant equations of the form (1) can be constructed for a 5-state order reversible Markov chain. Specifically, the 6 nonlinear equations of P_{ij} can be written as

$$\begin{aligned} P_{12}P_{23}P_{31} &= P_{13}P_{32}P_{21}, & P_{12}P_{24}P_{41} &= P_{14}P_{42}P_{21}, \\ P_{12}P_{25}P_{51} &= P_{15}P_{52}P_{21}, & P_{13}P_{34}P_{41} &= P_{14}P_{43}P_{31}, \\ P_{13}P_{35}P_{51} &= P_{15}P_{53}P_{31}, & P_{14}P_{45}P_{51} &= P_{15}P_{54}P_{41}. \end{aligned}$$

Thus, the order reversibility condition reduces the number of free parameters that need to be estimated.

Identifiability of mixture distribution

A HMM is a type of mixture model since an independent mixture model can be considered as a special case of HMM when the transition probabilities P_{ij} do not depend on i . A well known general problem with mixture models is identifiability of the parameters. In our model, because the support of the distributions for state 1 and state 2 (or state 4 and state 5) are negative (positive) and their distributions are assumed to have the same parametric form, we can exchange the indices of state 1 and state 2 (or state 4 and state 5) of the Markov chain without changing the mixture distribution. Without further restrictions or constraints there will not be a unique maximizing point of the likelihood, and numerical methods for estimating the parameters are likely to have problems.

One way to overcome the problems associated with the lack of identifiability of the parameters is by the imposition of an appropriate constraint on the parameters (McLachlan and Peel, 2000). Since states 1 and 4 are regions of under- or over-expression it is appropriate to assume that the transition probability from state 1 (4) to state 1 (4) should be large while the transition probability within singleton states, i.e. from state 2 (5) to state 2 (5) should be small. One way to ensure this in the estimation method is to impose constraints such as $P_{11} > P_{22}$ and $P_{44} > P_{55}$. Alternatively, we may impose prior distributions on these particular parameters, we do this by adding a corresponding penalty functions to the log-likelihood in the es-

timation method. Since all transition probabilities fall in the interval of (0, 1), a suitable prior for the transition probabilities can be chosen from a *beta* distribution. The consequence of this will be that genes in a region of over- or under-expression will tend to group together due to the high transition probability P_{11} or P_{44} and hence separate from singleton over or under-expressed genes.

Based on biological rationale we added a further constraint to the transition probability matrix, namely that $P_{23} = P_{33} = P_{53}$. The motivation for this is that the probability of transitioning into a normal state is the same, irrespective of whether the gene is in a singleton state or a normal state. Note that this constraint implies a non-Markov (independence) property of those three transition probabilities.

Penalized log-likelihood

Suppose $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ is the observed relative expression sequence for a chromosome of a tumor sample. The complete data is these observed expression values plus the unobserved hidden states. A HMM assumes independence between the observation given their hidden state, thus the penalized complete-data log-likelihood of the HMM for a single chromosome of a single sample can be written as

$$\begin{aligned} \log PL^C(\Psi) = & \sum_{j=1}^5 Z_j^t \cdot \log p_j + \sum_{t=2}^n \sum_{j'=1}^5 \sum_{j=1}^5 Z_{j'j}^t \cdot \log P_{j'j} + \sum_{t=1}^n \sum_{j=1}^5 Z_j^t \cdot \log f(x_t | \theta_j) \\ & + \log h_1(P_{11}) + \log h_2(P_{22}) + \log h_4(P_{44}) + \log h_5(P_{55}) \quad (2) \end{aligned}$$

where $p_j = Pr(S_1 = j)$ is the initial probability, $\Psi = \{\{\theta_j\}, \{p_j\}, \{P_{ij}\}\}$ are the set of all parameters, and h_1, h_2, h_4, h_5 are the densities of the *beta* prior distributions for $P_{11}, P_{22}, P_{44}, P_{55}$ respectively. Specifically,

$$\begin{aligned} h_1 &\sim beta(\alpha_1, 1), \quad h_2 \sim beta(1, \beta_2) \\ h_4 &\sim beta(\alpha_4, 1), \quad h_5 \sim beta(1, \beta_5). \end{aligned}$$

We will select values of the tuning parameters $(\alpha_1, \alpha_4, \beta_2, \beta_5)$ to be greater than 1 to ensure that P_{11} and P_{44} are large, and P_{22} and P_{55} are small. The variables Z_j^t and $Z_{j'j}^t$ are indicator variables for the true hidden states defined as

$$\begin{aligned} Z_j^t &= \begin{cases} 1, & \text{if } S_t = j \\ 0, & \text{otherwise} \end{cases} \\ Z_{j'j}^t &= \begin{cases} 1, & \text{if } S_{t-1} = j' \text{ and } S_t = j \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

The first 3 terms on the right hand side of (2) are the usual ones seen in the HMM literature. The whole penalized complete-data log-likelihood is then the sum over all chromosomes of all samples.

Estimation procedure

The parameters in (2) can be estimated by the method of maximum likelihood, which has been successfully used in a variety of mixture problems. For computation of the estimates it is natural to use the EM algorithm, Redner and Walker (1984). While the EM algorithm has good theoretical properties, it can be slow. In addition, in the E-step, the calculation of the expected value of Z_j^t , $Z_j^t = f_j^{(t)} b_j^{(t)} / \sum_{j=1}^5 f_j^{(t)}$, can be numerically unstable, when directly applying the forward-backward algorithm (Baum et al., 1970), as the forward and backward probabilities, $f_j^{(t)}$ and $b_j^{(t)}$ for gene t being in state j , become rapidly close to zero as n increases (Leroux, 1992). To address this issue, we use the method described by Leroux and Puterman (1992). We choose for each gene t a value of r for which $10^r \sum_j f_j^{(t)}$ lies between 0.1 and 1 and multiply $f_j^{(t)}$ by 10^r ($j = 1, \dots, 5$). A similar procedure is applied to $b_j^{(t)}$. Then $f_j^{(t)}$ and the $b_j^{(t)}$ can be reconstructed for the purpose of computing the expected values of Z_j^t and $Z_{j'j}^t$.

In the M-step, there exist simple closed forms for finding the updated parameters in component densities of exponential family distributions and Markov chain transition probabilities for standard HMMs. How-

ever, as shown in (1), the condition of order reversibility for a Markov chain imposes 6 nonlinear constraints on the transition probabilities. This adds complexity on the estimation for the 14 free parameters in the transition probability matrix. To solve this problem, we employ sequential linear programming to find the solution in each M-step. We utilize the software, AMPL,¹ in conjunction with an R program for the forward-backward algorithm to obtain the parameter estimates and posterior probabilities of expression states for each gene.

Choice of tuning parameters

Because a *beta* distribution was used to define the penalty terms for P_{11} , P_{22} , P_{44} , P_{55} , the last four terms in equation (2) can be written as

$$(\alpha_1 - 1) \log P_{11} + (\alpha_4 - 1) \log P_{44} + (\beta_2 - 1) \log (1 - P_{22}) + (\beta_5 - 1) \log (1 - P_{55}) + const.$$

Thus, $\alpha_1, \alpha_4, \beta_2, \beta_5$ become the tuning parameters in (2). In standard penalized regression and classification problems tuning parameters are typically chosen by cross-validation. This is not possible with a HMM because the true state is not observed. To overcome this, we employ the idea of false positive rate (FPR) and false negative rate (FNR) to select the tuning parameters. We first define an event E_i as a sequence of genes with identical expression state and N_{Ei} the number of gene in the corresponding event. For example, for the sequence of expression states

33111133332233333344433335

there are eight events with $E_1 = 3$, $E_2 = 1$, $E_3 = 3$ etc. and, $N_{E1} = 2$, $N_{E2} = 4$, N_{E3} , etc. From the definition of the 5 expression states, the HMM should be able to classify a short sequence of under-expressed (or over-expressed) genes as an event of state 2 (or 5). Similarly, a long sequence of under-expressed (or over-expressed) genes should be classified as an event of state 1 (or 4). If we are willing to assume that any event E_i of an under-expressed gene with N_{Ei} being 1 should be in state 2, and any event E_j of an under-expressed gene with $N_{Ej} \geq 3$ should be in state 1, then the FPR and FNR for regions of under-expression can be defined as

$$FPR = \frac{\#\{E_i \mid N_{Ei} \geq 3, E_i = 2\}}{\#\{E_i \mid N_{Ej} \geq 3, E_i = 1 \text{ or } 2\}}$$

$$FNR = \frac{\#\{E_i \mid N_{Ei} = 1, E_i = 1\}}{\#\{E_i \mid N_{Ej} = 1, E_i = 1 \text{ or } 2\}}$$

The FPR and FNR for regions of over-expression can be defined in an analogous way. Therefore, the tuning parameters for the penalty terms can be adjusted in a way that the FPR and FNR for regions of under-expression or over-expression will be under some pre-selected thresholds. We use 0.05 as the threshold for both. However, we should note that this is not a unique way to select the tuning parameters because many different sets of tuning parameters may result in the same FPR and FNR.

RESULTS AND DISCUSSION

Analysis of lung cancer microarray data

We applied our method to a gene expression microarray experiment to search for regions of interest. The data were published by Beer et al. (2002). The data consist of gene expression values measured by Affymetrix HuGeneFL chips for 86 lung adenocarcinomas samples. In addition, there are data from 9 normal lung tissues taken from patients with lung cancer. The probeset expression summaries were calculated using the method described in Beer et al. (2002). On the Affymetrix HuGeneFL chip, there are 7,129 probesets with each probeset measuring the expression of one gene. However, there is redundancy in that there can be

¹The reference is written by Fourer et al. (1993). The student edition can be downloaded for free from the web: www.ampl.com/cm/cs/what/ampl/DOWNLOADS.

more than one probeset for a given gene. To account for this we took the median value for any gene with multiple probesets, which resulted in 6,800 gene expression measurements. All the negative expression values were set at zero. We further added a constant 150 to all the expression values to avoid artificially large fold changes compared to the normal tissues for genes with small expression.

After this data preprocessing, the reference expression value for gene k on chromosome i is determined as the median of 9 normal lung samples, denoted as $N_i^{(k)}$. Then, the rescaled data for analysis are $X_{im}^{(k)} = \log_2(T_{im}^k/N_i^{(k)})$, where T_{im}^k is the pre-processed expression for gene k on chromosome i and tumor sample m . For each gene, we have mapped its base pair location along the chromosome as described by Levin et al. (2001). Since some of the genes are not mapped through this program, the final data includes 5,707 genes for each tumor sample. We assume the density of the normal expression state is Gaussian. Since we have 9 normal tissue samples, we estimate the parameters $(\hat{\mu}_3, \hat{\sigma}^2_3)$ for this density using all the data from these normal tissues. The resulting parameter estimates are $(\hat{\mu}_3, \hat{\sigma}^2_3) = (0, 0.3^2)$. These estimates were then used in the EM algorithm as fixed. In addition, we assume truncated lognormal distributions with parameters (μ_j, σ_j) , $j = 1, 2, 4, 5$ for the other four component densities $f(-x|\mu_1, \sigma_1)$, $f(-x|\mu_2, \sigma_2)$, $f(x|\mu_4, \sigma_4)$ and $f(x|\mu_5, \sigma_5)$. The truncated point is set at 0.4, so that there was no probability mass between -0.4 and 0.4 for genes in over- or under-expressed states. As described in the previous section, the prior we chose for the transition probabilities $\{P_{11}, P_{44}\}$ are $\text{beta}(\alpha_1, 1)$ and $\text{beta}(\alpha_4, 1)$ respectively, and the prior for $\{P_{22}, P_{55}\}$ are $\text{beta}(1, \beta_2)$ and $\text{beta}(1, \beta_5)$, respectively.

We ran the EM algorithm until the increment of the penalized log-likelihood is less than $10^{-4}\%$ of the penalized log-likelihood from the previous iteration. We tried many sets of tuning parameters $\{\alpha_1, \alpha_4, \beta_2, \beta_4\}$ to ensure that the false positive and false negative rates are less than 0.05. The expected value of Z_j 's in the E-step from the last EM iteration is used as the posterior state probabilities to assign the expression state for each gene. A gene will be determined to be in state j if its corresponding posterior state probability is greater than 0.6. Thus, genes for which there was no clear indication of the state were not assigned to any particular state.

The estimated transition probability matrix is

$$\begin{bmatrix} 0.721 & 5 \cdot 10^{-5} & 0.238 & 0.010 & 0.032 \\ 1 \cdot 10^{-5} & 0.070 & 0.877 & 0.017 & 0.042 \\ 0.004 & 0.057 & 0.877 & 0.009 & 0.053 \\ 0.006 & 0.028 & 0.316 & 0.649 & 2 \cdot 10^{-5} \\ 0.009 & 0.045 & 0.877 & 8 \cdot 10^{-6} & 0.070 \end{bmatrix}$$

Note that the constraint on the transition probability matrix, has forced $P_{23} = P_{33} = P_{53}$. However, these three estimated transition probabilities were very similar when this constraint was removed. This transition probability matrix reveals several interesting phenomena. First, the transition probability of under-expression region to under-expression region (P_{11}) is larger than that of over-expression region to over-expression region (P_{44}). This may be due to the fact that the overall magnitude of the gene expression values relative to normal tissues in regions of under-expression is greater than that in regions of over-expression. This can be seen in Figure 1 where a larger proportion of genes in under-expression regions have values further away from zero than in over-expression regions. A possible explanation of this is that deletion of genes will result in disappearance of gene expression, resulting in a many fold change in gene expression relative to normal.

Second, P_{31} and P_{34} are close to 0. This suggests that there is very small number of regions of under- and over-expression in this data. And the number of regions of under-expression is even smaller than that of regions of over-expression ($P_{31} < P_{34}$). On the other hand, both P_{13} and P_{43} are the second largest transition probability in the first and fourth row of this matrix. This is because the majority of the genes are normally expressed. An interpretation of this is that genes in a region of under- or over-expression either continue in their current state or transit into a normal expression state. Furthermore, we calculated the limiting probabilities from the estimated transition probabilities. The resulting probabilities for the five different expression states are 1.78% for region of under-expression, 5.94% for singleton of under-expression, 84.6% for normal expression, 2.11% for region of over-expression, and 5.57% for singleton of over-

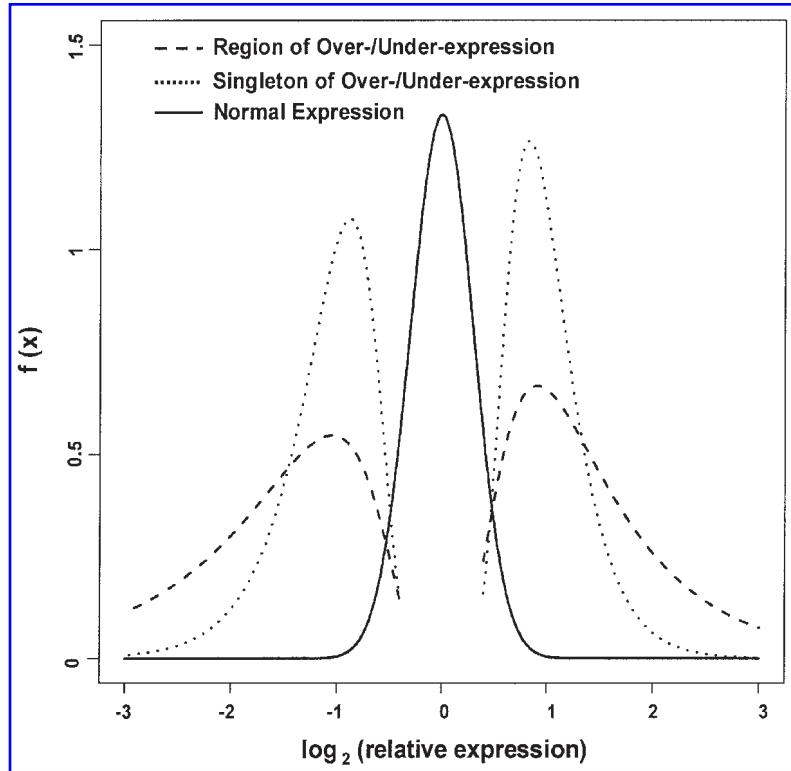


FIG. 1. Component densities of expression states estimated from hidden Markov model (HMM) for lung cancer data. The solid curve represents the density of relative expression values (in logarithm base 2 scale) in normal expression regions; dotted curves represent the densities of singleton of over- (right side) or under-expression (left side); dashed curves represent the density of region of over- (right side) or under-expression (left side). The relative expression is the ratio of the expression of a tumor over the median of nine normal expression values for any single gene.

expression. Note that the proportion of over-expressed genes (7.68%, state 4 and state 5) is close to the proportion of under-expressed genes (7.72%, state 1 and state 2).

An interesting result is found in Figure 1, which shows the five component densities. The densities suggest that the expression levels in regions of over-expression (under-expression) tend to be higher (lower) than those in singletons of over-expression (under-expression). Pollack et al. (2002) performed a parallel assessment of correlation between mRNA levels and DNA copy number changes using cDNA microarray and array CGH in a group of breast cancer lines and tumors. They observed that the overall patterns of gene amplification and elevated gene expression are quite concordant, and concluded that DNA copy number alterations have a pervasive global influence on gene expression. Thus, some of those regions we identified could be regions of gain or loss. Identification of the genes in those regions that are most highly expressed may lead to the discovery of oncogenes or genes important to the cancer process. Of the 86 × 5707 genes, the number of genes that are classified into state 1,2,3,4 and 5 are 2,419 (0.5%), 23,662 (4.8%), 422,218 (86.0%), 4,496 (0.9%), and 20,850 (3.5%), respectively. 17,159 (3.5%) genes are not classified into any of the five states because none of the five posterior state probabilities are greater than 0.6 for those genes.

Based on the definition of FPR and FNR in the previous section, the false positive and false negative rates for regions of under-expression (over-expression) are 5% (0.2%) and 0.5% (2.8%). The resulting number of regions and singletons of over- and under-expression is shown in Table 1. The definition of an event, E_i , in this table is given in the subsection, “Choice of tuning parameters.” For example, 367 events of 2 genes in a region of under-expression mean that those regions all have only 2 genes in a row that are called in state of under-expression.

TABLE 1. NUMBER OF EVENTS OF ABNORMAL EXPRESSION ESTIMATED FROM HIDDEN MARKOV MODEL FOR LUNG CANCER DATA

Events size	Region down	Singleton down	Region up	Singleton up
1 gene	99 (0.5%)	20,194	534 (2.8%)	18,837
2 genes	367 (16%)	1968	671 (40%)	1005
3+ genes	455 (95%)	24	708 (99.8%)	1

Association of regional abnormality and clinical measures

We investigated the impact of regions of over- and/or under-expression on clinical measures, such as tumor stage and patient survival. Since we are more confident with the calls for those regions of over- or under-expression with at least three genes in a row, we calculated for each of the 86 tumor samples the number of regions of over- and under-expression with $NE_i \geq 3$ (the total number of events in the whole data are 708 and 455, respectively as seen in Table 1).

We first observed an interesting result that a strong correlation exists between the number of regions of over-expression and the number of regions of under-expression (Spearman correlation 0.67, $p < 0.0001$). We then investigated the association of region number with tumor stage and differentiation. The patients' tumors are either stage 1 (67 patients) or stage 3 (19 patients), and the differentiation states: are well (23 patients), moderate (42 patients), and poor (21 patients). The three box plots in Figure 2 show the distribution of the number

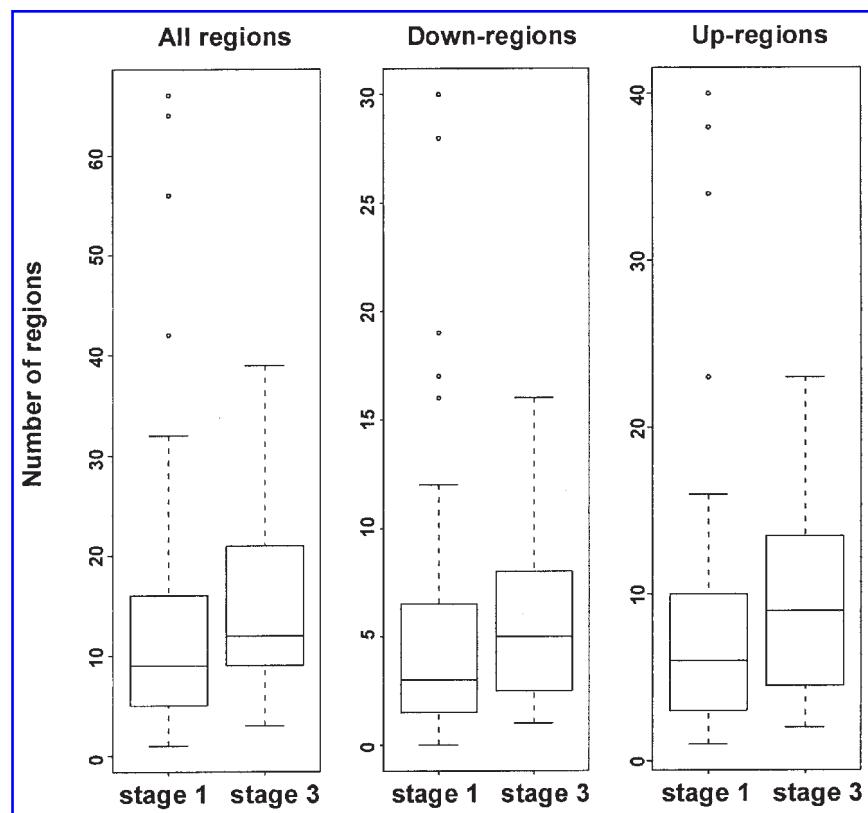


FIG. 2. The association of the number of regions of abnormal expression with stage in lung cancer data. Higher stage is associated with overall higher abnormal expression regions. The association is consistent for down- and up-regulated regions. However, four stage 1 tumors are found to have the highest numbers of abnormal expression regions in all three box plots.

of abnormal expression regions between stage 1 and stage 3 tumors. The number of all regions in the most left panel is the total number of regions of under- and over-expression. Although the difference of the number of abnormal expression regions between stage 1 and stage 3 tumors does not attain significant level (Wilcoxon rank sum test p value: 0.092, 0.067, and 0.203 for all regions, down regions, and up regions, respectively), stage 3 tumors tend to have more regions of over and/or under-expression than stage 1 tumors. The diluted significance level is largely due to the 4 stage 1 tumors that have most of abnormal expression regions. These four patients are all alive with survival times of 21.1, 87.7, 36, and 40 months (median survival time 29.5 months) with differentiation status being poor, poor, well, and moderate. With these 4 tumors removed, the tests show significant difference in the number of regions for stage (p -value: 0.028, 0.019, 0.078 for all regions, down regions, and up regions, respectively). An even stronger association can be seen in Figure 3 between the number of regions and the differentiation status. It is evident that the poorer the differentiation status of the tumor for each patient, the greater the number of regions of over- and/or under-expression. The Kruskal-Wallis rank sum tests result in p -values: 0.014, 0.017, and 0.013 for all regions, down-regions, and up-regions.

We finally examined the association of the number of regions with patient survival. We used Cox PH model (Cox, 1972) for survival analysis, treating the number of regions as a predictor. None of the p -values are significant for all 86 tumor samples. However, removing the four tumors results in p -values: 0.007, 0.003, and 0.049 for all regions, down-regions, and up-regions. Figure 4 is an example of this association in the all regions case. From the upper panel, we can see in general the greater number of regions of over- and under-expression for a patient the shorter the survival time. A Kaplan-Meier plot is shown in the lower panel, based on a dichotomization of the 82 patients based on a cut-off of 12 regions of over- and under-expression. It is evident that the group of patients with larger number of regions has a poor survival.

The process of carcinogenesis starts with the initiation of a normal cell which then progresses to a clonal

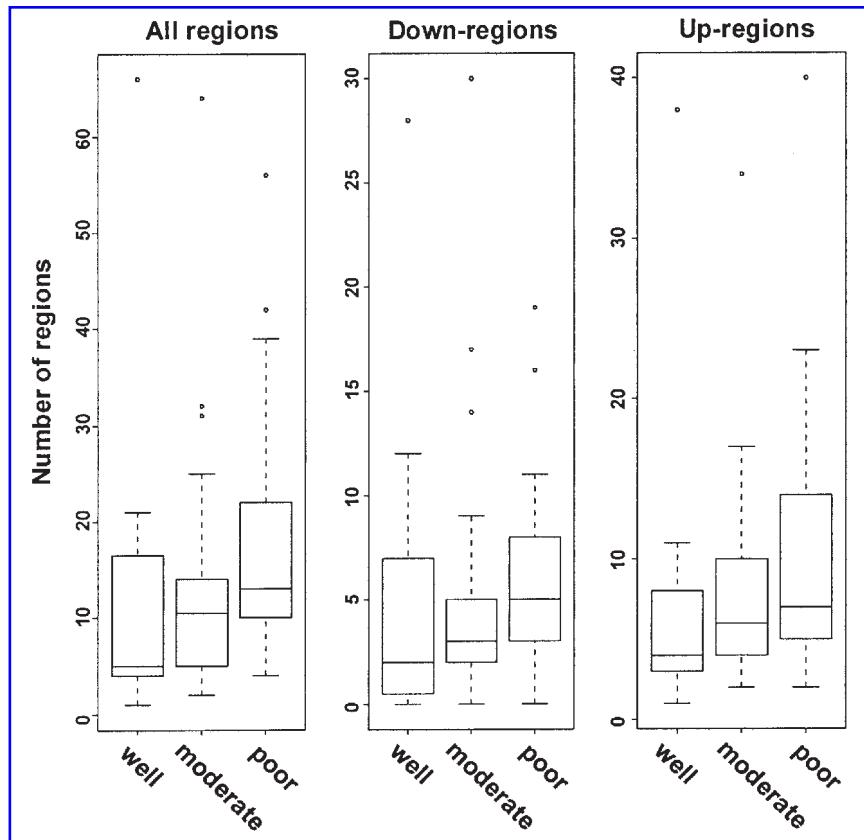


FIG. 3. The association of the number of regions of abnormal expression with differentiation status. More total abnormal, down- and up-regulated regions are found among the poorly differentiated tumors.

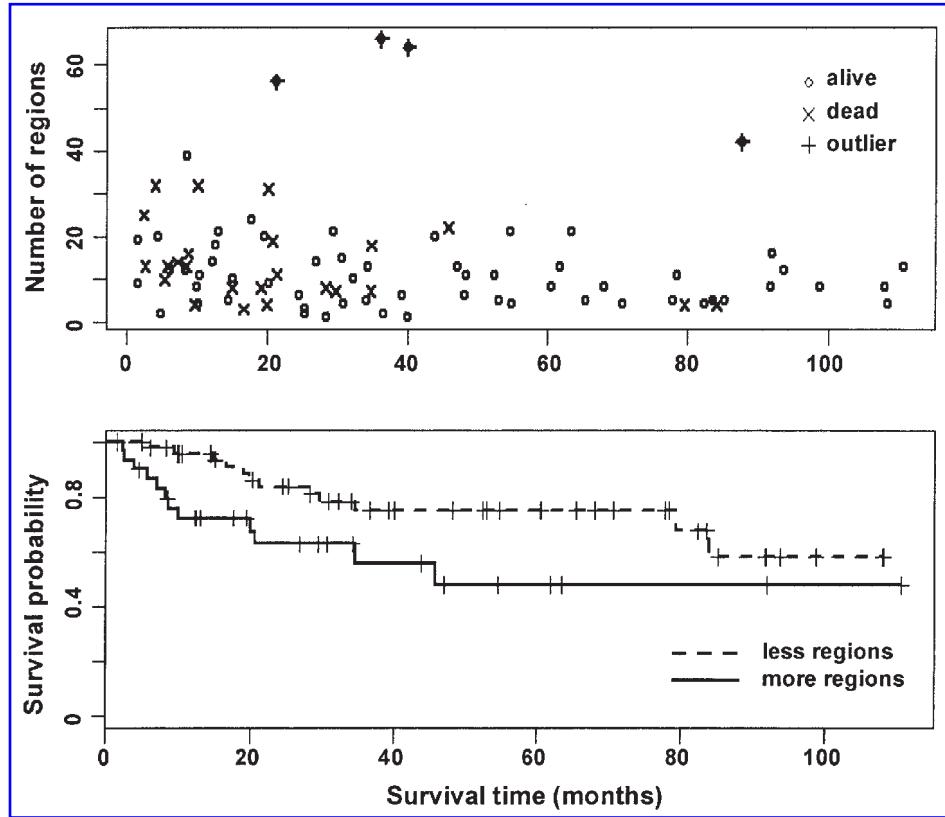


FIG. 4. The association of the number of regions of abnormal expression with patients' prognosis. Upper panel: The number of total abnormal expression regions (y-axis) is inversely associated with patients' survival time (x-axis). Lower panel: After removing four stage 1 tumors with the highest numbers of abnormal expression regions as observed in the upper panel, a significant difference of the survival distributions is observed between the two groups of patients with more (solid curve) or less (dashed curve) abnormal expression regions using 12 regions as the cutoff ($p=0.03$).

growth of malignant cells capable of invasion and metastasis. The mechanism driving this process is progressive DNA alterations, which occurs over time (Vogelstein and Kinzler, 2004). Under this model, it is the accumulation of DNA alterations in a cell which is involved in the progression from benign to the malignant state. Tumor stage and differentiation are well established surrogates of malignant potential. In this study, we provide a direct molecular measure of alterations from normal gene expression which reflects gross DNA damage (amplifications and deletions) as well as alterations in transcriptional control. As such, we expect these regional alterations in gene expression to reflect the accumulation of the DNA damage. Therefore, our findings that regional alterations in gene expression are correlated with stage and differentiation are consistent with the model of carcinogenesis described above.

Identification of regions of over- and underexpression

We further examine regions of over- or underexpression that occurs in many tumors and that may be of prognostic and therapeutic value. As the abnormal expression regions identified by our method may be caused by DNA copy number changes (amplifications or deletions), we will compare those regions to what have been reported in the literature such as Knuutila et al. (1998, 1999).

We determine that a gene is within a region of over- or under-expression if the posterior state probability is greater than 0.6. Since regions may contain various numbers of genes, we only select those that have at least three genes to avoid false positive calls for the small regions (one or two genes). However, each tumor is likely to have different genes that show abnormal expression levels in a specific chromosomal region. Therefore, we count the number of tumor samples in which a given gene is in a region of over- or underexpression.

We observed that some chromosomes have more abnormal expression regions than others. Regions of over-expression in chromosomes 1, 6, 7, 11, 12, 17, 19, 22, and X are more frequent. Many studies using CGH have reported amplifications in non-small-cell lung cancer (NSCLC) in these chromosomes as well. For example, 12p amplicon harbors the KRAS2 gene that has been detected in two lung adenocarcinomas (Bjorkqvist et al., 1998). In addition, amplifications of 6p12-pter and 17q24-qter are frequent in NSCLC (Knuutila et al., 1998). For these three regions, we observed in our data five genes (D6S51E, D6S52E, CSNK2B, CLIC1, and VARS2) in the 6p21.33-p22.1 region with frequent counts in four to seven tumor samples, three genes (TNFRSF1A, SCNN1A, and LTBR) in 12p13.2-p13.33 with frequent counts in six to seven tumor samples, and three genes (GALK1, ITGB4, and LLGL2) in 17q24.3-q25.3 with frequent counts in five tumor samples. We also report here the four most frequent regions of over-expression in our data: (1) 7p22.3 (JTV1, RAC1, KDEL2) with frequency of 10–15; (2) 21q22.3 (TFF1, TFF2, TFF3) with frequency of 10; (3) 11q23.3-q25 (SC5DL, unknown gene, SORL1) with frequency of 9; and (4) 11q12.1-q13.5 (EMS1, PPFI1A, FADD) with frequency of 3–8. We note that some of the survival genes reported in Beer et al. (2002) are in regions of over-expression and include VEGF, ERBB2, WNT10B, and IGFBP3, although the frequency is rare (1 or 2 only).

Regions of under-expression are much less frequent in our data as compared to the detection of regions of deletions in NSCLC (Knuutila et al., 1999). For example, we didn't find any region of under-expression at either region 3p14.2 or 9p21. We observed in our data two chromosomes, 16 and 19, that have the most regions of under-expression. We also observed 3 regions in chromosomes 6, 10, and 17. These regions are (1) 6p21.33 (HLA-DRA, HLA-DRB1, HLA-DRB5) with a frequency of 21; (2) 10q22.1-q22.2 (SFTPA1, SFTPA2, SFTPBD) with a frequency of 42; and (3) 17q11.2-q21.1 (SCYA14, SCYA5, SCYA4) with a frequency of 15. Although DNA deletion and amplification could account for expression changes in the regions identified above, the possibility of other mechanisms that may result in the expression abnormality should not be excluded. For example, DNA methylation, a type of chemical modification of DNA that can be inherited without changing the DNA sequence, can lead to gene silencing when combined with histone acetylation (Baylin et al., 2001). In addition, up-regulation of gene expression in a chromosomal region may be caused by the opening of the chromatin of an entire neighborhood as a result of activation of a target gene within the neighborhood (Spellman and Rubin, 2002). Overall, our findings suggest the possibility that there are regional changes associated with genomic structure changes; these merit further investigation for potential clinical utility.

CONCLUSION

In this paper, we have integrated gene expression data with genomic location information and have applied a HMM to discover chromosomal regions of abnormal expression. Several regions identified in this paper are consistent with known regions of amplification reported in the literature. This integrated analysis provides biological interpretation regarding the expression changes in certain lung tumors. Recent development of high throughput SNP arrays has allowed researchers to directly identify changes in the DNA level reflective of LOH, deletion, and amplification. Computational methods (Lin et al., 2004; Huang et al., 2004) have been developed to identify these abnormal DNA events. As cancers are complex biological systems, no single type of global profiling approach can entirely elucidate tumor behavior. Therefore, combining mRNA and DNA high throughput data in conjunction with sophisticated statistical algorithms may be needed to identify important molecular causes of carcinogenesis and characterize the fundamental processes of tumor growth.

ACKNOWLEDGMENTS

We acknowledge the contributions of Dr. Albert M. Levin towards the ideas presented in the manuscript and thank him for providing genome map location data. We also acknowledge the helpful comments of the reviewers of this manuscript. This work was supported by National Cancer Institute grant: U19 CA-85953.

REFERENCES

- BAYLIN, S.B., ESTELLER, M., ROUNTREE, M.R., et al. (2001). Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum Mol Genet* **10**, 687–692.
- BAUM, L.E., PETRIE, T., SOULES, G., et al. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *Ann Math Stat* **41**, 164–171.
- BEER, D.G., KARDIA, S.L.R., HUANG C.-C., et al. (2002). Gene expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* **8**, 816–824.
- BISOGNIN, A., BORTOLUZZI, S., and DANIELI, G.A. (2004). Detection of chromosomal regions showing differential gene expression in human skeletal muscle and in alveolar rhabdomyosarcoma. *BMC Bioinform* **5**, 68.
- BJORKQVIST, A.M., HUSGAFVEL-PURSIAINEN, K., ANTTILA, S., et al. (1998). DNA gains in 3q occur frequently in squamous cell carcinoma of the lung, but not in adenocarcinoma. *Genes Chromosomes Cancer* **22**, 79–82.
- CARON, H., VAN SCHAIK, B., VAN DER MEE, M., et al. (2001). The human transcriptome map: clustering of highly expressed genes in chromosomal domain. *Science* **291**, 1289–1292.
- COHEN, B.A., MITRA, R.D., HUGHES, J.D., et al. (2000). A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat Genet* **26**, 183–186.
- COX, D.R. (1972). Regression models and life tables *J R Soc Stat B* **34**, 187–220.
- CUI, Y., ZHOU, M., and WONG, W.H. (2004). Integrated analysis of microarray data and gene function information. *OMICS* **8**, 106–117.
- FOURER, R., GAY, D.M., KERNIGHAN, B.W. (1993). *AMPL: a modeling language for mathematical programming*. (Scientific Press, New York).
- FUJII, T., DRACHEVA, T., PLAYER, A., et al. (2002). A preliminary transcriptome map of non-small cell lung cancer. *Cancer Res* **62**, 3340–3346.
- GAO, F., FOAT, B.C., and BUSSEMAKER, H.J. (2004). Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinform* **5**, 31.
- HUANG, J., WEI, W., ZHANG, J., et al. (2004). Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics* **1**, 287–299.
- HUSING, J., ZESCHNIGK, M., BOES, T., et al. (2003). Combining DNA expression with positional information to detect functional silencing of chromosomal regions. *Bioinformatics* **19**, 2335–2342.
- KOSKI, T. (2002). *Hidden Markov models for bioinformatics* (Springer/Kluwer Academic Publishers, New York).
- KNUUTILA S., BJORKQVIST, A.M., AUTIO, K., et al. (1998). DNA copy number amplification in human neoplasms: review of comparative genomic hybridization studies. *Am J Pathol* **152**, 1107–1123.
- KNUUTILA, S., AALTO, Y., AUTIO, K., et al. (1999). DNA copy number losses in human neoplasms. *Am J Pathol* **155**, 683–694.
- LEE, T.I., RINALDI, H.J., ROBERT, F., et al. (2002). Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804.
- LEROUX, B.G. (1992). Maximum-likelihood estimation for hidden Markov models. *Stochastic Proc Appl* **40**, 127–143.
- LEROUX, B.G., and PUTERMAN, M.L. (1992). Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* **48**, 545–558.
- LEVIN, A.M., LEVIN, A., and KARDIA, S.L.R. (2001). A physical transcriptome map for chromosome level analysis of gene expression data. Presented at the American Society of Human Genetics 51st Annual Meeting, San Diego.
- LEVIN, A.M., GHOSH, D., CHO, K.R., et al. (2005). A model-based scan statistic for identifying extreme chromosomal regions of gene expression in human tumors. *Bioinformatics* **21**, 2867–2874.
- LIN, M., WEI, L.J., SELLERS, W.R., et al. (2004). dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* **20**, 1233–1240.
- MCLACHLAN, G., and PEEL, D. (2000). *Finite mixture model* (Wiley, New York).
- NARIAI, N., KIM, S., IMOTO, S., et al. (2004). Using protein-protein interactions for re-finig gene networks estimated from microarray data by Bayesian networks. *Proc Pac Symp Biocomput* **9**, 336–347.
- POLLACK, J.R., PEROU, C.M., ALIZADEH, A.A., et al. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23**, 41–46.
- POLLACK, J.R., SRLIE, T., PEROU, C.M., et al. (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99**, 12963–12968.
- RABINER, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* **77**, 257–286.
- REDNER, R.A. and WALKER, H.F. (1984). Mixture densities, maximum likelihood and the EM algorithm. *Soc Indust Appl Math Rev* **26**, 195–239.

- ROSS, S.M. (1993). *Introduction to Probability Models*, 5th ed. (Academic Press, San Diego).
- SPELLMAN, P.T., and RUBIN, G.M. (2002). Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *J Biol* **1**, 5.
- TAMADA, Y., KIM, S., BANNAI, H., et al. (2003). Estimating gene networks from gene expression data by combining Bayesian netowrk model with promoter element detection. *Bioinformatics Suppl.* **2**, 227–236.
- VOGELSTEIN, B., and KINZLER, K.W. (2004). Cancer genes and the pathways they control. *Nat Med* **10**, 789–799.
- ZHOU, X., KAO, M.J., and WONG, W.H. (2002). Transitive functional annotation by shortest path analysis of gene expression data. *Proc Natl Acad Sci USA* **99**, 12783–12788.

Address reprint requests to:

*Dr. Chiang-Ching Huang
Department of Preventive Medicine
Northwestern University
680 N. Lake Shore Dr., Ste. 1102
Chicago, IL 60611-4402*

E-mail: huangcc@northwestern.edu

Clinically Relevant Subsets Identified by Gene Expression Patterns Support a Revised Ontogenic Model of Wilms Tumor: A Children's Oncology Group Study^{1,2}

Samantha Gadd*, Vicki Huff[†],
Chiang-Ching Huang[‡], E. Cristy Ruteshouser[†],
Jeffrey S. Dome[§], Paul E. Grundy[¶],
Norman Breslow[#], Lawrence Jennings^{*},
Daniel M. Green^{**}, J. Bruce Beckwith^{††},
and Elizabeth J. Perlman*

*Department of Pathology, Northwestern University's Feinberg School of Medicine and Robert H. Lurie Cancer Center, Chicago, IL; [†]Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX; [‡]Department of Preventive Medicine, Northwestern University's Feinberg School of Medicine and Robert H. Lurie Cancer Center, Chicago, IL; [§]Division of Oncology, Children's National Medical Center, Washington, DC; [¶]Departments of Pediatrics and Oncology, Cross Cancer Institute and the University of Alberta, Edmonton, Alberta; [#]Department of Biostatistics, University of Washington and the Fred Hutchinson Cancer Research Center, Seattle, WA; ^{**}Department of Epidemiology and Cancer Control, Saint Jude Children's Research Hospital, Memphis, TN; ^{††}Department of Pathology and Human Anatomy, Loma Linda University, Missoula, MT

Abstract

Wilms tumors (WT) have provided broad insights into the interface between development and tumorigenesis. Further understanding is confounded by their genetic, histologic, and clinical heterogeneity, the basis of which remains largely unknown. We evaluated 224 WT for global gene expression patterns; *WT1*, *CTNNB1*, and *WTX* mutation; and 11p15 copy number and methylation patterns. Five subsets were identified showing distinct differences in their pathologic and clinical features: these findings were validated in 100 additional WT. The gene expression pattern of each subset was compared with published gene expression profiles during normal renal development. A novel subset of epithelial WT in infants lacked *WT1*, *CTNNB1*, and *WTX* mutations and nephrogenic rests and displayed a gene expression pattern of the postinduction nephron, and none recurred. Three subsets were characterized by a low expression of *WT1* and intralobar nephrogenic rests. These differed in their frequency of *WT1* and *CTNNB1* mutations, in their age, in their relapse rate, and in their expression similarities with the intermediate mesoderm *versus* the metanephric mesenchyme. The largest subset was characterized by biallelic methylation of the imprint control region 1, a gene expression profile of the metanephric mesenchyme, and both interlunar and perilobar

Abbreviations: WT, Wilms tumor; ICR1, imprint control region 1; ICR2, imprint control region 2; MM, metanephric mesenchyme; LOI, loss of imprinting; LOH, loss of heterozygosity; ROI, retention of imprinting; UPD, uniparental disomy

Address all correspondence to: Elizabeth J. Perlman, MD, Children's Memorial Hospital, 2300 Children's Plaza, Box 17, Chicago, IL 60614. E-mail: eperlman@childrensmemorial.org

¹This work is supported by grants from the National Institutes of Health: U10CA42326 (N.B., D.M.G., E.J.P.), U10CA98543 (J.S.D., P.E.G., E.J.P.), UO1CA88131 (E.J.P., C.C.H.), CA34936 (V.H.), DK069599 (V.H.), CA16672 (V.H.), RP100329 (V.H.), and RP110324 (V.H.).

²This article refers to supplementary materials, which are designated by Tables W1 and W2 and are available online at www.neoplasia.com.

Received 23 April 2012; Revised 28 June 2012; Accepted 4 July 2012

nephrogenic rests. These data provide a biologic explanation for the clinical and pathologic heterogeneity seen within WT and enable the future development of subset-specific therapeutic strategies. Further, these data support a revision of the current model of WT ontogeny, which allows for an interplay between the type of initiating event and the developmental stage in which it occurs.

Neoplasia (2012) 14, 742–756

Introduction

The initiation and progression of the most common adult cancers result from a stepwise accumulation of multiple genetic events within a finite number of pathways occurring over many years. In contrast, the initiation of neoplasia in children results from one to two genetic events that occur over the course of months rather than years. These events usually involve genes responsible for normal development and result in tumors that closely resemble cells within the developing embryo. They are also often the same genetic events that participate in the development of adult tumors. Therefore, pediatric embryonal neoplasms provide invaluable insights into normal development and into both adult and childhood neoplasia. The investigation of Wilms tumor (WT), one of the most common tumors of childhood, is a remarkable illustration of this. This unique success is due in part to the fact that WT is the only embryonal neoplasm that arises within precursor lesions known as nephrogenic rests, of which there are two predominant types: perilobular and intralobular [1]. WTs are also capable of showing a striking spectrum of appearances ranging from undifferentiated “blastemal” tumors to “teratoid” tumors composed of a mixture of differentiated skeletal muscle, chondroid, and a variety of epithelial cell types. This heterogeneity implies a complexity to the underlying causes of WT that has fascinated investigators for decades.

Two genetic loci have consistently been associated with the pathogenesis of WT, the *WT1* gene at 11p13, and the *WT2* locus at 11p15. *WT1* encodes a transcription factor important in multiple phases of normal renal, gonadal, and cardiac development [2,3]. Germline mutations of *WT1* result in syndromes, including Denys-Drash and Wilms tumor-aniridia-genitourinary malformation-mental retardation: both are characterized by an increased risk of WT and abnormal genitourinary development [4–6]. Somatic mutations of *WT1* are seen in 10% to 20% of sporadic WT [7–9]. Frequently accompanying *WT1* mutation is canonical Wnt activation, most commonly due to activating mutation of β-catenin (*CTNNB1*) [10,11]. Inactivating mutations of *WTX*, a protein that contributes to β-catenin degradation, may also occur in 15% to 20% of patients with WT, regardless of their *WT1* mutation status [12–15]. While canonical Wnt- activating mutations likely occur subsequent to *WT1* mutation [16,17], whether or not Wnt activation is required for tumor development after *WT1* mutation is not clear. Nor is the role of Wnt activation in WT that lack *WT1* mutation known.

The *WT2* chromosomal region came to scientific attention with the observations of 11p15 loss of heterozygosity (LOH) or loss of imprinting (LOI) in a large proportion of sporadic WT [18], and 11p15 uniparental disomy (UPD) or duplication in patients with Beckwith-Wiedemann syndrome (BWS), which carries an increased risk of WT and developmental abnormalities including organ and limb overgrowth [19–22]. 11p15 methylation abnormalities resulting in WT are accompanied by aberrant methylation at imprint control region 1

(ICR1), resulting in biallelic expression of *IGF2*, a gene normally expressed only from the paternally inherited allele [21,23]. Although 11p15 clearly plays a critical role in the pathogenesis of WT, the observation of 11p15 LOH in normal tissue from some WT patients [24] and the lack of tumors arising in mutant mice with ICR1 LOI [25] imply that biallelic expression of *IGF2* alone is insufficient for tumor development.

Other loci implicated in WT infrequently are the familial predisposition loci FWT1 at 17q12-q21 and FWT2 at 19q13.4 [26,27]. The documented association between relapse and LOH for 1p and 16q [28] is being used to stratify patients within the current Children's Oncology Group therapeutic protocols. The critical genes within these regions are not known.

In summary, despite the wealth of knowledge that the investigation of WT has provided, much remains unknown and further progress is made difficult by the genetic, histologic, and clinical heterogeneity that characterizes WT. The goal of this study was to investigate patterns of global gene expression and known genetic and epigenetic changes in a large number of prospectively identified WTs to identify and characterize distinctive subsets that may merit therapeutic stratification or respond to specific therapies. In addition, the recent availability of large data sets of gene expression patterns identified in microdissected samples of different embryologic stages and in different cell types during normal renal development [29,30] offers a unique opportunity to place each of these WT subsets within their developmental context. We have accomplished this, and here we provide a revised ontogenetic model for the development of WT.

Materials and Methods

Clinical Samples

Samples were taken from a case-cohort sampling previously described [31]. Briefly, all patients with Favorable Histology Wilms Tumor (FHWT) registered on the National Wilms Tumor Study 5 for whom pretreatment tumor tissues were available were identified. From the resulting 1451 patients, all those known to have relapsed and a random sample including approximately 30% of the remaining were identified. The resulting 600 patients were randomly divided into two groups of 300 patients each. The use of case-cohort sampling allows for the resource-efficient investigation of clinical outcome in a tumor characterized by a low relapse rate. Institutional review board approval and informed consent were obtained for all tumor specimens. Frozen sections of each sample confirmed more than 80% viable cellular tumor. Pathologic features (diagnosis, local stage, histologic pattern, skeletal muscle quantification, presence, and type of nephrogenic rests) were recorded prospectively at the time of central pathology review. Skeletal muscle quantification represents the estimated proportion of the tumor volume containing cells with cross striations.

Gene Expression Analysis

Samples were hybridized to Affymetrix U133A arrays, scanned, subjected to quality control standards, and normalized as previously described [32]. Gene expression patterns were identified through unsupervised analysis using average linkage clustering with CLUSTER and displayed with TreeView (<http://rana.lbl.gov/EisenSoftware.htm>) [33]. Gene Set Enrichment Analysis v.2.5 (GSEA) was used to identify those gene lists that best define the different tumor subsets. GSEA ranks the expression of each gene based on its correlation with one of two phenotypes being compared. It then determines the presence and ranking of each gene within available independent gene lists queried. From this ranking, it calculates an enrichment score that reflects the degree to which genes in the independent gene list are overrepresented. The normalized enrichment score (NES) takes into account the number of genes within the independent gene set. Leading-edge genes are those that account for the gene set's enrichment signal.

Loss of Heterozygosity for 1p, 16q

Polymerase chain reaction (PCR) microsatellite analysis was performed prospectively during National Wilms Tumor Study 5 as previously described in detail [28].

Methylation Analysis at 11p15 ICR1 and ICR2

Methylation of the paternally methylated ICR1 (controlling *H19* and *IGF2* expression) and the maternally methylated ICR2 (*KvDMR1*, whose methylation is independent of ICR1) was determined using specific restriction sites recognized by methylation-sensitive enzymes as previously described [34]. Retention of imprinting (ROI) was defined as 30% to 70% methylation of both ICR1 and ICR2, LOI was defined as 80% to 100% methylation of ICR1 and 30% to 70% methylation of ICR2, and LOH was defined as 80% to 100% methylation of ICR1 and 0% to 20% methylation of ICR2. Tumors with values outside these ranges were not classified.

11p15 Copy Number Analysis

Samples with 11p15 LOH were analyzed for 11p15 copy number using multiplex ligation-dependent probe amplification (MLPA) to assess for UPD (two identical copies of 11p15) using four probe sets to loci on 11p15, as previously described [34].

Mutation Analysis for *WT1*, *CTNNB1*, and *WTX*

Tumor DNAs were assessed for *WT1* point mutations by analysis of PCR products from all 10 exons and for partial or complete deletion of *WT1* by quantitative real-time PCR analysis using amplicons for the promoter/exon 1, exon 2/3, exon 4, exon 5, exon 6, exon 7, exon 8/9, and exon 10 of *WT1*, as previously described [14,34,35]. Tumor DNAs were analyzed for point mutations in *CTNNB1* (exons 3, 7, and 8), as previously described [10,14]. To detect *WTX* mutations, the entire coding region was amplified as a single PCR product and sequenced. Quantitative PCR was performed to detect deletions, as previously described [14].

Results

Identification of Subsets of WT by Gene Expression Patterns

We first analyzed the global gene expression pattern of the 224 samples from the first group of 300 patients that passed quality control

parameters. Unsupervised hierarchical clustering using the top 2000, 4000, and 10,000 most variable genes demonstrated two subsets (S1 and S2) that were stable in each analysis, identifying the same samples within each subset (Figure 1A, blue and red bars). S1 and S2 tumors were then removed, the most highly variable genes were re-identified from the remaining tumors, and hierarchical clustering was again performed. This did not reveal new subsets with consistent expression patterns.

A key gene differentially expressed in S2 tumors is *WT1* (Figure 1A). To better assess the role of *WT1* in the overall expression pattern, 54 unique genes were identified due to their Pearson Correlation Coefficient (PCC) scores of less than -0.6 or greater than 0.6 for both available *WT1* probe sets compared with all probe sets on the array. Hierarchical analysis using these 54 genes again identifies S1 and S2 (Figure 1B, blue and red bars), as well as an additional subset, S3 (green bar in Figure 1B). Lastly, two small groups of tumors flank the S2 tumors and are grouped together as S4 (Figure 1B, purple bars). There is no evidence of clustering of the S3 and S4 tumors within the original unsupervised analysis (Figure 1A). Tumors outside of S1 to S4 (the majority of tumors) are classified as S5. The top 100 genes differentially expressed in each of S1, S2, S3, and S4 compared with S5 as determined by GSEA are provided in Table W1. The gene expression data are deposited in National Center for Biotechnology Information's Gene Expression Omnibus, accessible at GEO Series accession no. GSE31403 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31403>).

Clinical and Pathologic Features of FHWT Subsets

The features of WT within S1 to S4 are listed in Table 1, and these results are compared with the remaining S5 tumors in Table 2 and are summarized below:

- Subset 1 tumors ($n = 11$) occur in infants and all have a distinctive epithelial tubular differentiated histologic pattern throughout. All lack both nephrogenic rests and skeletal muscle differentiation, and none relapsed.
- Subset 2 tumors ($n = 23$) present at an early median age, have a mixed (triphasic) histology and commonly arise within intralobar nephrogenic rests (ILNRs). Of 22 evaluable S2 tumors, 9, 12, and 1 display muscle differentiation comprising 5% or less, 10% to 25%, and 80% of the tumor volume, respectively. Two relapses occurred, and two additional patients developed contralateral tumors after therapy.
- Subset 3 tumors ($n = 21$) are pathologically similar to S2, with less skeletal muscle differentiation. Of 21 evaluable tumors, 13, 4, and 4 showed absent, 5% or less, and 10% to 25% skeletal muscle, respectively. Seven relapses occurred.
- Subset 4 tumors ($n = 11$) are also pathologically similar to S2: skeletal muscle differentiation was identified in 6 of 10 evaluable S4 tumors, with 3, 2, and 1 showing 5% or less, 10%, and 40% skeletal muscle, respectively. Six patients relapsed.
- Subset 5 tumors ($n = 158$): 70 tripasic (mixed), 76 blastemal predominant, 9 epithelial predominant, and 3 stromal predominant. Skeletal muscle is present in 27 (17%) of 156 evaluable tumors (24/27 with <5%, 2 with 5%-15%). Of 154 evaluable tumors, 38 (25%) contain perilobar nephrogenic rests (PLNR) and 5 of 38 also contain ILNRs. ILNRs alone are identified in 23 (15%) of 154 tumors. Five of the 224 patients in this study had bilateral WT at diagnosis, and all belong to S5. Fifty S5 patients relapsed. It should be noted that while the relapse rates in these five subsets show clear differences, they do not achieve statistical significance; the small sample size may contribute to this.

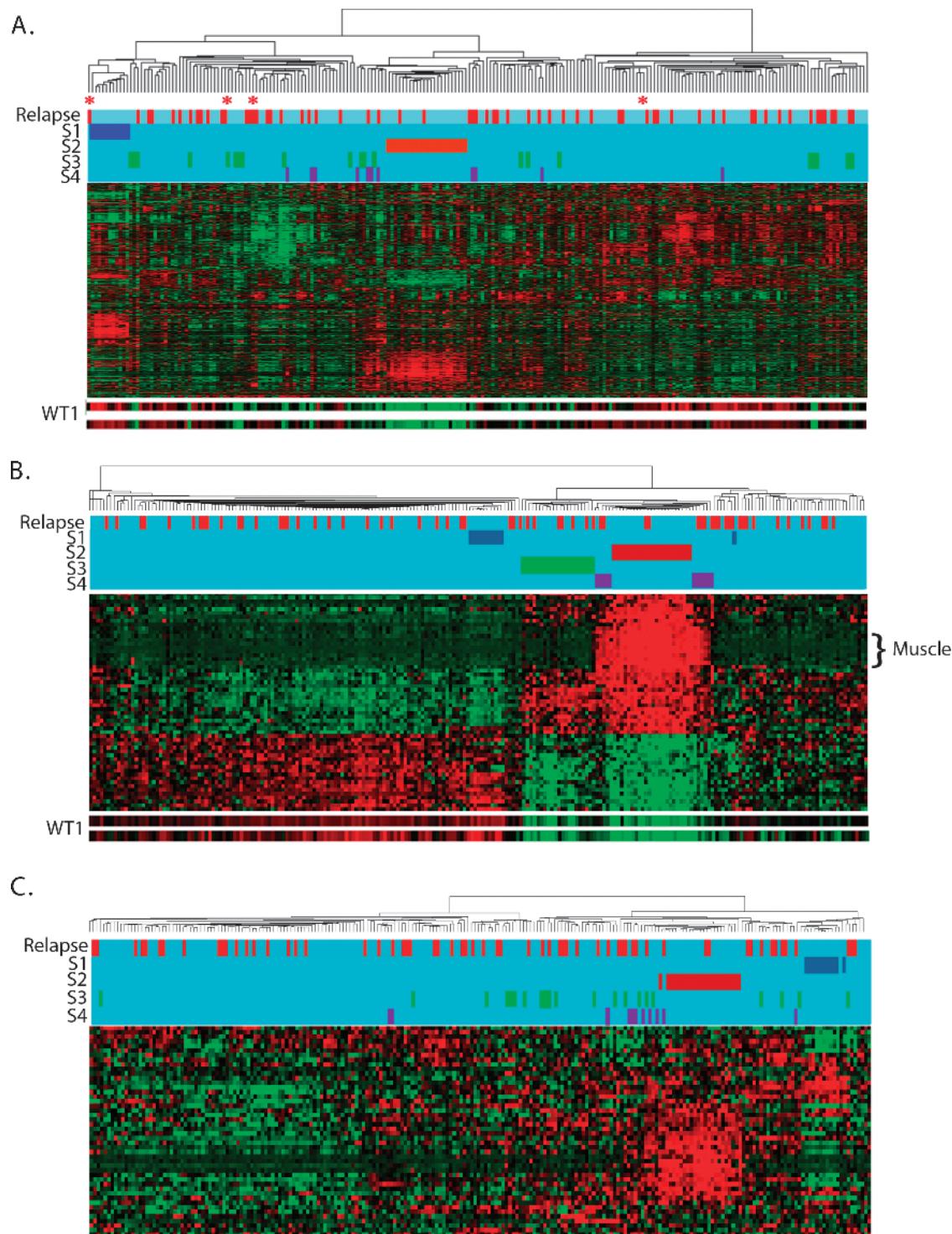


Figure 1. Hierarchical analysis of 224 FHWT: (A) Unsupervised hierarchical analysis using the 4000 probe sets with the highest coefficient of variation. Two subsets are designated by the blue (S1) and red bars (S2). Also shown are the tumors identified in B in green (S3) and purple (S4). Tumors that relapsed are designated in red at the top of the dendrogram. Expression of genes, clustered on the *y* axis, is shown with levels ranging from high (red) to low (green). The expressions of the two *WT1* probe sets are illustrated separately at the bottom. Four tumors outside S1 that show solely epithelial tubular differentiation are marked with a red asterisk. See Table W1 for the top 100 genes differentially expressed in S1 to S4 compared with S5. (B) Hierarchical analysis using the 54 genes with a Pearson correlation coefficient greater than 0.60 or less than -0.60 for both available *WT1* alleles. S1 and S2 are readily identified (blue and red bars). Two additional subsets are now apparent, indicated by green (S3) and purple bars (S4). Genes associated with muscle differentiation are marked. (C) Hierarchical analysis of Wnt targets with a coefficient of variation greater than 0.06. S1 (blue) and S2 (red) clustered tightly together. Tumors in S3 (green) and S4 (purple) did not show evidence of strong Wnt activation and did not cluster.

Table 1. Features of Tumors in Subsets.

	Relapse	Stage	Histology (Weight)	Age (months)	Rests	11p15 Status	1p LOH	16q LOH	Mutations		
									WT	CTNNB1	WTX
Subset 1											
WT00-082	No	I	Epithelial (390 g)	8	None	ROI	NO	NO	No	No	No
WT00-269	No	I	Epithelial (789 g)	10	None	ROI	NO	NO	No	No	No
WT00-177	No	I	Epithelial (333 g)	13	None	ROI	NO	NO	No	No	No
WT00-175	No	I	Epithelial (580 g)	14	None	ROI	NO	NO	No	No	No
WT00-054	No	I	Epithelial (207 g)	17	None	ROI	NO	NO	No	No	No
WT00-002	No	I	Epithelial (31 g)	18	None	ROI	NA	NA	No	No	No
WT00-173	No	I	Epithelial (86 g)	39	None	ROI	NO	NO	No	No	No
WT00-167	No	II	Epithelial (393 g)	91	None	LOI	NO	NO	No	No	No
WT00-153	No	I	Epithelial (485 g)	15	None	ROI	NO	NO	No	No	No
WT00-086	No	I	Epithelial (230 g)	6	None	ROI	NO	NO	No	No	No
WT00-185	No	I	Epithelial (315 g)	7	NE	ROI	NO	NO	No	No	No
Subset 2											
WT00-196	No	I	Mixed	3	ILNR	ROI	NO	NO	No	No	No
WT00-270	No	I	Mixed	14	ILNR	LOH	NO	NO	No	No	No
WT00-171	No	II	Mixed	43	None	LOH	LOH	NO	No	No	No
WT00-237	No	III	Mixed	11	ILNR	LOH	NO	NO	No	No	Yes (XX)
WT00-216	No	II	Mixed	20	ILNR	ROI	NO	NO	No	Exon 8	No
WT00-264	No	I	Mixed	41	None	LOI	NO	NO	No	Exon 8	No
WT00-166	No	I	Mixed	37	None	LOH	NO	NO	No	Exon 8	Yes (XX)
WT00-128	No	I	Mixed	57	ILNR	LOH	NO	NO	No	Exon 7	Yes (XY)
WT00-074	No	II	Mixed	7	ILNR	LOH	NA	NA	No	Exon 3	No
WT00-290	Yes	II	Mixed	27	ILNR	ROI	NO	NO	No	Exon 3	No
WT00-027	No	III	Stromal	5	ILNR	LOH	NO	NO	Yes	Exon 3	No
WT00-232	No	I	Mixed	12	ILNR	ROI	NO	NO	Yes	Exon 3	No
WT00-152	No	II	Mixed	70	ILNR	ROI	NO	NO	Yes	Exon 3	No
WT00-254	No	III	Mixed	35	None	ROI	NO	NO	Yes	Exon 3	No
WT00-288	No	II	Mixed	8	ILNR	LOH	NO	NO	Yes	Exon 3	Yes* (XY)
WT00-156	No	II	Mixed	9	ILNR	LOH	NO	NO	Yes	Exon 3	No
WT00-061	Yes	II	Stromal	9	ILNR	LOH	NO	NO	Yes	Exon 3	No
WT00-072	No	II	Mixed	13	ILNR	LOH	ND	ND	Yes	Exon 3	No
WT00-113	No	III	Stromal	9	ILNR	LOH	NO	NO	Yes	Exon 3	No
WT00-244	No	III	Mixed	12	ILNR	ROI	NO	NO	Yes	Exon 3	No
WT00-064	No	III	Mixed	13	None	LOH	NO	NO	Yes	Exon 8	Yes (XY)
WT00-193	No	II	Mixed	111	ILNR	LOH	ND	ND	Yes	Exon 7	Yes (XY)
WT00-112	No	III	Blastemal	39	None	ND	NO	LOH	ND	ND	ND
Subset 3											
WT00-110	No	II	Mixed	57	None	LOI	NO	NO	No	No	No
WT00-285	No	III	Mixed	57	ILNR	LOH	NO	NO	No	No	No
WT00-142	No	III	Mixed	55	None	ROI	NO	NO	No	No	No
WT00-127	No	I	Mixed	17	ILNR	LOI	NO	NO	No	No	No
WT00-135	Yes	II	Mixed	5	ILNR	ROI	NO	NO	No	No	No
WT00-220	No	II	Mixed	36	ILNR	LOH	LOH	LOH	No	No	No
WT00-277	Yes	III	Mixed	120	NA	ROI	NO	NO	No	No	No
WT00-209	No	III	Mixed	35	None	ROI	NO	NO	No	No	Yes (p.XX)
WT00-098	No	II	Mixed	41	ILNR	LOI	NO	NO	No	No	Yes (p.XX)
WT00-291	Yes	III	Mixed	28	ILNR	ROI	NO	NO	No	Exon 3	No
WT00-205	No	I	Mixed	44	ILNR	ROI	NO	NO	Yes	Exon 3	Yes (p.XY)
WT00-214	No	II	Mixed	0.2	ILNR	ROI	NO	NO	Yes	Exon 3	No
WT00-038	No	I	Mixed	21	ILNR	ROI	NO	NO	Yes	Exon 3	No
WT00-145	Yes	III	Mixed	10	ILNR	ROI	NO	NO	Yes	Exon 3	No [†]
WT00-106	Yes	I	Mixed	10	ILNR	LOH	NO	NO	Yes	Exon 3	No
WT00-023	No	I	Mixed	11	ILNR	LOH	LOH	LOH	Yes	Exon 3	No
WT00-297	No	II/IV	Mixed	39	ILNR	LOH	NO	NO	Yes	Exon 8	No
WT00-192	No	III	Blastemal	36	ILNR	ROI	NO	NO	Yes	No	Yes (XY)
WT00-121	Yes	III/IV	Blastemal	87	ILNR	LOH	NO	NO	Yes	No	No [†]
WT00-052	Yes	II	Blastemal	53	ILNR	LOI	NO	NO	Yes	No	Yes (XX)
WT00-122	No	III	Mixed	58	None	LOH	NO	NO	Yes(p)	No	No
Subset 4											
WT00-026	Yes	I	Mixed	8	ILNR	LOH	NO	NO	Yes	No	No
WT00-204	Yes	IV	Mixed	73	None	ROI	NO	NO	No	Exon 8	Yes (XX)
WT00-076	No	III	Mixed	72	None	LOI	NO	NO	No	No	Yes (XY)
WT00-172	No	II	Mixed	38	None	LOH	NO	NO	No	No	No
WT00-206	Yes	III	Mixed	65	ILNR	LOH	NO	LOH	No	Exon 8	No [†]
WT00-276	No	I	Mixed	20	ILNR	LOH	NO	NO	No	Exon 3	No
WT00-018	No	II	Mixed	56	Incon	LOH	ND	ND	No	Exon 3	No
WT00-021	No	II	Mixed	17	None	LOI	NO	NI	No	Exon 8	No
WT00-050	Yes	I	Mixed	15	None	ROI	NO	NO	No	No	No
WT00-190	Yes	II	Mixed	73	None	ND	NO	NO	ND	ND	ND
WT00-168	Yes	IV	Blastemal	37	NE	LOI	NO	LOH	No	No	No

ND indicates not done; NE, not evaluable; NI, not informative; p, partial.

^{*}Missense mutation, unknown significance.[†]Missense mutation, known single nucleotide polymorphism.

Table 2. Summary of Features in Subsets.

	No. Patient (%)	Nephrogenic Rests (%)	Pathologic Features		Clinical Features		ICR1, ICR2 Methylation (%)		1p LOH (%)		16q LOH (%)		Mutation (%)	
			ILNR	PLNR	Predominant Histology	Skeletal Muscle (%)	Median Age (months)	No. Relapse (%)	11p15 LOH	11p15 ROI	WT1	CTNNB1 Exon 3	WTX	
S1	11 (6.2%*)	0	0	0	Epithelial	0	14	0	0	0	0	0	0	0
S2	23 (12.1%*)	74	0	Mixed	87	13	2 (3%*)	63.5	4.5	32	5	54	55	22
S3	21 (9.4%*)	76	0	Mixed	38	39	7 (11%*)	33.3	19.1	47.6	9.5	52	33	24
S4	11 (3.8%*)	30	0	Mixed	60	38	6 (16%*)	50.0	30.0	20.0	0.0	22.0	10	10
S5	158 (68.5%*)	15	25	Variable	17	43.5	50 (13%*)	37.2	43.8	19.0	14.3	26.8	ND	ND
Validation subsets														
SA	5	0	0	0	Epithelial	0	10	0	20	0	80	0	0	0
SB	7	86	0	Mixed	100	28	1	86	0	14	0	0	0	0
SC	8	38	0	Mixed	40	23	5	57.1	28.6	14.3	0	0	16.7	
SD	80	19	16	Variable	15	42	25	51	31	12	20	20	25	

*Percentage adjusted for case-cohort design.

Characterizing Subsets of FHWT Based on Gene Expression

GSEA was used to characterize the gene expression patterns in the different subsets using the gene lists provided in Table 3. S1 to S4 tumors were individually compared with S5 tumors; S3 and S4 tumors were also compared with S2 tumors. Those gene lists with a false discovery rate (FDR) of 20% or less and nominal $P \leq .05$ are provided in Table W2. These are grouped into categories according to their common leading edge genes, and these categories are provided in Table 4. The results of the above analyses are summarized below, beginning with the largest group, S5. The expression patterns of illustrative genes (designated with an asterisk [*] in the text) are shown in Figure 2.

S5 tumors, the comparison group for this study, show a pattern of expression previously reported by others in WT [32,36,37]. This includes expression of *SIX1*, *PAX2*, *EYA1*, *SALL2*, *HOXA11**, *HOXA9*, *MEOX1*, *MEIS2**, *PRAME*, *NNAT*, *CRABP2*, *FZD7*, *COL2A1*, *GPR64*, *WASF*, *HMGAA2**, *UCHL1*, and *CCND2*, most of which are known to be expressed within the early metanephric mesenchyme (MM). S5 shows gene expression heterogeneity (Figure 1A); however, no expression patterns are sufficiently stable to enable identification or validation. When analyzed alone, S5 tumors did not cluster based on 11p15 methylation, 1p or 16q LOH, nephrogenic rest, presence of muscle differentiation, histologic pattern, or relapse status.

S1 tumors show significantly decreased expression levels of the above renal developmental genes expressed in S5, and negative enrichment of genes expressed in WTs compared with fetal kidney (LI) and genes expressed in the preinduction MM (Brunskill A), illustrated by *HMGAA2**, *MEIS2**, and *HOXA11**. Instead, genes expressed subsequent to mesenchymal-to-epithelial transition (Brunskill groups D and L) are enriched in S1. Several metabolic processes are significantly enriched in S1 tumors, largely driven by *AKR1C3**, a gene expressed in the maturing epithelial component of the developing kidney [29,38]. Similarly, the Notch signaling pathway, key to renal epithelial differentiation, is significantly enriched in S1 tumors. To evaluate whether S1 gene expression was driven solely by epithelial differentiation, four non-S1 tumors with the same histologic pattern were identified from all 224 tumors. They demonstrated different gene expression patterns (*red asterisks* in Figure 1A), were older at diagnosis (56, 58, 71, and 85 months), and presented at higher stages (one stage I, one stage II, and two stage III tumors) when compared with S1. Two were associated with nephrogenic rests, two of four relapsed, and all demonstrated LOI at 11p15. These features all differ from those of S1 and indicate that the S1 expression profile is not simply a function of epithelial differentiation.

Subset 2 tumors show enrichment of a very large number of GO Biologic Processes gene lists, many of which have in common early muscle development genes, including *MYH3** (myosin), *TTN** (titin), and *ACTA1** (actin), to name a few. Notably, the degree of expression of these muscle-related genes far exceeded the histologic evidence of muscle differentiation in most tumors, which was often quite focal. A wide variety of processes involved in the morphogenesis of other cell types is also enriched. Noteworthy is the increased expression of transcription factors *TWIST1** and *PITX2**. Both cell proliferation and programmed cell death gene lists are enriched, largely driven by *TBX3**, *MYC*, *LGALS1*, *PLAGL1*, *FABP7*, and *IGF1*. Signal transduction was likewise enriched, particularly Ras signaling, driven by *IGF1*. Of the curated canonical pathways, NFAT and TGFB signaling were enriched in S2. This supports the premise previously proposed that *WT1* loss may lead to TGFB activation, resulting in NFAT induction,

Table 3. Gene Lists Analyzed by Gene Set Enrichment Analysis.

Gene Lists	Reference
All lists in GSEA v 2.5 from Gene Ontology Biologic Processes with >50 genes (C5): 225 lists	www.broadinstitute.org/gsea
All lists in GSEA v 2.5 from curated canonical pathways with >50 genes (C2): 88 lists	www.broadinstitute.org/gsea
List comparing WT to fetal kidney (LI)	[36]
Lists comparing WT with and without <i>WT1</i> mutation (TYCKO)	[42]
Lists comparing WT with and without <i>CTNNB1</i> mutation (ZIRN)	[66]
STANFORD Wnt signaling	www.stanford.edu/~rnusse/wntwindow.html
15 lists characterizing different stages and cell types within the developing kidney (Brunskill A-0)	[29]
List characterizing the intermediate mesoderm (LIN)	[30]

which in turn mediates the switch of TGFB from growth inhibitor to growth promoter [39,40]. The genes differentially expressed between WTs with and without *WT1* mutation (TYCKO) are concordantly differentially expressed in S2 tumors, including striking up-regulation of *WIFI1** and down-regulation of *HAS2* and *KLK6**. Genes differentially expressed in WT with *CTNNB1* mutation (ZIRN) are likewise enriched. Lastly, S2 tumors show significant enrichment of genes normally expressed before induction (LIN intermediate mesoderm, Brunskill group A) and in the renal interstitium (Brunskill groups M, N, and O). Taken together, S2 tumors show evidence of loss of *WT1* expression, Wnt activation, and divergent mesenchymal differentiation, all occurring quite early in renal development.

Subset 3 tumors show an overall expression pattern similar to that of S5 (Figure 1A). This is supported by the absence of enrichment of C5 gene ontology biologic processes, C2-curated canonical pathway gene lists, and renal development gene lists analyzed when comparing S3 to S5 (Table 4). However, there are significant differences between S3 and S5. More than 95% of the top genes differentially expressed between S3 and S5 tumors are also coordinately differentially expressed when comparing S2 to S5 (Table W1). Similarly, the gene lists differentially expressed in both WT1 (TYCKO) and CTNNB1 (ZIRN) mutant WT and the Stanford Wnt targets showed the same pattern in S3 and S2 tumors. These observations are consistent with the fact that S3 tumors were identified based on genes coordinately expressed

Table 4. GSEA of S1 to S4 Compared with S5.

Subset 1	Median NES: S1 vs S5	Median NOM P: S1 vs S5	Median FDR: S1 vs S5				
(A) GO biologic processes							
Mitosis/cell cycle	-1.65	.03	0.07				
DNA replication	-1.56	.05	0.08				
Metabolic processes	1.67	.004	0.08				
(B) Curated and published lists							
LI_FETAL_VS_WT_KIDNEY_DN	-1.64	.02	0.02				
LI_FETAL_VS_WT_KIDNEY_UP	1.55	.03	0.08				
NOTCH_SIGNALING_PATHWAY	1.78	<.001	<0.01				
(C) Renal developmental lists							
Brunskill A: Early MM	-1.41	.03	0.17				
Brunskill D: Renal vesical, S body	1.56	.03	0.09				
Brunskill L: Epithelial differentiation	1.5	.05	0.07				
S2, S3, S4	Median NES: S2 vs S5	Median NOM P: S2 vs S5	Median FDR: S2 vs S5	Median NES: S3 vs S5	Median NES: S4 vs S5	Median NES: S3 vs S2	Median NES: S4 vs S2
(A) GO biologic processes							
Muscle development	1.75	<.001	0.01	NS	1.62	-1.74	NS
Morphogenesis and development	1.71	<.008	0.02	NS	1.71	-1.67	NS
Regulation of catalytic activity	1.61	<.001	0.05	NS	1.43	-1.54	NS
Regulation of transferase/kinase	1.46	.001	0.02	NS	NS	NS	NS
Cell cycle	1.49	.012	0.12	NS	NS	-1.47	NS
Signal transduction	1.77	.002	0.01	NS	1.71	-1.65	NS
Metabolic process	1.59	.006	0.06	NS	NS	NS	NS
(B) Curated and published lists							
TYCKO_UP_IN_WT1_WILDTYPE_WT	-1.68	<.001	0.01	-1.71	-1.55	1.50	1.51
TYCKO_UP_IN_WT1_MUT_WT	1.40	.002	0.13	1.5	1.49	-1.41	-1.54
ZIRN_UP_IN_CTNNB1_MUT_WT	1.60	<.001	0.03	1.58	1.74	-1.49	NS
ZIRN_DOWN_IN_CTNNB1_MUT_WT	-1.66	.004	0.01	-1.79	NS	NS	NS
STANFORD_WNT_GENES_UP	1.86	<.001	0.01	1.92	1.76	-1.67	NS
NFATPATHWAY	1.80	<.001	0.06	NS	1.70	-1.82	NS
TGF_BETA_SIGNALING_PATHWAY	1.51	.02	0.18	NS	NS	NS	NS
(C) Renal developmental gene lists							
LIN intermediate mesoderm	1.6	.004	0.004	NS	1.66	-1.64	NS
Brunskill A: Early MM	1.58	.003	0.03	NS	1.60	-1.57	NS
Brunskill M, N, O: interstitium	1.66	.003	0.03	NS	1.59	-1.62	NS

Categories of gene lists with an FDR of 20% or less and nominal $P \leq .05$.

Individual gene lists and their leading-edge genes are provided in Table W2.

FDR indicates false discovery rate (median of all gene lists in category); NES, nominal enrichment Score (median of all gene lists in category); NOM, nominal P value (median of all gene lists in category); NS, not significant ($P > .05$ or FDR > 0.20).



◆ S1 ■ S2 ▲ S3 △ S4 ◆ S5

Figure 2. Patterns of gene expression within the different subsets of FHWT: The log expression levels (low to high) of selected genes are plotted on the y axis. The x axis reflects an arbitrary tumor number, grouping the different tumor types starting with S1 in blue, followed S2 in red, S3 in green, S4 in purple, and the remaining tumors (S5) in turquoise.

with *WT1* and with the observed down-regulation of *WT1* in S3 tumors (Figure 1B). Direct comparison of S3 with S2 tumors reveals negative enrichment of most of the same GO categories, canonical pathways, and renal developmental categories previously identified when comparing S2 with S5. In particular, this includes a low expression of genes involved with skeletal muscle differentiation in S3 (Figure 1B), as well as NFAT, TGFB, and Ras signaling. Lastly, whereas S3 tumors show greater canonical Wnt activation than S5 tumors do ($\text{NES} = 1.9$, nominal $P < .001$, FDR = 1%), there is negative enrichment of Stanford Wnt targets in S3 compared with S2 ($\text{NES} = -1.67$, nominal $P < .001$, FDR = 0.03). In summary, while similarities between S2 and S3 point toward low *WT1* expression, the differences lie in decreased canonical Wnt activation and decreased divergent mesenchymal differentiation in S3 tumors and disruption of S3 tumors later in renal development than that of S2 tumors.

S4 tumors show an overall gene expression pattern quite similar to that of S2 (Figure 1B), and none of the biologic processes, canonical pathways, or renal developmental gene lists are significantly enriched in the comparison of S2 with S4. The one list that differed included genes differentially expressed in tumors with *WT1* mutation, which were downregulated in S4 (TYCKO; Table 4). Comparing S4 to S5, more than 95% of the top differentially expressed genes are also concordantly differentially expressed in S2, and S4 shows significant enrichment of many of the same gene lists enriched in S2 compared with S5, including one of two Wnt signaling gene lists (Table 4). Wnt activation was not evident in Figure 1. In summary, the S4 gene expression pattern is quite similar to S2, supporting an origin very early in renal development, although with somewhat less evidence of Wnt activation and *WT1* loss.

WT1, CTNNB1, and WTX Mutation Analysis

S1 to S5 show differences in expression of genes associated with *WT1* loss and Wnt activation. Therefore, mutation analysis was performed for *WT1*, *CTNNB1*, and *WTX* within S1 to S4 tumors (Table 1). S1 lacks mutations in any of these genes. Whereas ~50% of S2 and S3 tumors contain abnormalities of *WT1*, only 10% of S4 tumors contain *WT1* genetic events, corresponding to the different enrichment pattern seen in the TYCKO gene lists. The documentation of *WT1* mutation in only ~50% of S2 to S4 tumors that show loss of *WT1* expression suggests that other mechanisms for *WT1* mRNA loss exist. S2 to S4 show variable evidence of *CTNNB1* exon 3 mutations (those known to stabilize β -catenin); the highest frequency is seen in S2, and the lowest, in S4. This parallels the Wnt target expression patterns in these subsets. *WTX* mutation is identified in 13 (24%) of 53 S2 to S4 tumors and did not vary significantly across these subsets. As previously reported, *WTX* mutation was associated with *CTNNB1* exon 7, 8 mutations [41]. Comparison of gene expression patterns did not show detectable differences when comparing those S2 to S4 tumors with and without *WTX* mutations (data not shown).

Canonical Wnt Activation

To further investigate Wnt signaling and to determine whether mutation analysis was merited for S5, hierarchical analysis was performed using the Stanford Wnt targets (Figure 1C). S1 tumors cluster together because of increased expression of targets such as *LEF1* and *FZD2* and decreased expression of *CCND1* and *JAG1*. S2 tumors show strong expression of a large number of canonical Wnt targets including those involved in muscle differentiation (*PITX**, *ISLR*) and extracellular inhibitors of Wnt signaling (*WIF1**, *DKK1*, and

*DKK2**). Overexpression of these inhibitors has been attributed to cellular resistance to feedback inhibition in the setting of constitutive Wnt activation [42]. Three S2 tumors lacking *CTNNB1* or *WTX* mutations show strong evidence of Wnt activation, suggesting the presence of other mechanisms for canonical Wnt activation, as reviewed previously [43]. S3 and S4 tumors show no clear evidence of canonical Wnt activation and relative decrease in the expression of Wnt inhibitors *WIF1**, *DKK1*, and *DKK2**, suggesting insufficient constitutive Wnt activation to trigger this feedback inhibition. S5 lacks overt patterns of differential expression of Wnt targets in these analyses.

11p15 Analysis

Methylation analysis of ICR1 and ICR2 was performed on 205 of 224 tumors with sufficient remaining frozen tissue (Tables 1 and 2). This confirms the presence of tumors with the biallelic paternal but not maternal methylation patterns (Figure 3A). The considerable majority of S1 tumors retain the normal methylation pattern. LOI is rare in S2 and increases in both S3 and S4. In contrast, LOH is high in S2 and decreases in S3. All of the S2 to S4 tumors with 11p15 LOH are disomic by copy number analysis, consistent with UPD. Lastly, the majority of S5 tumors show biallelic methylation of ICR1 resulting from either 11p15 LOH (37.2%) or LOI (43.8%). This does not result in a significant differential expression of *IGF2* because *IGF2** is highly expressed in virtually all tumors outside S1.

Because 11p15 LOH is usually copy neutral in WT due to UPD, the gene expression consequences of 11p15 LOH and 11p15 LOI are presumed to be largely the same. This is supported by a comparison of S5 tumors with 11p15 LOH versus those with LOI. Significant differential expression of only a few genes was identified, one of which is *CDKN1C* (fold change [FC] = 0.5, $P = 2 \times 10^{-6}$), whose expression is controlled by ICR2 (with biallelic methylation in LOH). Other genes controlled by ICR2 (*PHLDA2*, *SLC22A1L*, *KCNQ1DN*, *KCNQ1*, and *MTR1*) were not significantly differentially expressed, and no other genes of biologic relevance were identified.

1p and 16q LOH

The considerable majority of tumors showing LOH at these two regions reside in S5. S5 tumors did not cluster according to 1p and 16q LOH status.

Validation of Subsets in an Independent Tumor Set

Gene expression and 11p15 methylation analysis were performed on an independent set of 100 tumors from the second group of 300 samples. Hierarchical analysis using the top genes provided in Table W1 demonstrates three subsets of tumors, classified as SA, SB, and SC (Figure 3B). The clinical, pathologic, and 11p15 methylation features of these three subsets, and of the remaining 80 tumors (classified as SD), are provided in Table 2. Subsets A, B, C, and D show the same gene expression and 11p15 methylation patterns, and the same clinicopathologic features as S1, S2, S3, and S5, respectively. An exception is the difference in median age noted between S2 and S3, which was not observed in SB and SC. This is likely a reflection of the broad age ranges seen in S2 and S3 combined with the smaller number of tumors in the validation set (Table 1). Because the top gene lists for S2 and S4 in Table W1 are very similar, it is possible that the above analysis using these lists would not adequately differentiated S2 and S4. To better evaluate for the presence of S4-like tumors within the validation set, we clustered the initial 224 samples combined with the 100 validation samples using the 54 genes associated with *WT1* expression (the

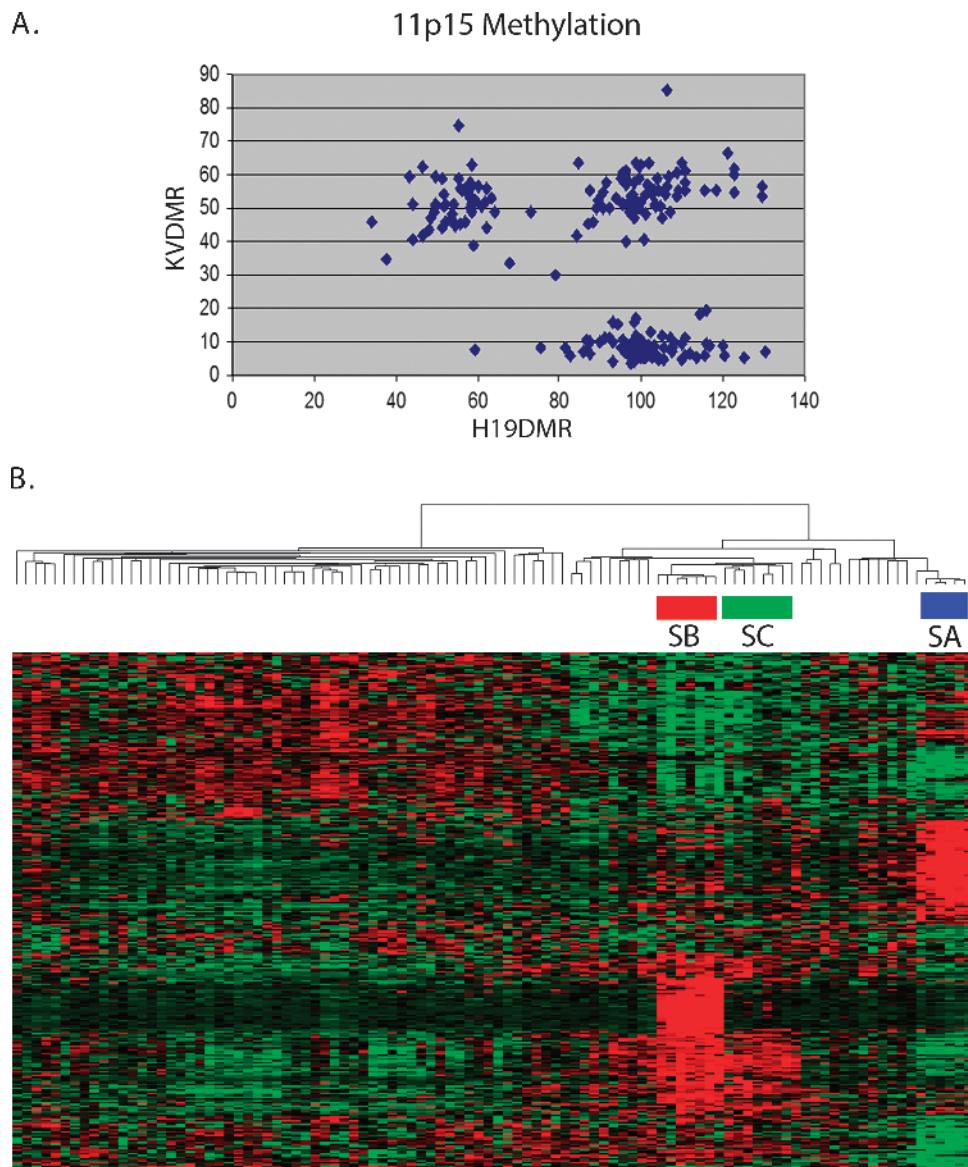


Figure 3. 11p15 methylation and subset validation. (A) ICR1 and ICR2 methylation: Three patterns of methylation were identified: 11p15 LOH (80%-100% methylation of ICR1 and 0%-20% methylation of ICR2), 11p15 LOI (80%-100% methylation of ICR1 and 30%-70% methylation of ICR2), and 11p15 ROI (30%-70% methylation of both ICR1 and ICR2). Tumors with values outside these ranges were not classified. (B) Validation with an independent set of 100 FHWT: Hierarchical analysis was performed using the top genes from Table W1. This demonstrates three subsets of FHWT with the same clinical and pathologic features of S1, S2, and S3 of the training set.

analysis that detected S4 tumors originally). This resulted in clustering of all SA, SB, and SC tumors with S1, S2, and S3 tumors, respectively (data not shown). None of the validation tumors clustered with S4.

Discussion

Wilms tumors (WT), like most pediatric embryonal neoplasms, have been postulated to arise from cells undergoing differentiation during organogenesis. WT commonly display a triphasic histology remarkably similar to cells at different stages of differentiation of the normal MM: blastemal cells similar to early-stage MM, stromal cells, and epithelialized cells arranged in duct-like structures. This histology and the known clonality of these tumors imply that a cell from which a tumor arises is not only very plastic in its differentiation capacity but that progeny cells can exhibit very different cell fates. The current study has been able to delineate, based on gene expression analysis, distinct

subsets of tumors with different molecular, clinical, and pathologic features. These subsets exhibit different mutation patterns as well as different gene expression profiles characteristic of different stages of differentiation. This suggests that tumor subset identity is a result of the interplay of the type of initiating event, the stage of differentiation of the tumor cell of origin, and also its cellular context.

The genitourinary system develops from the intermediate mesoderm, from which the gonads, the mesonephros, and the metanephros (including both the ureteric bud and the primitive MM) develop. Reciprocal interactions between the ureteric bud and the MM are responsible for induction within the MM, which occurs through canonical Wnt signaling. After initiation of induction, mesenchymal-to-epithelial transition begins, enabling the formation of the nephron [44]. Epithelial segmentation, mediated largely by Notch signaling, leads to the development of the glomerular pole proximally and the tubular pole

distally. As has been reported previously for WT, most WTs in our study (those of S5) express genes that characterize the MM [32,36,37]. Because the current study analyzes a very large panel of relapse enriched but otherwise unselected WTs for which substantial clinical and molecular information is available, we were able to identify additional subsets of WT with different clinical and pathologic features that show evidence of disruption at different times during normal development. S1 tumors show gene expression patterns similar to the late postinduction epithelial phase of renal development. The gene expression patterns of S2 and S4 show similarities to the intermediate mesoderm or very early MM. S3 and S5, while displaying unique gene expression patterns, do share a profile similar to the MM. To arrive at the ontogeny of WT, these developmental profiles need to be analyzed in conjunction with the genetic events known to be pathogenetically important in WT.

WT1 in Renal Development and in WT

As demonstrated by the *Wt1*^{-/-} mouse, *WT1* expression is low but essential for cell survival within the intermediate mesoderm [2,45]. This is followed by increased *WT1* expression in the MM. Somatic ablation of *Wt1* in mice soon after nephrogenesis commences results in a block in the formation of epithelial structures [25], consistent with its critical role in the mesenchymal-to-epithelial transition [46]. The expression of *WT1* within the different tumor subsets reflects the changing role of *WT1* during development as well as the phenotypic effects after its loss, often through mutation. *WT1* is highly expressed in S5, in keeping with its expression within the MM, where *WT1* is normally highly expressed. *WT1* is also highly expressed in S1, in keeping with high *WT1* expression normally seen within early epithelial elements. S2 to S4 tumors show low *WT1* expression (often because of mutation), which is frequently associated with stromal elements. This is consistent with experimental data demonstrating that loss of *WT1* disrupts the normal differentiation of epithelial nephronic elements but has no salient effect on stromal development [25].

Canonical Wnt Activation in Renal Development and in WT

Within the MM, canonical Wnt activation is required for the mesenchymal to epithelial transition to occur; however, subsequent Wnt down-regulation is also required to allow the fully epithelial state of the renal vesicle to develop [47–49]. S1 to S5 tumors show variation in canonical Wnt activation concordant with the timing of their developmental arrest or secondary to genetic changes. No evidence of Wnt activation using these analytic tools is identified in S5, in keeping with their origin around the time of induction. Similarly, S1 tumors show evidence of the canonical Wnt down-regulation required for full epithelial differentiation. Strong constitutive canonical Wnt activation is identified in S2 tumors, with only weak Wnt activation in S3 and S4 tumors, correlating with the incidence of *CTNNB1* exon 3 mutation in the subsets (55%, 33%, and 20%, respectively). Canonical Wnt signaling is required for mesenchymal stem cell self-renewal and several of the Wnt pathway genes overexpressed in S2 tumors are mesenchymal stem cell markers, including *MET* and *LGR5* [50,51]. Therefore, for cells of the intermediate mesoderm or early MM that have lost *WT1* (and are therefore unable to complete epithelial differentiation as outlined above), constitutive Wnt activation would be predicted to result in the continued blastemal and mesenchymal cell proliferation and development of S2 tumors. *WT1* loss somewhat later in development (after the normal reduction of canonical Wnt activation) may be less sensitive to *CTNNB1* mutation and require a different

secondary genetic event to achieve abnormal proliferation sufficient to result in tumor development, as discussed further below. It is important to note that Wnt signaling is a remarkably complex and cell type- and stage-dependent. Therefore, the above studies do not fully evaluate canonical or noncanonical Wnt signaling in any of the subsets described.

11p15 Alterations in Renal Development and WT

IGF2 expression is high in the undifferentiated MM. After induction, *IGF2* is dramatically reduced in the epithelial cells but remains high in the stromal cells [52]. Therefore, the patterns of expression of *IGF2* and *WT1* during renal development are complementary, and differences in S1 to 5 tumors would therefore be expected. S1 shows retention of the normal methylation pattern at 11p15 and lower *IGF2* expression, in keeping with postinduction epithelium. In contrast, S5 tumors show high *IGF2* expression, consistent with the very high prevalence of biallelic methylation of ICR1 (81%). S2 to S4 tumors require more careful consideration: 49% of S2 to S4 tumors show evidence of copy neutral 11p15 LOH, known as UPD. UPD occurs when one chromosome (or part thereof) is lost and the second copy is duplicated, usually due to chromosome loss and reduplication or mitotic recombination. In cells that sustain a *WT1* mutation in one allele, UPD for chromosome 11 represents a common mechanism for losing the remaining normal *WT1* allele. UPD (LOH) at 11p15 will also occur in this situation as an epiphenomenon. Because WTs displaying LOH invariably lose the maternally derived chromosome/chromosome segment [53], *WT1* mutant tumors often therefore carry two copies of all paternally imprinted genes. In contrast, the mechanisms that cause 11p15 LOI are confined to 11p15 and can be independent of 11p13 and the *WT1* gene. S2 to S4 tumors show an increasing prevalence of 11p15 LOI (4.5%, 19%, and 30%, respectively). LOI was identified in 6 of 16 S2 to S4 tumors lacking *CTNNB1* mutations, but in none of the 21 S2 to S4 tumors with *CTNNB1* exon 3 mutations. These observations suggest that 11p15-biallelic methylation (either LOH or LOI) may be a second genetic event in the pathogenesis of at least some S2 to S4 tumors and may functionally substitute for *CTNNB1* mutations with regard to malignant development. This provides support to a previous study that carefully detailed the 11p15 copy number and methylation analysis, *WT1* mutation parameters, and the *CTNNB1* status of 36 WT with *WT1* alterations provided by Haruta et al. [54]. Interestingly, a mouse model was recently reported that required the introduction of both *WT1* mutation and *IGF2* overexpression to produce murine WT [25].

A Revised Model for WT Ontogeny

The most prevalent model for WT development is bifurcated. In one arm, biallelic *WT1* mutation results in the development of an ILNR followed by additional genetic changes (e.g., Wnt-activating mutations), resulting in WT development. In the second arm, genetic or epigenetic changes result in biallelic 11p15 ICR1 methylation and development of a PLNR, followed perhaps by additional genetic changes leading to WT development. The data presented in our study confirms these two groups but supports additional heterogeneity that requires revision of this model, as described below (Figure 4).

Subset 1 tumors are characterized by 1) enrichment of genes expressed in the renal epithelium after induction, 2) high *WT1* expression, 3) decreased *IGF2* expression, and 4) decreased expression of Wnt targets. They lack *WT1*, *CTNNB1*, and *WTX* mutations, retain heterozygosity for 1p, and 16q, and retain the normal imprinting pattern at 11p15. We propose that S1 tumors arise in the late

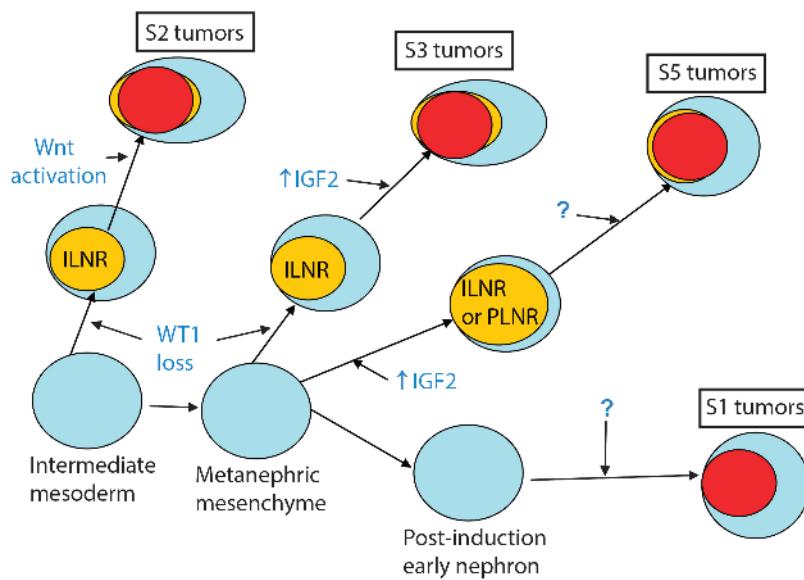


Figure 4. Revised model for WT ontogeny. See text for description.

postinduction MM within a cell characterized by high expression of *WT1*, low expression of *IGF2*, and repression of canonical Wnt signaling. S1 tumors do not arise within nephrogenic rests, and the responsible genetic event(s) has not been identified.

Subset 2 tumors show 1) enrichment of genes expressed in the intermediate mesoderm and early MM, 2) loss of *WT1* expression often due to *WT1* mutation, 3) canonical Wnt activation accompanied by a high *CTNNB1*-exon 3 mutation frequency, and 4) divergent mesenchymal differentiation, previously attributed to *WT1* loss and/or canonical Wnt activation [41,55–58]. In addition, 25% of tumors showed *WTX* mutation, which has been associated with expansion of early mesenchymal precursors within the developing kidney [59]. We propose that S2 tumors arise within cells of the intermediate mesoderm or early MM after inactivation of one *WT1* allele followed by loss of the second *WT1* allele, commonly through chromosomal mechanisms resulting in UPD for 11p13. (Incidental UPD for 11p15 also occurs; however, during this developmental window, biallelic expression of *IGF2* would not be expected to be as significant because the *IGF2* levels are already quite high.) Loss of *WT1* perturbs normal nephron development with preferential expansion of mesenchymal elements (which do not require *WT1*), resulting in an ILNR. Within the ILNR, an additional genetic event (such as *CTNNB1* or *WTX* mutation) results in Wnt activation accompanied by continued proliferation of primitive mesenchymal cells and increased diversion into alternate pathways, including muscle differentiation. The mixed histology observed in S2 tumors supports an origin before the delineation of nephron- or stroma-fated compartments occurs.

Subset 3 tumors show similarities with S2 tumors owing to their common primary pathogenetic feature (low *WT1* expression). However, they differ by their decreased evidence of canonical Wnt activation and decreased divergent mesenchymal differentiation. In addition, S3 tumors lack enrichment of genes expressed in the intermediate mesoderm and instead show the same expression of MM genes seen in S5 tumors. We therefore propose that S3 tumors show a similar sequence of genetic events to S2, although these events occur in cells of the MM later in development, at a time when *WT1* expression is increasing and *IGF2* expression is decreasing. In addition, the different timing is asso-

ciated with a different sensitivity to (or requirement for) Wnt-activating mutations and therefore to differences in divergent mesenchymal differentiation. Our study provides data supporting the hypothesis previously reported that 11p15 ICR1 biallelic methylation may represent an alternative second event [54].

S4 tumors have a gene expression pattern similar to S2, although they have a lower incidence of *WT1*, *CTNNB1*, and *WTX* mutations (and corresponding expression patterns) and a high incidence of biallelic methylation of ICR1 (80%) similar to S5. Nonetheless, they are defined by their *WT1*-associated gene expression pattern. These tumors may therefore have alternative genetic/epigenetic abnormalities up- or downstream to *WT1*. The small number of tumors within this subset and our inability to validate this subset suggests caution is needed with regard to S4. The intriguing clinical and genetic differences, particularly the high relapse rate, support the continued retention of this subset pending further knowledge.

S5 tumors display the gene expression pattern of the MM and a high frequency of biallelic methylation of ICR1 at 11p15. This supports assertions that approximately 70% of WT arise, in part, due to abnormal expression of normally imprinted 11p15 genes [21,23]. It also corroborates an association between an older age at presentation, histologic pattern, and 11p15 LOI in WT [60]. S5 is notable for being the only subset in which PLNRs were observed, although ILNRs were also seen and some tumors contained both. Patients with BWS are known to develop both ILNRs and PLNRs [1]. This suggests that factors such as the cell type or the timing of the initial genetic event may determine the type of nephrogenic rest and the heterogeneous WT histology that results. We propose that S5 tumors arise within cells of the MM at a time when *IGF2* expression is diminishing. Biallelic methylation of ICR1 during this developmental window results in increased *IGF2* expression and the development of a nephrogenic rest. If biallelic methylation of ICR1 occurs before induction, the persistent elevation of *IGF2* causes preferential mesenchymal proliferation and prevents nephron development, resulting in an ILNR. If manifested after induction within early nephronic cells, persistent *IGF2* elevation prevents terminal epithelial differentiation, resulting in the development of a PLNR. This is supported by the

presence of high *WT1* and *IGF2* expression in primitive epithelial structures of WT, whereas *IGF2* expression is low in normal terminally differentiated epithelial renal elements [52]. Within the resulting nephrogenic rest (be it ILNR or PLNR), a second genetic event likely occurs resulting in tumor development. The absence of tumor development in a mutant mouse strain with biallelic *Igf2* expression, but no other genetic alteration, would suggest that a second event is required [25], although the nature of this putative second event in S5 is still unknown.

An important question that remains is whether S5 is a unique entity or simply a conglomeration of tumors that failed to cluster within S1, S2, S3, or S4. Our study accumulates evidence supporting the unity of S5, as outlined below:

- More than 80% of S5 tumors show somatic biallelic methylation of ICR1, a constitutional abnormality seen in patients with BWS, which is associated with WT development.
- The clinical and pathologic features of tumors arising in patients with BWS (including precursor lesions) are the same as those identified in S5 tumors that are not associated with BWS.
- We could not identify other clusters within S5 based on gene expression, even when S1 to S4 were removed.
- S5 tumors did not show subclusters based on any clinical or pathologic parameter analyzed in this study.
- We have considerable experience observing the variation of gene expression within pathogenetically homogeneous populations of tumors, contrasted with the differences in expression between different tumor types, even when they are histologically similar [31,32,61–63]. The S5 tumors show the range of variation that is expected within a group of tumors that share an underlying pathogenesis.

In summary, while we cannot exclude the presence of biologically distinctive subsets within the larger group of S5 tumors, in our opinion, the prevailing evidence supports the hypothesis that S5 tumors represent a single entity united by their underlying pathogenesis.

Significance

The categorization by gene expression analysis of prospectively identified WTs resulted in the delineation of biologically unique subsets with distinctive mutational spectra and clinical outcomes. Not only does this provide insight into the pathogenesis of these subsets and an explanation for their heterogeneity, but more importantly, defining subsets driven by different genetic events may allow for both subset-specific and targeted therapeutic strategies. Our data suggest that therapies targeting the IGF receptor may be broadly applicable to WT outside S1, whereas the use of therapies involving the Wnt, TGFB, and NFAT pathways, when available, may have activity restricted to particular subsets. The future therapy of infants with low-stage disease will be particularly affected by subset-specific strategies for reducing chemotherapy. Currently, patients younger than 24 months with stage I FHWT weighing less than 550 g are defined as very low risk WT (VLRWT) and are treated with surgery alone. S1 seems to be responsible for approximately 30% of VLRWT who did not receive adjuvant chemotherapy [34,60,64]. In the current study, S1 tumors also occurred in patients older than 24 months, with nephrectomy weights more than 550 g, yet still retained an excellent survival. This may allow for the removal of the arbitrary age and tumor weight restrictions for S1 tumors, thereby expanding the number of patients able to be treated with nephrectomy alone. S2 tumors (likewise common in infants) had an excellent outcome in the current study (in which all

patients received adjuvant chemotherapy). In our previous study of VLRWT, when such patients did not receive chemotherapy, there was an increased risk of relapse [34,65], suggesting they may benefit from chemotherapy. Therefore, these studies provide the opportunity to define groups of VLRWT with different relapse risks using appropriate biologic markers that are now being validated using samples from the current protocol.

More globally, our findings illustrate that while pediatric embryonal tumors have a restricted number of genetic events, they cannot be characterized solely by activation of a single gene or pathway. Rather, the clinical and biologic phenotype may be determined in part by the developmental context in which genetic lesions are introduced. In particular, our study highlights considerable complexity with regard to the role of *WT1*. The impact of *WT1* loss seems to change depending on the developmental timing and/or cell of origin of its occurrence, as is seen in S2 and S3 tumors. We provide further evidence suggesting that activation of key signaling pathways (NFAT, TGFB, Ras) is restricted to those subsets arising through loss of *WT1* expression quite early in development (S2 and S4). Lastly, our study points toward mechanisms of *WT1* loss of expression other than mutation that need to be identified. All these observations are of broad interest, as the critical role of *WT1* is increasingly demonstrated to play an important and broad role in development and disease that extends well beyond the kidney [46].

Acknowledgments

The authors are grateful for the technical expertise of Patricia Beezhold and Donna Kersey.

References

- [1] Beckwith JB, Kiviat NB, and Bonadio JF (1990). Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor. *Pediatr Pathol* **10**, 1–36.
- [2] Kreidberg JA, Sariola H, Loring JM, Maeda M, Pelletier J, Housman D, and Jaenisch R (1993). *WT1* is required for early kidney development. *Cell* **74**, 679–691.
- [3] Scharnhorst V, van der Eb AJ, and Jochemsen AG (2001). *WT1* proteins: functions in growth and differentiation. *Gene* **273**, 141–161.
- [4] Pelletier J, Bruening W, Kashtan CE, Mauer SM, Manivel JC, Striegel JE, Houghton DC, Junien C, Habib R, and Fousher L (1991). Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell* **67**, 437–447.
- [5] Call KM, Glaser T, Ito CY, Buckler AJ, Pelletier J, Haber DA, Rose EA, Kral A, Yeger H, and Lewis WH (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* **60**, 509–520.
- [6] Gessler M, Poustka A, Cavenee W, Neve RL, Orkin SH, and Bruns GA (1990). Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. *Nature* **343**, 774–778.
- [7] Varanasi R, Bardeesy N, Ghahremani M, Petruzzoli MJ, Nowak N, Adam MA, Grundy P, Shows TB, and Pelletier J (1994). Fine structure analysis of the *WT1* gene in sporadic Wilms tumors. *Proc Natl Acad Sci USA* **91**, 3554–3558.
- [8] Gessler M, Konig A, Arden K, Grundy P, Orkin S, Sallan S, Peters C, Ruyle S, Mandell J, Li F, et al. (1994). Infrequent mutation of the *WT1* gene in 77 Wilms' tumors. *Hum Mutat* **3**, 212–222.
- [9] Huff V (1998). Wilms tumor genetics. *Am J Med Genet* **79**, 260–267.
- [10] Maiti S, Alam R, Amos CI, and Huff V (2000). Frequent association of β -catenin and *WT1* mutations in Wilms tumors. *Cancer Res* **60**, 6288–6292.
- [11] Koesters R, Ridder R, Kopp-Schneider A, Betts D, Adams V, Niggli F, Briner J, and von Knebel DM (1999). Mutational activation of the β -catenin proto-oncogene is a common event in the development of Wilms' tumors. *Cancer Res* **59**, 3880–3882.
- [12] Major MB, Camp ND, Berndt JD, Yi X, Goldenberg SJ, Hubbert C, Biechele TL, Gingras AC, Zheng N, MacCoss MJ, et al. (2007). Wilms tumor suppressor WTX negatively regulates WNT/ β -catenin signaling. *Science* **316**, 1043–1046.

- [13] Rivera MN, Kim WJ, Wells J, Driscoll DR, Brannigan BW, Han M, Kim JC, Feinberg AP, Gerald WL, Vargas SO, et al. (2007). An X chromosome gene, *WTX*, is commonly inactivated in Wilms tumor. *Science* **315**, 642–645.
- [14] Ruteshouser EC, Robinson SM, and Huff V (2008). Wilms tumor genetics: mutations in *WT1*, *WTX*, and *CTNNB1* account for only about one-third of tumors. *Genes Chromosomes Cancer* **47**, 461–470.
- [15] Perotti D, Gamba B, Sardella M, Spreafico F, Terenziani M, Collini P, Pession A, Nantron M, Fossati-Bellani F, and Radice P (2008). Functional inactivation of the *WTX* gene is not a frequent event in Wilms' tumors. *Oncogene* **27**, 4625–4632.
- [16] Fukuzawa R, Heathcott RW, More HE, and Reeve AE (2007). Sequential *WT1* and *CTNNB1* mutations and alterations of β -catenin localisation in intralobar nephrogenic rests and associated Wilms tumours: two case studies. *J Clin Pathol* **60**, 1013–1016.
- [17] Uschkereit C, Perez N, de Torres C, Kuff M, Mora J, and Royer-Pokora B (2007). Different *CTNNB1* mutations as molecular genetic proof for the independent origin of four Wilms tumours in a patient with a novel germ line *WT1* mutation. *J Med Genet* **44**, 393–396.
- [18] Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, and Feinberg AP (1993). Relaxation of imprinted genes in human cancer. *Nature* **362**, 747–749.
- [19] Ping AJ, Reeve AE, Law DJ, Young MR, Boehnke M, and Feinberg AP (1989). Genetic linkage of Beckwith-Wiedemann syndrome to 11p15. *Am J Hum Genet* **44**, 720–723.
- [20] Weksberg R, Shen DR, Fei YL, Song QL, and Squire J (1993). Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. *Nat Genet* **5**, 143–150.
- [21] Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, Smith PJ, and Reeve AE (1993). Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. *Nature* **362**, 749–751.
- [22] Ohlsson R, Nyström A, Pfeifer-Ohlsson S, Tohonen V, Hedborg F, Schofield P, Flam F, and Ekstrom TJ (1993). IGF2 is parentally imprinted during human embryogenesis and in the Beckwith-Wiedemann syndrome. *Nat Genet* **4**, 94–97.
- [23] Steenman MJ, Rainier S, Dobry CJ, Grundy P, Horon IL, and Feinberg AP (1994). Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nat Genet* **7**, 433–439. Erratum in *Nat Genet* 1994;8:203.
- [24] Chao LY, Huff V, Tomlinson G, Riccardi VM, Strong LC, and Saunders GF (1993). Genetic mosaicism in normal tissues of Wilms' tumour patients. *Nat Genet* **3**, 127–131.
- [25] Hu Q, Gao F, Tian W, Ruteshouser EC, Wang Y, Lazar J, Strong L, Behringer RR, and Huff V (2011). Wt1 ablation and Igf2-up-regulation in mice result in Wilms tumors with elevated ERK1/2 phosphorylation. *J Clin Invest* **121**, 174–183.
- [26] Rahman N, Abidi F, Ford D, Arbour L, Rapley E, Tonin P, Barton D, Batcup G, Berry J, Cotter F, et al. (1998). Confirmation of FWT1 as a Wilms' tumour susceptibility gene and phenotypic characteristics of Wilms' tumour attributable to FWT1. *Hum Genet* **103**, 547–556.
- [27] MacDonald JM, Douglass EC, Fisher R, Geiser CF, Krill CE, Strong LC, Virshup D, and Huff V (1998). Linkage of familial Wilms' tumor predisposition to chromosome 19 and a two-locus model for the etiology of familial tumors. *Cancer Res* **58**, 1387–1390.
- [28] Grundy PE, Breslow NE, Li S, Perlman E, Beckwith JB, Ritchey ML, Shamberger RC, Haase GM, D'Angio GJ, Donaldson M, et al. (2005). Loss of heterozygosity for chromosomes 1p and 16q is an adverse prognostic factor in favorable-histology Wilms tumor: a report from the National Wilms Tumor Study Group. *J Clin Oncol* **23**, 7312–7321.
- [29] Brunskill EW, Aronow BJ, Georgas K, Rumballe B, Valerius MT, Aronow J, Kaimal V, Jegga AG, Yu J, Grimmond S, et al. (2008). Atlas of gene expression in the developing kidney at microanatomic resolution. *Dev Cell* **15**, 781–791.
- [30] Lin SA, Kolle G, Grimmond S, Zhou Q, Doust E, Little MH, Aronow BJ, Ricardo SD, Pera MF, Bertram JF, et al. (2010). Subfractionation of differentiating human embryonic stem cell populations allows the isolation of a mesodermal population enriched for intermediate mesoderm and putative renal progenitors. *Stem Cells Dev* **19**, 1637–1648.
- [31] Huang CC, Gadd S, Breslow NB, Cutcliffe C, Sredni S, Helenowski IB, Dome JS, Grundy PE, Green DM, Fritsch MK, et al. (2009). Predicting relapse in favorable histology Wilms tumor using gene expression analysis. A report from the renal tumor committee of the Children's Oncology Group. *Clin Cancer Res* **15**, 1770–1778.
- [32] Huang CC, Cutcliffe C, Coffin C, Sorensen PH, Beckwith JB, and Perlman EJ (2006). Classification of malignant pediatric renal tumors by gene expression. *Pediatr Blood Cancer* **46**, 728–738.
- [33] Eisen MB, Spellman PT, Brown PO, and Botstein D (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* **95**, 14863–14868.
- [34] Perlman EJ, Grundy PE, Anderson JR, Jennings LJ, Green DM, Dome JS, Shamberger RC, Ruteshouser EC, and Huff V (2011). *WT1* mutation and 11P15 loss of heterozygosity predict relapse in very low-risk Wilms tumors treated with surgery alone: a Children's Oncology Group study. *J Clin Oncol* **29**, 698–703.
- [35] Huff V, Jaffe N, Saunders GF, Strong LC, Villalba F, and Ruteshouser EC (1995). *WT1* exon 1 deletion/insertion mutations in Wilms tumor patients, associated with di- and trinucleotide repeats and deletion hotspot consensus sequences. *Am J Hum Genet* **56**, 84–90.
- [36] Li CM, Guo M, Borczuk A, Powell CA, Wei M, Thaker HM, Friedman R, Klein U, and Tycko B (2002). Gene expression in Wilms' tumor mimics the earliest committed stage in the metanephric mesenchymal-epithelial transition. *Am J Pathol* **160**, 2181–2190.
- [37] Dekel B, Metsuyanim S, Schmidt-Ott KM, Fridman E, Jacob-Hirsch J, Simon A, Pinthus J, Mor Y, Barasch J, Amarliglio N, et al. (2006). Multiple imprinted and stemness genes provide a link between normal and tumor progenitor cells of the developing human kidney. *Cancer Res* **66**, 6040–6049.
- [38] Azzarello JT, Lin HK, Gherezghiher A, Zakharov V, Yu Z, Kropp BP, Culkin DJ, Penning TM, and Fung KM (2009). Expression of *AKR1C3* in renal cell carcinoma, papillary urothelial carcinoma, and Wilms' tumor. *Int J Clin Exp Pathol* **3**, 147–155.
- [39] Dey BR, Sukhatme VP, Roberts AB, Sporn MB, Rauscher FJ III, and Kim SJ (1994). Repression of the transforming growth factor- β 1 gene by the Wilms' tumor suppressor *WT1* gene product. *Mol Endocrinol* **8**, 595–602.
- [40] Singh G, Singh SK, Konig A, Reutlinger K, Nye MD, Adhikary T, Eilers M, Gress TM, Fernandez-Zapico ME, and Ellenrieder V (2010). Sequential activation of NFAT and c-Myc transcription factors mediates the TGF- β switch from a suppressor to a promoter of cancer cell proliferation. *J Biol Chem* **285**, 27241–27250.
- [41] Fukuzawa R, Anaka MR, Weeks RJ, Morison IM, and Reeve AE (2009). Canonical WNT signalling determines lineage specificity in Wilms tumour. *Oncogene* **28**, 1063–1075.
- [42] Li CM, Kim CE, Margolin AA, Guo M, Zhu J, Mason JM, Hensle TW, Murty VV, Grundy PE, Fearon ER, et al. (2004). *CTNNB1* mutations and overexpression of *Wnt/ β -catenin* target genes in *WT1*-mutant Wilms' tumors. *Am J Pathol* **165**, 1943–1953.
- [43] Corbin M, de Reynies A, Rickman DS, Berrebi D, Boccon-Gibod L, Cohen-Gogo S, Fabre M, Jaubert F, Faussillon M, Yilmaz F, et al. (2009). WNT/ β -catenin pathway activation in Wilms tumors: a unifying mechanism with multiple entries? *Genes Chromosomes Cancer* **48**, 816–827.
- [44] Schwab K, Patterson LT, Aronow BJ, Luckas R, Liang HC, and Potter SS (2003). A catalogue of gene expression in the developing kidney. *Kidney Int* **64**, 1588–1604.
- [45] Moore AW, McInnes L, Kreidberg J, Hastie ND, and Schedl A (1999). YAC complementation shows a requirement for Wt1 in the development of epicardium, adrenal gland and throughout nephrogenesis. *Development* **126**, 1845–1857.
- [46] Miller-Hedges E and Hohenstein P (2012). *WT1* in disease: shifting the epithelial-mesenchymal balance. *J Pathol* **226**, 229–240.
- [47] Park JS, Valerius MT, and McMahon AP (2007). *Wnt/ β -catenin* signaling regulates nephron induction during mouse kidney development. *Development* **134**, 2533–2539.
- [48] Iglesias DM, Hueber PA, Chu L, Campbell R, Patenaude AM, Dziarmaga AJ, Quinlan J, Mohamed O, Dufort D, and Goodyer PR (2007). Canonical WNT signaling during kidney development. *Am J Physiol Renal Physiol* **293**, F494–F500.
- [49] Schmidt-Ott KM, Masckauchan TN, Chen X, Hirsh BJ, Sarkar A, Yang J, Paragas N, Wallace VA, Dufort D, Pavlidis P, et al. (2007). β -Catenin/TCF/Lef controls a differentiation-associated transcriptional program in renal epithelial progenitors. *Development* **134**, 3177–3190.
- [50] Etheridge SL, Spencer GJ, Heath DJ, and Genever PG (2004). Expression profiling and functional analysis of wnt signaling mechanisms in mesenchymal stem cells. *Stem Cells* **22**, 849–860.
- [51] Lindsley RC, Gill JG, Kyba M, Murphy TL, and Murphy KM (2006). Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development* **133**, 3787–3796.
- [52] Yun K, Molenaar AJ, Fiedler AM, Mark AJ, Eccles MR, Becroft DM, and Reeve AE (1993). Insulin-like growth factor II messenger ribonucleic acid

- expression in Wilms tumor, nephrogenic rest, and kidney. *Lab Invest* **69**, 603–615.
- [53] Schroeder WT, Chao LY, Dao DD, Strong LC, Pathak S, Riccardi V, Lewis WH, and Saunders GF (1987). Nonrandom loss of maternal chromosome 11 alleles in Wilms tumors. *Am J Hum Genet* **40**, 413–420.
- [54] Haruta M, Arai Y, Sugawara W, Watanabe N, Honda S, Ohshima J, Soejima H, Nakadate H, Okita H, Hata J, et al. (2008). Duplication of paternal IGF2 or loss of maternal IGF2 imprinting occurs in half of Wilms tumors with various structural WT1 abnormalities. *Genes Chromosomes Cancer* **47**, 712–727.
- [55] Schumacher V, Schneider S, Figge A, Wildhardt G, Harms D, Schmidt D, Weirich A, Ludwig R, and Royer-Pokora B (1997). Correlation of germ-line mutations and two-hit inactivation of the *WT1* gene with Wilms tumors of stromal-predominant histology. *Proc Natl Acad Sci USA* **94**, 3972–3977.
- [56] Miyagawa K, Kent J, Moore A, Charlieu JP, Little MH, Williamson KA, Kelsey A, Brown KW, Hassam S, Briner J, et al. (1998). Loss of *WT1* function leads to ectopic myogenesis in Wilms' tumour [letter]. *Nat Genet* **18**, 15–17.
- [57] Fukuzawa R, Breslow NE, Morison IM, Dwyer P, Kusafuka T, Kobayashi Y, Becroft DM, Beckwith JB, Perlman EJ, and Reeve AE (2004). Epigenetic differences between Wilms' tumours in white and East-Asian children. *Lancet* **363**, 446–451.
- [58] Fukuzawa R, Heathcott RW, Sano M, Morison IM, Yun K, and Reeve AE (2004). Myogenesis in Wilms' tumors is associated with mutations of the *WT1* gene and activation of Bcl-2 and the Wnt signaling pathway. *Pediatr Dev Pathol* **7**, 125–137.
- [59] Moisan A, Rivera MN, Lotinun S, Akhavanfar S, Coffman EJ, Cook EB, Stoykova S, Mukherjee S, Schoonmaker JA, Burger A, et al. (2011). The *WTX* tumor suppressor regulates mesenchymal progenitor cell fate specification. *Dev Cell* **20**, 583–596.
- [60] Ravenel JD, Broman KW, Perlman EJ, Niemitz EL, Jayawardena TM, Bell DW, Haber DA, Uejima H, and Feinberg AP (2001). Loss of imprinting of insulin-like growth factor-II (*IGF2*) gene in distinguishing specific biologic subtypes of Wilms tumor. *J Natl Cancer Inst* **93**, 1698–1703.
- [61] Cutcliff C, Kersey D, Huang CC, Hasan C, Walterhouse D, and Perlman EJ (2005). Clear cell sarcoma of the kidney: up-regulation of neural markers with activation of the sonic hedgehog and Akt pathways. *Clin Cancer Res* **11**, 7986–7994.
- [62] Gadd S, Beezhold P, Jennings L, George D, Leuer K, Huang CC, Huff V, Tognon C, Sorensen PH, Triche T, et al. (in press). Mediators of receptor tyrosine kinase activation in infantile fibrosarcoma: a Children's Oncology Group study. *J Pathol*.
- [63] Gadd S, Sredni ST, Huang CC, and Perlman EJ (2010). Rhabdoid tumor: gene expression clues to pathogenesis and potential therapeutic targets. *Lab Invest* **90**, 724–738.
- [64] Green DM, Breslow NE, Beckwith JB, Ritche ML, Shamberger RC, Haase GM, D'Angio GJ, Perlman E, Donaldson M, Grundy PE, et al. (2001). Treatment with nephrectomy only for small, stage I/favorable histology Wilms' tumor: a report from the National Wilms' Tumor Study Group. *J Clin Oncol* **19**, 3719–3724.
- [65] Sredni S, Gadd S, Huang CC, Breslow N, Grundy P, Green DM, Dome J, Shamberger RC, Beckwith B, and Perlman EJ (2009). Subsets of very low risk Wilms tumors show distinctive gene expression, histologic, and clinical features. *Clin Cancer Res* **15**, 6800–6809.
- [66] Zirn B, Samans B, Wittmann S, Pietsch T, Leuschner I, Graf N, and Gessler M (2006). Target genes of the WNT/β-catenin pathway in Wilms tumors. *Genes Chromosomes Cancer* **45**, 565–574.

Table W1. Top 100 Genes Differentially Expressed in Each Subset by GSEA.

S1 vs S5			S2 vs S5			S3 vs S5			S4 vs S5		
Gene	P	FC	Gene	P	FC	Gene	P	FC	Gene	P	FC
<i>MGAM</i>	2.7e - 10	12.7	<i>MYH3</i>	3.4e - 20	43.7	<i>GAS2</i> * [†]	2.3e - 13	7.2	<i>MYH3</i>	2.3e - 06	11.4
<i>KCNJ16</i>	2.5e - 07	8.9	<i>KBTBD10</i>	4.2e - 16	34.0	<i>MEOX2</i> [†]	1.3e - 05	4.5	<i>KBTBD10</i>	7.4e - 05	10.4
<i>HLA-DMA</i>	7.2e - 09	20.2	<i>TNNC1</i>	9.3e - 17	39.7	<i>SIM2</i> * [†]	2.7e - 12	4.5	<i>TNNC1</i>	2.0e - 05	8.8
<i>GUCY1B3</i>	7.9e - 12	10.9	<i>TTN</i>	2.8e - 17	25.1	<i>NAV3</i> * [†]	1.6e - 13	4.3	<i>TTN</i>	2.5e - 05	10.0
<i>AKR1C3</i>	6.4e - 08	9.4	<i>WIF1</i>	1.7e - 11	19.2	<i>EVI1</i> [†]	1.8e - 05	3.5	<i>ACTA1</i>	1.9e - 04	9.8
<i>IL1RI</i>	8.8e - 09	16.5	<i>ACTA1</i>	2.9e - 14	21.2	<i>NEF3</i> *	3.5e - 05	2.8	<i>TNNC2</i>	3.0e - 05	8.0
<i>AMIGO2</i>	7.5e - 06	8.9	<i>NEB</i>	3.8e - 16	16.1	<i>CNR1</i> [†]	5.7e - 07	4.6	<i>MYL1</i>	2.5e - 05	8.0
<i>TMEM23</i>	1.6e - 08	6.1	<i>TNNC2</i>	4.2e - 12	23.3	<i>WIF1</i> * [†]	5.3e - 04	3.1	<i>NEB</i>	1.9e - 05	4.9
<i>FLJ20273</i>	2.7e - 11	10.3	<i>MYOZ2</i>	3.6e - 10	13.5	<i>GADI</i> * [†]	6.0e - 05	4.9	<i>ACTC1</i>	1.2e - 06	8.6
<i>FBLN2</i>	2.9e - 09	13.6	<i>MYOT</i>	3.3e - 11	9.0	<i>LMO2</i> [†]	1.3e - 05	2.8	<i>MEOX2</i>	4.2e - 04	4.0
<i>C11ORF32</i>	5.2e - 07	11.0	<i>MYL1</i>	2.4e - 14	16.4	<i>PLK2</i> [†]	4.3e - 06	2.6	<i>TPM2</i>	6.3e - 06	5.5
<i>SQRDL</i>	1.7e - 10	8.4	<i>OSR2</i>	3.2e - 10	8.3	<i>NPTX2</i> [†]	4.1e - 07	2.7	<i>EVI1</i>	7.5e - 04	4.3
<i>FRY</i>	6.8e - 06	6.5	<i>ACTC1</i>	3.4e - 22	23.6	<i>TGFB2</i> [†]	2.2e - 05	2.7	<i>WIF1</i>	5.9e - 04	3.5
<i>TNFSF10</i>	3.6e - 04	7.0	<i>TPM2</i>	2.5e - 16	13.5	<i>MID1</i> * [†]	5.5e - 09	2.8	<i>GAS2</i>	2.4e - 05	3.5
<i>LGII</i>	1.0e - 03	11.9	<i>DKK2</i>	2.2e - 09	6.1	<i>DKFZP686A01247</i> [†]	2.2e - 07	2.6	<i>MYLPF</i>	9.1e - 06	5.5
<i>HEY2</i>	4.0e - 06	6.1	<i>MYLPF</i>	5.6e - 14	16.8	<i>CXCR7</i> [†]	2.0e - 04	2.1	<i>NAV3</i>	6.0e - 08	3.6
<i>EGFL6</i>	2.8e - 05	6.9	<i>SULT1E1</i>	5.5e - 08	8.6	<i>TSPAN8</i> [†]	8.6e - 04	2.0	<i>COL15A1</i>	3.1e - 03	4.7
<i>ARHGPAP29</i>	1.4e - 06	7.5	<i>TNNI1</i>	3.0e - 12	9.2	<i>RP6-213H19.1</i> [†]	1.2e - 09	2.1	<i>MYOZ2</i>	4.4e - 03	3.8
<i>C20ORF103</i>	5.8e - 07	10.3	<i>OGN</i>	2.7e - 08	6.1	<i>WNT5A</i> * [†]	4.1e - 04	3.0	<i>CHODL</i>	2.5e - 03	4.1
<i>WISP3</i>	1.4e - 06	4.7	<i>CKM</i>	1.3e - 08	8.9	<i>NPY</i>	2.1e - 03	2.9	<i>CFH</i>	3.8e - 04	2.9
<i>BHLHB3</i>	2.5e - 04	4.3	<i>TWIST1</i>	2.5e - 07	11.3	<i>SART2</i> [†]	1.1e - 06	2.5	<i>OSR2</i>	3.6e - 04	2.6
<i>CLU</i>	9.9e - 08	8.7	<i>NEF3</i>	6.0e - 15	4.9	<i>EDNRA</i> * [†]	5.3e - 04	2.2	<i>LUM</i>	2.6e - 03	3.3
<i>GRB14</i>	1.5e - 04	4.5	<i>LUM</i>	2.3e - 20	6.7	<i>BCHE</i> [†]	1.2e - 02	2.6	<i>MYOT</i>	2.2e - 03	2.6
<i>CLEC4M</i>	5.2e - 05	7.8	<i>TNNI2</i>	2.5e - 09	9.6	<i>CXCL6</i> [†]	5.3e - 03	2.0	<i>SRPX</i>	7.7e - 03	5.4
<i>CCND1</i>	2.4e - 13	5.4	<i>PITX2</i>	3.8e - 08	3.9	<i>FBN2</i> [†]	1.1e - 05	2.0	<i>BCHE</i>	1.1e - 02	3.9
<i>CLIC5</i>	1.3e - 05	3.3	<i>BCHE</i>	1.5e - 07	4.6	<i>CYP26A1</i> [†]	3.0e - 05	2.2	<i>MAL</i>	2.4e - 03	2.9
<i>CUGBP2</i>	6.9e - 24	8.0	<i>GAS2</i>	2.6e - 11	5.2	<i>CDH12</i>	1.1e - 02	2.3	<i>SPARCL1</i>	1.1e - 04	3.6
<i>FER1L3</i>	1.8e - 06	5.5	<i>TMEM47</i>	4.2e - 08	5.4	<i>LRRTMD</i> [†]	5.2e - 03	2.0	<i>PLSCR4</i>	8.0e - 03	2.9
<i>CTGF</i>	1.1e - 10	6.1	<i>FABP7</i>	5.8e - 07	11.4	<i>CCND1</i> [†]	8.9e - 06	1.5	<i>DLK1</i>	5.8e - 04	2.3
<i>CRTAC1</i>	2.4e - 06	8.2	<i>IL17B</i>	1.6e - 10	4.8	<i>PROM1</i> [†]	9.7e - 09	1.7	<i>EMX2</i>	3.0e - 03	2.5
<i>QPCT</i>	5.6e - 05	4.1	<i>DPT</i>	1.1e - 08	8.9	<i>GPR177</i> * [†]	3.5e - 07	2.1	<i>SERPINF1</i>	5.5e - 04	3.6
<i>FGF14</i>	7.9e - 07	4.3	<i>NAV3</i>	1.8e - 15	6.8	<i>DNM3</i> [†]	2.7e - 03	2.1	<i>MATN2</i>	7.1e - 03	3.1
<i>OLFML2A</i>	2.9e - 09	7.2	<i>COL15A1</i>	8.2e - 11	5.9	<i>DKK1</i> [†]	1.2e - 02	3.3	<i>OGN</i>	5.7e - 04	1.9
<i>HS3ST2</i>	2.8e - 09	6.2	<i>COL21A1</i>	5.1e - 09	5.3	<i>EYA2</i> [†]	1.7e - 04	1.9	<i>ZFHXB1</i>	3.6e - 03	2.3
<i>PLOD2</i>	4.5e - 14	5.8	<i>MEF2C</i>	8.7e - 13	6.6	<i>LOXL1</i> [†]	4.3e - 09	1.9	<i>TNNI1</i>	9.3e - 04	2.8
<i>PPARGC1A</i>	2.3e - 06	3.8	<i>MYH8</i>	2.0e - 12	7.6	<i>IGFB2BP3</i> [†]	5.2e - 03	1.4	<i>MEF2C</i>	5.1e - 06	3.0
<i>SATB2</i>	3.2e - 05	3.5	<i>DKK1</i>	9.1e - 07	6.6	<i>KIAA0960</i> [†]	3.1e - 03	2.0	<i>NEF3</i>	4.3e - 03	2.3
<i>SLC12A1</i>	3.1e - 03	7.1	<i>PLK2</i>	3.8e - 17	4.8	<i>IGFBP5</i> [†]	1.2e - 06	1.9	<i>CKM</i>	5.6e - 04	2.6
<i>C14ORF105</i>	1.5e - 03	4.2	<i>CHODL</i>	1.2e - 09	5.2	<i>RASL11B</i> * [†]	4.8e - 10	2.1	<i>ALDH1A1</i>	1.3e - 02	2.6
<i>IRX5</i>	5.2e - 05	3.6	<i>ZFHXB1</i>	8.7e - 09	3.3	<i>LRIG1</i> [†]	1.2e - 07	1.8	<i>SIM2</i>	4.0e - 03	3.3
<i>GPX3</i>	6.3e - 08	7.8	<i>MYL4</i>	1.9e - 11	5.2	<i>IRX5</i> [†]	1.0e - 04	1.4	<i>COL21A1</i>	6.1e - 03	2.6
<i>ASTN2</i>	8.8e - 06	4.7	<i>DLK1</i>	4.0e - 15	3.7	<i>MYC</i> [†]	2.5e - 05	1.6	<i>PDGFR4</i>	1.2e - 04	2.8
<i>PDZD2</i>	4.0e - 07	4.0	<i>SERPINF1</i>	1.4e - 17	5.1	<i>EGFL6</i> [†]	6.6e - 04	1.9	<i>LDB2</i>	1.1e - 04	2.5
<i>TRHDE</i>	1.5e - 03	3.8	<i>EGFL6</i>	2.1e - 09	4.3	<i>CHGN</i> [†]	5.0e - 04	1.9	<i>CXCR7</i>	1.3e - 03	2.2
<i>LIFR</i>	2.9e - 05	3.9	<i>MEOX2</i>	7.9e - 07	3.6	<i>NEBL</i> [†]	2.1e - 05	2.4	<i>AGTR1</i>	2.0e - 02	3.4
<i>CD19</i>	1.0e - 12	4.4	<i>SPARCL1</i>	7.5e - 18	5.2	<i>DNAPTP6</i> * [†]	2.2e - 08	1.8	<i>NFIB</i>	5.4e - 03	3.0
<i>SPOCK1</i>	3.0e - 06	6.2	<i>TPM3</i>	2.6e - 07	4.2	<i>KCNJ15</i> [†]	7.0e - 05	1.9	<i>FABP4</i>	3.8e - 02	7.0
<i>KRT8</i>	7.6e - 09	4.3	<i>TNNT2</i>	1.3e - 07	3.6	<i>PCOLCE</i> * [†]	1.7e - 06	1.9	<i>DKFZP686A01247</i>	2.4e - 04	2.5
<i>CRYAB</i>	2.2e - 06	4.3	<i>PDGFRA</i>	1.2e - 21	4.6	<i>SATB2</i> [†]	7.1e - 03	1.6	<i>EDNRA</i>	1.1e - 04	3.2
<i>ADAMTS1</i>	1.7e - 07	5.4	<i>CXCL14</i>	1.2e - 06	4.2	<i>PCDH9</i> [†]	1.8e - 03	1.9	<i>TSPAN8</i>	2.7e - 03	1.9
<i>MEIS2</i>	2.2e - 07	0.1	<i>WT1</i>	1.4e - 10	0.2	<i>HAS2</i> * [†]	5.1e - 12	0.2	<i>HAS2</i>	1.3e - 04	0.3
<i>LPHN2</i>	1.1e - 07	0.2	<i>KLK6</i>	3.0e - 29	0.2	<i>KLK6</i> * [†]	1.8e - 14	0.3	<i>TMEM100</i>	2.7e - 04	0.3
<i>HMGAA2</i>	1.6e - 08	0.1	<i>ITPR1</i>	3.6e - 19	0.2	<i>TMEM100</i> [†]	2.8e - 09	0.2	<i>TCEAL2</i>	3.0e - 04	0.4
<i>TCF21</i>	9.1e - 07	0.2	<i>ITGA8</i>	8.2e - 15	0.2	<i>WT1</i> * [†]	7.8e - 07	0.3	<i>BMPER</i>	4.0e - 06	0.4
<i>PLAGL1</i>	7.1e - 08	0.2	<i>TMEM100</i>	1.7e - 12	0.2	<i>TCEAL2</i> [†]	6.3e - 09	0.3	<i>ZBTB10</i>	1.4e - 04	0.5
<i>CCND2</i>	5.2e - 05	0.2	<i>STAT4</i>	7.6e - 30	0.2	<i>ITGA8</i> * [†]	2.1e - 07	0.3	<i>WT1</i>	1.0e - 07	0.4
<i>PLAG1</i>	3.3e - 10	0.2	<i>LYPD1</i>	6.7e - 11	0.2	<i>ITPR1</i> * [†]	2.2e - 10	0.3	<i>STAT4</i>	2.1e - 04	0.3
<i>SIX2</i>	7.9e - 07	0.2	<i>BMPER</i>	3.7e - 33	0.3	<i>STAT4</i> [†]	2.8e - 16	0.2	<i>GPR64</i>	2.2e - 03	0.4
<i>POSTN</i>	8.5e - 05	0.2	<i>ADAMTS3</i>	3.8e - 15	0.4	<i>IFI16</i> [†]	3.1e - 05	0.3	<i>CD1D</i>	3.7e - 15	0.4
<i>HTR2B</i>	1.9e - 06	0.2	<i>WASF3</i>	5.0e - 16	0.2	<i>ADAMTS3</i> * [†]	1.1e - 08	0.5	<i>TBL1XR1</i>	3.9e - 03	0.5
<i>CDH7</i>	1.7e - 13	0.2	<i>HAS2</i>	1.5e - 13	0.3	<i>FLRT3</i> [†]	5.0e - 04	0.5	<i>SLC7A11</i>	4.2e - 08	0.3
<i>TUBB2B</i>	1.9e - 07	0.5	<i>TCEAL2</i>	1.2e - 08	0.3	<i>KRT19</i> [†]	4.0e - 03	0.8	<i>BEX1</i>	1.7e - 06	0.4
<i>RHOBTB3</i>	5.2e - 10	0.4	<i>TCF21</i>	1.3e - 07	0.4	<i>PCTP</i> * [†]	1.0e - 11	0.5	<i>CITED1</i>	1.6e - 03	0.4
<i>FZD2</i>	3.0e - 06	0.2	<i>KRT19</i>	4.4e - 09	0.4	<i>CITED1</i> [†]	3.4e - 04	0.6	<i>HTR2B</i>	9.4e - 03	0.4
<i>SNAI2</i>	9.3e - 06	0.3	<i>MEOX1</i>	1.3e - 06	0.4	<i>SLC7A11</i> [†]	4.7e - 06	0.3	<i>HOXC10</i>	2.7e - 02	0.5
<i>CHRNA1</i>	6.6e - 14	0.3	<i>NTRK2</i>	4.2e - 09	0.3	<i>HOXC6</i> [†]	3.5e - 03	0.5	<i>GCH1</i>	1.7e - 03	0.5
<i>FZD10</i>	8.5e - 09	0.3	<i>ID4</i>	4.1e - 11	0.2	<i>ADAMTS5</i> [†]	3.1e - 04	0.5	<i>ADAMTS3</i>	6.3e - 03	0.6
<i>LEF1</i>	2.8e - 07	0.3	<i>DBC1</i>	4.4e - 23	0.3	<i>PNMA2</i> [†]	7.7e - 09	0.5	<i>ITGA8</i>	2.5e - 03	0.5
<i>TMEM45A</i>	3.7e - 07	0.3	<i>FZD6</i>	2.4e - 08	0.5	<i>MMP23</i> [†]	5.2e - 07	0.5	<i>MEIS2</i>	2.7e - 02	0.5
<i>MEIS1</i>	2.6e - 07	0.2	<i>CD1D</i>	1.6e - 18	0.4	<i>GCH1</i> [†]	1.6e - 03	0.6	<i>KLK6</i>	1.1e - 03	0.5
<i>ELAVL4</i>	3.2e - 24	0.3	<i>SLC7A11</i>	1.9e - 17	0.2	<i>IL15</i> [†]	2.4e - 11	0.6	<i>ITPR1</i>	3.7e - 03	0.5
<i>ASPN</i>	1.3e - 03	0.3	<i>CDH7</i>	1.4e - 13	0.4	<i>WNT5B</i> [†]	3.0e - 07	0.6	<i>HOXD11</i>	1.4e - 02	0.5

Table W1. (continued)

S1 vs S5			S2 vs S5			S3 vs S5			S4 vs S5		
Gene	P	FC	Gene	P	FC	Gene	P	FC	Gene	P	FC
<i>IGF2BP3</i>	2.2e - 10	0.1	<i>ADCY2</i>	8.9e - 19	0.3	<i>MGAM</i> [†]	1.4e - 09	0.5	<i>SALL1</i>	5.2e - 04	0.5
<i>C3ORF52</i>	1.5e - 24	0.3	<i>FOXD1</i>	4.1e - 06	0.4	<i>HOXD1</i> [†]	1.1e - 14	0.4	<i>LYPD1</i>	9.8e - 05	0.4
<i>TRIM28</i>	1.8e - 04	0.3	<i>SPOCK2</i>	8.2e - 14	0.3	<i>WWP1</i> [†]	5.9e - 08	0.5	<i>MGC24039</i>	1.5e - 06	0.6
<i>COL6A3</i>	1.2e - 05	0.2	<i>FGFR2</i>	1.8e - 16	0.4	<i>PLVAP</i> [†]	3.9e - 09	0.5	<i>PBK</i>	5.9e - 03	0.6
<i>LRRTM2</i>	2.0e - 10	0.3	<i>HTR2B</i>	1.4e - 04	0.4	<i>PLCH1</i> [†]	8.1e - 10	0.5	<i>HOXD10</i>	1.0e - 02	0.6
<i>UCHL1</i>	1.6e - 07	0.2	<i>EYA1</i>	5.5e - 08	0.4	<i>BMPER</i> [†]	4.3e - 05	0.5	<i>C14ORF109</i>	4.9e - 04	0.6
<i>MMP23</i>	1.8e - 09	0.3	<i>ANKRD15</i>	4.1e - 16	0.4	<i>DUSP5</i> [†]	3.8e - 14	0.5	<i>PTPN14</i>	2.2e - 05	0.6
<i>DBC1</i>	4.3e - 09	0.3	<i>PTPN14</i>	1.9e - 24	0.5	<i>LYPD1*</i> [†]	4.5e - 06	0.4	<i>ADCY2</i>	1.2e - 05	0.5
<i>VASH2</i>	1.2e - 07	0.4	<i>HOXD10</i>	4.6e - 09	0.4	<i>MEIS2</i> [†]	8.6e - 03	0.6	<i>HOXC6</i>	7.0e - 02	0.6
<i>C4ORF31</i>	1.0e - 12	0.3	<i>TACSTD1</i>	2.0e - 07	0.4	<i>RP11-301I17.1</i> [†]	4.5e - 07	0.6	<i>ANKRD6</i>	6.1e - 04	0.6
<i>PRAME</i>	1.8e - 04	0.2	<i>IL13RA1</i>	9.9e - 13	0.4	<i>DNM1</i> [†]	3.0e - 09	0.5	<i>CYP51A1</i>	8.2e - 13	0.5
<i>COL6A2</i>	1.0e - 05	0.3	<i>HOXD1</i>	9.6e - 28	0.4	<i>GRIK2</i> [†]	1.3e - 06	0.5	<i>NRXN1</i>	3.5e - 09	0.6
<i>SLC19A2</i>	2.0e - 09	0.4	<i>MMP23</i>	8.9e - 12	0.5	<i>ECELI</i> [†]	6.5e - 12	0.5	<i>UTX</i>	5.8e - 04	0.6
<i>HOXA5</i>	3.8e - 05	0.3	<i>HOXD11</i>	3.6e - 06	0.4	<i>PON2*</i> [†]	7.5e - 08	0.5	<i>SH3GLB2</i>	3.2e - 04	0.6
<i>SACS</i>	1.1e - 16	0.4	<i>LSR</i>	1.2e - 08	0.5	<i>WASF3*</i> [†]	7.8e - 08	0.5	<i>DOCK4</i>	3.9e - 04	0.6
<i>BMPER</i>	3.5e - 13	0.3	<i>DUSP9</i>	3.3e - 13	0.5	<i>IL13RA1*</i> [†]	2.5e - 06	0.6	<i>CLGN</i>	3.1e - 04	0.6
<i>ISLR</i>	5.1e - 06	0.4	<i>VAMP8</i>	5.9e - 08	0.5	<i>CD1D</i> [†]	1.2e - 02	0.7	<i>CDH7</i>	1.7e - 03	0.6
<i>TNC</i>	2.3e - 07	0.5	<i>DPP6</i>	6.1e - 17	0.5	<i>RAB27A</i> [†]	5.9e - 11	0.5	<i>MIR16</i>	1.6e - 04	0.6
<i>TTK</i>	5.1e - 05	0.4	<i>UNC5C</i>	2.0e - 17	0.5	<i>SPOCK2</i> [†]	8.2e - 09	0.5	<i>LOC63920</i>	1.1e - 04	0.6
<i>ANK2</i>	2.3e - 06	0.4	<i>CXXC4</i>	1.7e - 14	0.5	<i>GPR64</i> [†]	7.4e - 03	0.6	<i>FOXD1</i>	2.3e - 03	0.5
<i>GJA1</i>	1.4e - 03	0.3	<i>SALL1</i>	4.9e - 06	0.5	<i>FAM59A*</i> [†]	7.8e - 05	0.6	<i>DBC1</i>	7.2e - 04	0.5
<i>LRRC17</i>	1.2e - 06	0.4	<i>MGAM</i>	8.9e - 15	0.4	<i>CYP51A1</i> [†]	8.2e - 09	0.6	<i>SMAD4</i>	2.2e - 03	0.6
<i>RP11-301I17.1</i>	1.6e - 13	0.4	<i>LAMA4</i>	1.7e - 14	0.4	<i>LPPR4</i> [†]	2.3e - 08	0.6	<i>MAK10</i>	5.7e - 03	0.6
<i>AH11</i>	1.1e - 06	0.4	<i>PAX2</i>	6.2e - 07	0.4	<i>BCL2*</i> [†]	8.2e - 11	0.5	<i>ZNF516</i>	3.1e - 04	0.6
<i>KIF26B</i>	8.1e - 14	0.4	<i>GCH1</i>	1.4e - 06	0.5	<i>HOXC10</i>	3.4e - 02	0.7	<i>KIF24</i>	6.2e - 05	0.6
<i>PHLDA1</i>	8.8e - 08	0.2	<i>DSG2</i>	1.0e - 08	0.5	<i>ANKRD15</i> [†]	6.3e - 06	0.6	<i>IL15</i>	1.3e - 06	0.6
<i>ANKRD6</i>	1.6e - 10	0.4	<i>PLCB1</i>	2.9e - 07	0.4	<i>DLK1</i>	1.8e - 03	0.3	<i>PCTP</i>	4.7e - 05	0.6
<i>COL3A1</i>	3.5e - 04	0.3	<i>PCTP</i>	7.1e - 17	0.5	<i>S100AI</i> [†]	6.1e - 05	0.6	<i>HAT1</i>	6.0e - 03	0.6

^{*}PCC with *WT1* > ±0.6.[†]Also coordinately significantly (*P* < .001) differentially expressed in comparison S2 versus S5.

Classification of Malignant Pediatric Renal Tumors by Gene Expression

Chiang-Ching Huang,² Colleen Cutcliffe,¹ Cheryl Coffin,³ Poul H.B. Sorensen,⁴ J. Bruce Beckwith,⁵ and Elizabeth J. Perlman^{1*} for the Renal Tumor Committee of the Children's Oncology Group

Background. The most common malignant renal tumors of childhood are Wilms tumor (WT), clear cell sarcoma of the kidney (CCSK), cellular mesoblastic nephroma (CMN), and rhabdoid tumor of the kidney (RTK). Because these tumors present significant diagnostic difficulties, the goal was to define diagnostically useful signatures based on gene expression. **Procedures.** Gene expression analysis using oligonucleotide arrays was performed on a training set of 47 tumors (10 CCSKs, 9 CMNs, 8 RTKs, and 20 WTs). Classifiers were developed for each tumor type using variations of compound covariate class predictor. The classifiers were applied to an independent test set of 72 tumors (3 CMN, 7 CCSK, 4 RTK, and 58 WT). Central review diagnosis was utilized as the gold standard. Correlation with the institutional diagnosis and qualitative estimation of confidence levels at the time of central review were noted.

Key words: clear cell sarcoma of kidney; gene expression; mesoblastic nephroma; pediatric renal tumors; rhabdoid tumor; Wilms tumor

Results. Within the training set, classifiers resulted in no errors when >10 genes were utilized. Top genes in each classifier were verified using quantitative reverse transcription-polymerase chain reaction (RT-PCR). Applying the classifiers to the test set, 71 of 72 tumors were correctly classified with a confidence level of >99%. The exception was incorrectly classified by the gold standard. In comparison, by histopathology 31% of the non-WT were not accurately classified by the local institution, and 29% were classified with <95% confidence on central review. **Conclusions.** Classifiers based on gene expression provide diagnostic confidence and accuracy greater than that of pathologic analysis alone. Tumors that show ambiguous gene expression profiles are those that are also pathologically and molecularly ambiguous and merit further analysis. Pediatr Blood Cancer 2006;46:728–738. © 2006 Wiley-Liss, Inc.

INTRODUCTION

The most common malignant renal tumor of childhood is Wilms tumor (WT, 85%). This is followed by three sarcomas that have a predilection for the kidney, namely clear cell sarcoma of the kidney (CCSK, 5%), cellular congenital mesoblastic nephroma (CMN, 4%), and rhabdoid tumor of the kidney (RTK, 3%). While most WTs are easily recognized by their characteristic combination of epithelial, blastemal, and stromal elements, approximately 5% of WT are composed exclusively of undifferentiated blastemal or stromal elements [1]. Such tumors are easily mistaken for CCSK, CMN, and RTK. It has been the experience of the Children's Oncology Group (COG) Renal Tumor Pathology Center that approximately 25% of these tumors are not correctly classified by local pathologists. Even after central pathology review, a small but significant number of tumors are diagnosed with less than certainty. The importance of an accurate diagnosis is most evident when the differences in therapy and prognosis among the four different tumors are considered. While tools such as immunohistochemistry, molecular analysis for fusion genes, and detection of gene mutations and deletions may help distinguish individual tumor types, such tests are often unreliable or not clinically available. The goals of this study were to determine if gene expression could accurately and reliably distinguish CCSK, CMN, RTK, and WT, and to define diagnostically useful gene expression signatures for each entity. In-depth analysis of corresponding protein expression and pathway analysis will be performed for each entity and reported separately.

MATERIALS AND METHODS

Specimen Collection

Analyzed specimens were submitted to the Renal Tumor Bank of the COG by institutions participating in the National Wilms Tumor Study-5 (NWTS-5) clinical protocol. The samples were snap frozen immediately following surgical removal and mailed on dry ice. Tissues were submitted by the Tumor Bank to the authors after review and approval of the research plan by the IRB of Children's Memorial Hospital. For all cases, the pathologic diagnosis was established at the time of initial registration by central pathology review from 1995 to 2002 with

¹Department of Pathology, Robert H. Lurie Comprehensive Cancer Center, Northwestern University's Feinberg School of Medicine, Chicago, Illinois; ²Department of Preventive Medicine, Robert H. Lurie Comprehensive Cancer Center, Northwestern University's Feinberg School of Medicine, Chicago, Illinois; ³Department of Pathology, Primary Children's Medical Center, University of Utah School of Medicine, Salt Lake City, Utah; ⁴Department of Pathology and Pediatrics, Children's and Women's Health Centre of British Columbia, Vancouver, British Columbia; ⁵Department of Pathology and Human Anatomy, Loma Linda University, Loma Linda, California

Grant sponsor: NIH; Grant numbers: U01 CA88131, U01-CA114757, U10 CA 42326, U10CA98543.

*Correspondence to: Elizabeth J. Perlman, Children's Memorial Hospital, 2300 Children's Plaza Box 17, Chicago, IL 60614.
E-mail: eperlman@childrensmemorial.org

Received 15 September 2005; Accepted 16 December 2005

full knowledge of the institutional diagnosis. The central pathology diagnosis was provided by JBB from 1995 to 1999, and by EJP from 2000 to 2002. All cases were again reviewed by both JBB and EJP prior to their inclusion in the study.

Training set. Frozen tissue samples from 60 pediatric renal tumors were obtained. Ten WT were selected for their monophasic undifferentiated histologic pattern. This is the WT pattern that is often unable to be confidently diagnosed. The rest of the tumors in the training set were unselected. Frozen sections from each tumor were stained by hematoxylin and eosin and evaluated prior to RNA isolation. Samples containing less than 80% viable tumor cellularity (six tumors) were rejected. After quality control (see below for details), seven additional samples were rejected due to poor RNA quality. The remaining 47 tumors (10 CCSKs, 9 cellular CMNs, 8 RTKs, and 20 WTs all of favorable histology) formed the training set for this study.

Test set. To further assess gene expression signatures derived from the training set, a test set was analyzed in a blinded fashion. Frozen tissue samples from an independent test set of 87 randomly selected tumors were analyzed for gene expression using the same methods described above for the training set. Eight tumors were rejected after histologic analysis and seven additional tumors were rejected due to poor RNA quality. The pathologic diagnoses of the 72 tumors within the test set included 3 CMN, 7 CCSK, 4 RTK, and 58 favorable histology WT.

Diagnostic Confidence Assessment

The classifiers based on gene expression that were generated in this study are associated with statistical confidence levels in the form of probability of error (as described below). We sought to compare the confidence that resulted from these classifiers with the confidence that resulted from pathologic classification as applied during NWTS-5. To accomplish this, the pathology confidence levels were assessed in two ways. First, the diagnosis provided at the original institution was compared to the diagnosis provided at central pathology review. Second, the subjective confidence level of the central pathology reviewer with his/her final diagnosis (a diagnosis which took into account all information available at the time of registration) was estimated by evaluating the original report as well as any information that subsequently became available. This was possible due to the central pathology review practice of providing in-depth reports that indicated any diagnostic uncertainties that existed following central review. Confidence levels of >95%, 90–95%, 75–89%, and <75% were accordingly assigned. It should be emphasized that the goal was not to determine the relative utility of particular diagnostic tests (such as reverse transcription-polymerase chain reaction (RT-PCR), or immunohistochemistry), but to determine how gene expression classifiers compare to the current general practice of pathologists.

Gene Expression Measurement

RNA isolation. Total RNA was isolated from 75 to 100 mg frozen tissue by Trizol (Invitrogen, Carlsbad, CA) extraction and treated with DNase I (Roche, Indianapolis, IN). No amplification of RNA was performed. The RNA was purified by RNeasy Mini Kit (Qiagen, Valencia, CA) and concentration determined by optical density as measured by a GeneQuant DNA/RNA calculator (Amersham Biosciences, Piscataway, NJ). Samples with A_{260}/A_{280} of less than 1.8 were excluded. The quality of the RNA was further assessed by a 1% SDS-PAGE gel.

cDNA synthesis. cDNA synthesis was performed with SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen). The manufacturer's protocol was modified by the use of an HPLC-purified T7-(dT)₂₄ primer (GenSet Oligos, ProLigo, Boulder, CO) and incubation at 42°C rather than 37°C. The double-stranded cDNA product was purified by phenol/chloroform extraction using Phase Lock Gels (Eppendorf Scientific, Westbury, NY).

Biotin-labeled cRNA synthesis and hybridization. In vitro transcription (IVT) reaction was performed on cDNA synthesized from 10 µg of RNA using the BioArray High Yield RNA Transcript Labeling Kit (Enzo, Farmingdale, NY). The resulting biotin-labeled cRNA was purified using RNeasy Mini Kit (Qiagen) and quantified using a GeneQuant DNA/RNA calculator (Amersham Biosciences). Samples were rejected from further analysis if the yield from the IVT reaction was less than 30 µg. Labeled cRNA (17 µg) was fragmented, hybridized to the HG-U133A chip, and scanned according to Affymetrix protocol (described at <http://www.affymetrix.com/products/arrays/specific/hgu133.affx>).

Quality control. Array images were assessed by eye to confirm scanner alignment and absence of significant bubbles or scratches. Ratios (3'/5') for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were confirmed to be less than 3.4, and *BioB* spike controls were confirmed as present on 100% of the chips. *BioC*, *BioD*, and *cre* were also present and in increasing intensity. When scaled to a target intensity of 2,500, scaling factors for all arrays were within threefold of each other. Background (range 33.63–122.30), raw Q-values (range 1.34–3.71), and mean intensities were within acceptable limits in accordance with Affymetrix guidelines for assessing array quality. Percent present calls ranged from 30 to 55.1.

Real-time quantitative RT-PCR. Based on availability of additional sample, RNA from unselected samples of 7 CCSKs, 7 CMNs, 7 RTKs, and 7 WTs that were examined for gene expression were further analyzed by RT-PCR to confirm RNA expression levels of two genes from each classifier. The ABI Prism 7700 Sequence Detection System (Applied Biosystems, Foster City, CA) was utilized. Each 50 µl RT-PCR reaction volume included 25 ng total RNA, 5 µl Taqman Buffer A (500 mM KCl, 100 mM Tris-HCl, 0.1 M EDTA, 600 nM passive reference dye, pH 8.3 at room temperature),

TABLE I. Primers and Probes Utilized

Gene	Taqman probe	Forward primer	Reverse primer
<i>FOXF1</i>	cttgcgtatgtgaggcgcccg	caggtcacctaccaagacatcaag	ccgacggttataccctcgagaag
<i>NPTX1</i>	tggaatgagccgaggatttaatcagttc	accttgttggagctgacttact	catgtAACCCaaGGTTCTtatcaa
<i>HLX1</i>	ttctccaacctcgagaggaaaggcct	ccatgeccagacgtaca	gcttggcacgtacttctgaatctc
<i>SPRY4</i>	accgtctgcgcgcgcctg	acggctgcgtgaaget	tgctgcagtttttagatgtcaga
<i>NQO1</i>	tggtcagaaggaaattgtcagagaaggtaa	tccaaatgttattaaatcacctctgtta	cttgtctaggcatattgtactgtct
<i>PLP2</i>	cgttcggcagccaagacatacagca	tctacctgtatcacccatgttgc	gcctfacaccggccatct
<i>EYA1</i>	atggaggaccacgttattttcatcagaacagc	gcctggatgcgcaggat	cgttagttaatgtggcagacacata
<i>PAX2</i>	cagcaacgtgtcaggccacacagacatac	tgaagtcaacgtcgagtctatctgcatt	gagtgggtctcgccatgc
<i>ACTIN</i>	atgcctccccatgcacatctgcgt	tcacccacactgtgcacatcta	cagcggaaaccgcgtatgcacatgg

These were designed using Primer Express software (Applied Biosystems). Taqman probes, labeled with reporter dye FAM and quencher dye QSY7, and the actin primers used as a control were synthesized by Megabases, Inc. (Evanston, IL); the remaining primers were synthesized by Operon (San Francisco, CA).

10 μ l of 25 mM MgCl₂, 1.5 μ l of each dNTP (10 mM dATP, dCTP, and dGTP and 20 mM dUTP), 0.5 μ l forward and reverse primers (10 μ M), 1 μ l of the corresponding Taqman probe (5,000 nM), 0.25 μ l AmpliTaq Gold supplied at 5U/ μ l, and 0.25 μ l MuLV Reverse Transcriptase. RT-PCR cycle parameters were 48°C for 30 min, 95°C for 15 min, followed by 40 cycles at 95°C for 15 sec and 59°C for 1 min. The primers and probes used in the study were designed using Primer Express software (Applied Biosystems). Taqman probes, labeled with the reporter dye FAM and the quencher dye QSY7, and the actin primers used as a control were synthesized by Megabases, Inc. (Evanston, IL); the remaining primers were synthesized by Operon (San Francisco, CA). Sequences of probes and primers are shown in Table I. β -actin was used as the endogenous control. Each threshold cycle (C_T), the cycle at which an increase in reporter fluorescence goes slightly over the optimal value line, was determined and the C_T for β -actin was subtracted from this for normalization purposes.

Data Analysis

Normalization and filtering. The microarray data consists of 22,283 probe sets from Affymetrix HG-U133A chip as described in <http://www.affymetrix.com/products/arrays/specific/hgu133.affx>. To translate the scanned images into expression analysis files, we used positional-dependent-nearest-neighbor model (PDNN) [2], a software package (PerfectMatch) available online at <http://odin.mdacc.tmc.edu/~zhangli/PerfectMatch/>. PDNN assigns a weighted average of probe intensity values for each probe set and automatically normalizes the data from all arrays. After normalization, the range of intensity for all genes varied from 4.95 to 10.79 using a log scale. Genes with low expression (maximum expression less than a log scale of 6 across the 47 tumors in the training set) were filtered out. Genes showing low variation (defined as a range of expression less than 0.6 log scale across all 47 tumors) and Affymetrix control genes also were filtered out. Filtering left a subset of 11,517 probe sets for analyses.

Unsupervised analysis. To discover gene expression patterns among the different tumor types, hierarchical clustering and principal component analysis were utilized. Average-linkage hierarchical clustering was performed using CLUSTER and the results were displayed using TREEVIEW (software at <http://rana.lbl.gov/EisenSoftware.htm> as described by Eisen et al. [3]). Principal component analysis (PCA) is a dimension reduction technique in which the first few components can capture the most data variance. This was performed using the R statistical package (<http://www.r-project.org/>) using singular value decomposition method.

Supervised analysis. To identify genes from which mathematical rules can be formulated to assign tumors to specific classes, we used a compound covariate class predictor (CCP-1) described by Hedenfalk et al. [4]. This method compares the gene expression of the most informative genes for each class of tumor to all remaining tumors in the training set and calculates the t-statistics for each gene. This is defined as the difference between the mean gene expression within the class and the mean gene expression of all remaining classes divided by the square root of the pooled variance in gene expression. The K genes (from 1 to 100 genes) with the largest positive t-statistics are selected for the classifier of that tumor type. Using the expression levels of the genes in the classifier for each sample, this method then constructs a tumor-class score that is a linear combination of the weighted expression of selected genes within the classifier. The weighting is provided by the t-statistics. The higher the score for a particular classifier, the more likely it is that the tumor belongs to that category. Using these methods, four classifiers were generated, one for each tumor type.

Different selection criteria for genes within a classifier may have merits and weaknesses. A large t-statistic could result from small pooled variance of gene expression while the difference in mean expression could be small and of questionable biological significance. We therefore varied the CCP by selecting the top K genes using only the difference between the mean gene expression within the class and the

mean gene expression in all remaining classes after standardizing the gene expression values. Because log scale was used, this value approximates fold-change. To preserve the robustness that t-statistics offers with regard to reliability of expression, the weighting utilized in the classifier continued to be the t-statistic for each corresponding gene. The four new classifiers were designated CCP-2.

Assigning tumor type to samples using gene expression. A logistic regression model using tumor-class score as the only covariate was used to assign a predictive probability for each sample to a class based on expression of the genes in the classifier for all tumors. To make the decision call for each tumor sample, it was assigned to the class that had the largest predictive probability. However, no decision was made if the largest predictive probability was less than 0.5.

Leave-one-out cross validation (LOOCV). To test the performance of each classifier, LOOCV [5] was utilized in the training set. In brief, for each CCP method above, a single tumor was removed from the dataset and a new classifier was built using the remaining 46 samples. The tumor that was removed was classified using the new classifiers. The process was repeated for each tumor and the cumulative error rate was recorded. The top K genes in each LOOCV procedure that were examined ranged from 1 to 100.

RESULTS

Pediatric Renal Tumors Can Be Distinguished by Gene Expression Patterns

The training set was analyzed for expression of the 11,517 probe sets that remained following filtering of probe sets that showed low expression and low variation across all the tumors. That dataset was analyzed first using two unsupervised methods: hierarchical clustering and PCA. A global visualization of the gene expression patterns using hierarchical clustering is illustrated in Figure 1A. Each malignant pediatric renal tumor demonstrated distinct gene expression patterns, and clustered correctly with tumors of the same pathologic subtype. Notably, each type of tumor had a unique gene expression signature compared with all the other types. Visualized through PCA, most tumors populated discrete areas of the graph, indicating a high degree of similarity in gene expression within each tumor type (Fig. 1B).

Classifiers Can Be Developed that Correctly Identify Each Tumor in the Training Set

The unsupervised methods described above document the native dissimilarity of gene expression patterns between the different histologic types of renal tumors, and the corresponding similarity within each tumor type. To

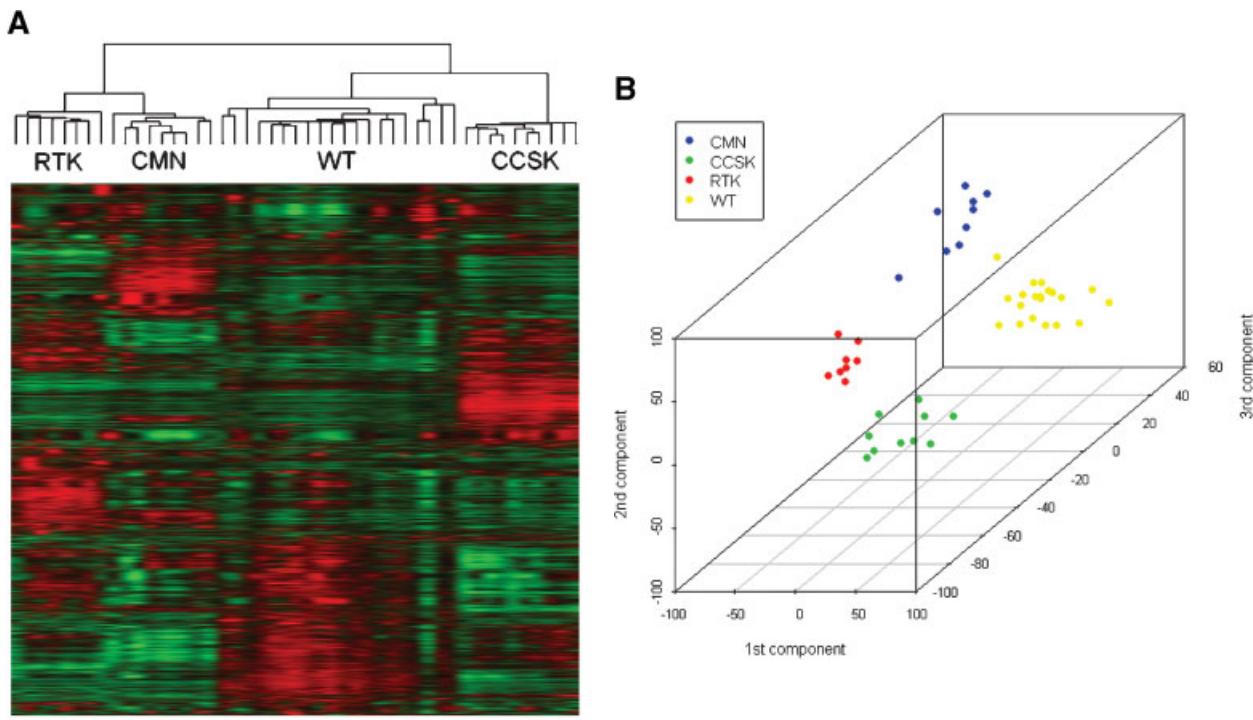


Fig. 1. Classification of pediatric renal tumors by gene expression. **A:** Hierarchical clustering of normalized log-transformed expression of 11,517 genes. Each malignant pediatric renal tumor shows distinct gene expression patterns and clusters correctly with like tumors. **B:** Principal component analysis of 47 tumors in the training set using 11,517 genes. Each tumor type populates a discrete area of the graph, indicating a high degree of similarity in gene expression within each tumor type. CMN, cellular mesoblastic nephroma; WT, Wilms tumor; CCSK, clear cell sarcoma of kidney; RTK, rhabdoid tumor of the kidney.

develop a specific set of genes (a classifier) that could be used to accurately and consistently assign a tumor to a specific category or class, we relied on two variations of compound covariate class predictor (CCP-1 and CCP-2). LOOCV was used in the training set to test the performance of the eight classifiers developed by CCP-1 and CCP-2. The total error number was zero if the classifier used more than

the 10 and 5 top genes in CCP1 and CCP2, respectively (Fig. 2A). We investigated the performance of each individual tumor sample further. For all but one tumor sample, the predictive probability remained close to 1 for the true class and close to 0 for the wrong classes regardless of the number of top K genes used (Fig. 2B). This is a good indication of the stability of the classifiers. The only

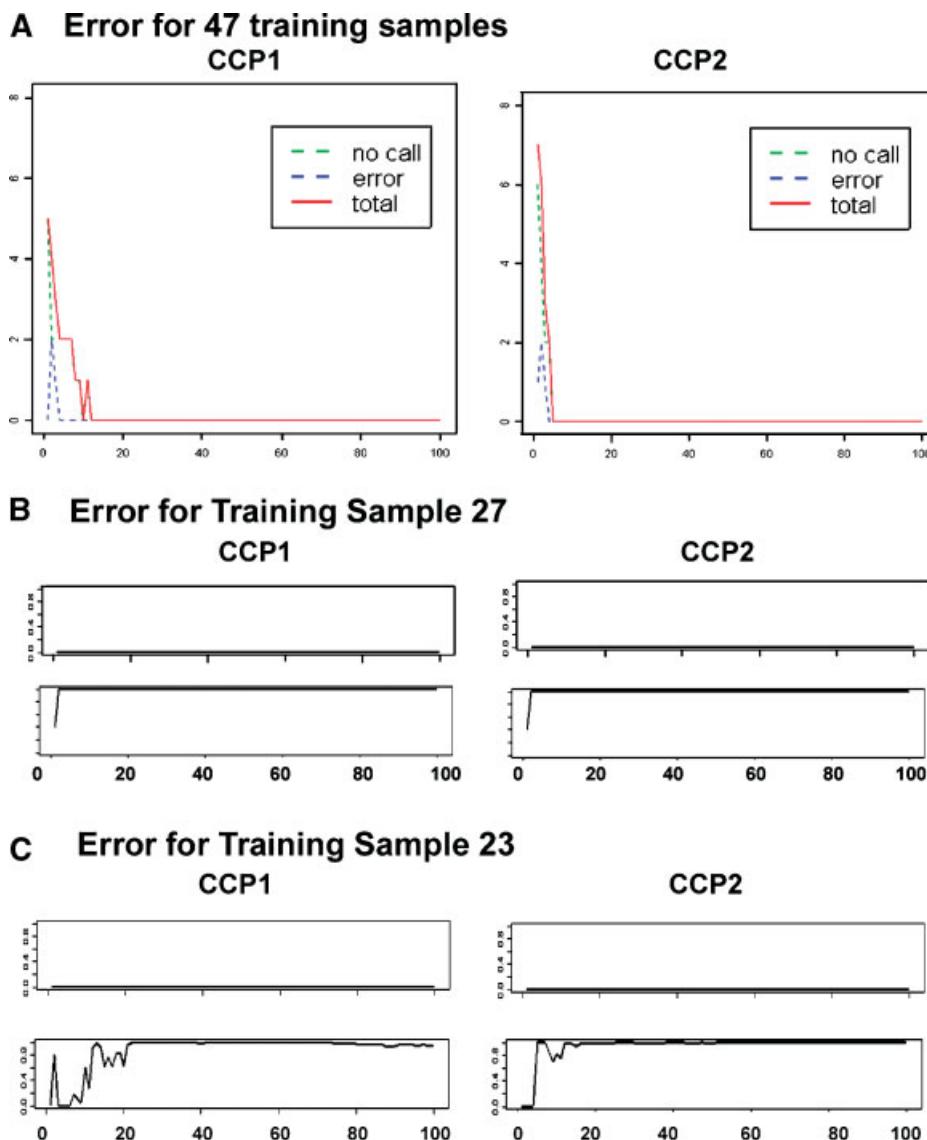


Fig. 2. Analysis of error within the training set. **A:** The total number of leave-one-out errors (red solid line), the number of 47 tumor samples misclassified (blue dashed line) and the number of tumor samples for which no decision was made (green dashed line) on the Y-axis versus the number of genes used in the classifier for CCP-1 and CCP2 on the X-axis. **B:** Error performance for a training sample that is representative of 46 tumors in the training set. The CCSK, CMN, and WT classifiers demonstrated zero probability regardless of the number of genes in the classifier (only the CCSK classifier is shown, top graph). The RTK classifier demonstrates close to 100% probability regardless of the number of genes in the classifier, and regardless of whether CCP-1 or CCP-2 was utilized (lower graph). The number of genes utilized in the classifier is indicated on the X-axis and the probability of correctly classifying the tumors is indicated on the Y-axis, ranging from 0 to 1.0. **C:** Classification error of training sample 23, which demonstrated a probability of 0 for the CMN, CCSK, and WT classifiers regardless of the number of genes utilized in the classifier (only CCSK classifier is shown, upper). When the RTK classifier was applied, there was a low probability when using a small number of genes (<10) in CCP-1 and CCP-2, however the tumor was successfully classified as RTK with greater than 99% probability when more than 20 genes were utilized (lower). The number of genes utilized in the classifier is indicated on the X-axis and the probability of correctly classifying the tumors is indicated on the Y-axis ranging from 1 to 1.0.

exception was training sample 23, which demonstrated no call with a very low probability for classifiers when using a small number of genes (<10) in both CCP-1 and CCP-2. However, it was successfully classified as RTK with greater than 99% probability when more than 20 and 10 genes were used in CCP-1 and CCP-2, respectively (Fig. 2C). Training sample 23 was further analyzed to determine any sources of variation that could explain this performance. Clinically, the tumor was Stage 4 at presentation and had the typical histologic appearance of RTK. The tumor was widely infiltrative in the initial nephrectomy sample, and the patient died a few months after diagnosis. These clinical and pathologic features are quite typical of RTK.

Genes Within the Classifiers Can Be Verified Using Quantitative RT-PCR

Analysis of the LOOCV indicated that the predominant diagnostic power lies in the top five genes of each classifier. Therefore, we selected two of the five top genes in each classifier for verification, based on their biologic function, and analyzed seven unselected tumors from each tumor type using quantitative RT-PCR. These findings confirmed the differential expression of each gene tested but highlight the differences in the relative strength of each gene as a classifier (Fig. 3). In particular, *NQO1* and *PLP2* demonstrated upregulation in RTKs, but this was not as dramatic as the upregulation of *NPTX1* in CCSKs or *EYA1* and *PAX2* in WTs.

Applying the Classifiers to an Independent Data Set Results in Accurate Diagnosis

We further examined the effectiveness of our classifiers using CCP-1 and CCP-2 on an independent test set of 72 pediatric renal tumors. Each classifier derived from the 47 tumors in the training set was applied to all 72 tumors, varying the number of genes in each classifier from 1 to 100. All but

two of these 72 tumors demonstrated unequivocally correct classifications, with a probability of falling into the true category of >99% and a probability of <1% of belonging to each of the false categories providing the number of genes in the classifier was greater than 10. The overall error experience corresponding to the number of genes utilized in the classifier is illustrated in Figure 4A. Two tumors are considered in greater detail below due to diagnostic ambiguities and because they provide useful insight into the different methods of classification, as well as the optimal number of genes that should be used with such classifiers.

Test sample 1 was confidently diagnosed by the original central pathology reviewer (JBB) as CMN at the time of original diagnosis. Retrospective pathologic analysis of this tumor (EJP) during this study confirmed this pathologic diagnosis. The tumor was then submitted to an external soft tissue sarcoma expert (CC), who classified the tumor as a primitive mesenchymal neoplasm with overlapping features of infantile myofibromatosis and CMN. Evaluation by RT-PCR failed to demonstrate expression of the *ETV6-NTRK* fusion product (data not shown). In addition, fluorescence *in situ* hybridization studies failed to reveal splitting of either the *ETV6* or *NTRK* loci (data not shown). Therefore, a variant translocation is unlikely. By gene expression profiling, this tumor was classified as CMN when greater than five genes were used in a classifier developed by CCP-2 with a probability of greater than 0.95. CCP-1 was more ambiguous, although significant probability remained that this tumor was correctly classified as CMN using a variable number of genes from 35 to 60 (Fig. 4B). When the other classifiers (CCSK, RTK, and WT) were applied, these demonstrated a probability of less than 1% of falling in these categories. In summary, this tumor does not demonstrate the fusion gene characteristic of CMN, yet has the gene expression features of CMN, and has pathologic features that are within the spectrum of infantile myofibromatosis and CMN.

Test sample 9 was classified as CCSK by central pathology review. This tumor represented a diagnostic quandary pathologically. The tumor was composed for the most part of small cells with inconspicuous nucleoli and scant cytoplasm. Such areas were similar to CMN. Areas of prominent hyalinization were present and small foci demonstrated prominent nucleoli and eccentric eosinophilic cytoplasm. Such areas were more characteristic of RTK. Other regions demonstrated a vascular pattern and sclerosis that suggested a diagnosis of CCSK. Immunohistochemistry revealed nonreactivity for epithelial, muscle, and neural markers. On central review, in the absence of positive evidence to support the diagnosis of RTK, a diagnosis of CCSK was made, although both the local pathologist and the central review pathologist remained highly skeptical of this diagnosis. The tumor was Stage 1, and the patient was alive 6 years after presentation. The experience of NWTSG-5 is that 50% of patients with Stage 1 RTK survive long term. The CCP-1 classification of this tumor was RTK using 20–55 genes, and

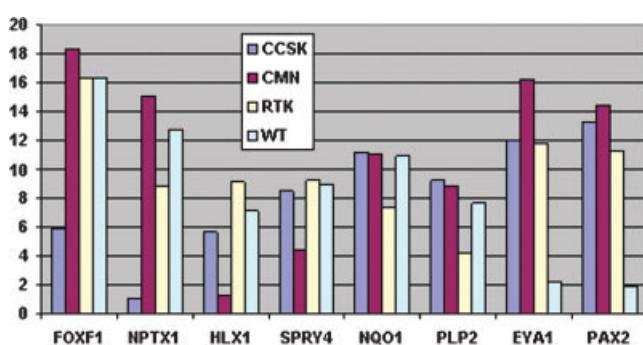


Fig. 3. Verification by quantitative RT-PCR. RNA from a subset of CCSK, WT, RTK, and CMN samples previously examined by microarrays were further analyzed by quantitative RT-PCR to confirm RNA expression levels of eight genes, two from each classifier. β -actin was used as the endogenous control. Shown are the average ΔC_T in arbitrary units for each tumor category. Increased levels of expression result in lower ΔC_T . This RT-PCR quantification resulted in similar fold-change when compared to gene expression analysis (data not shown).

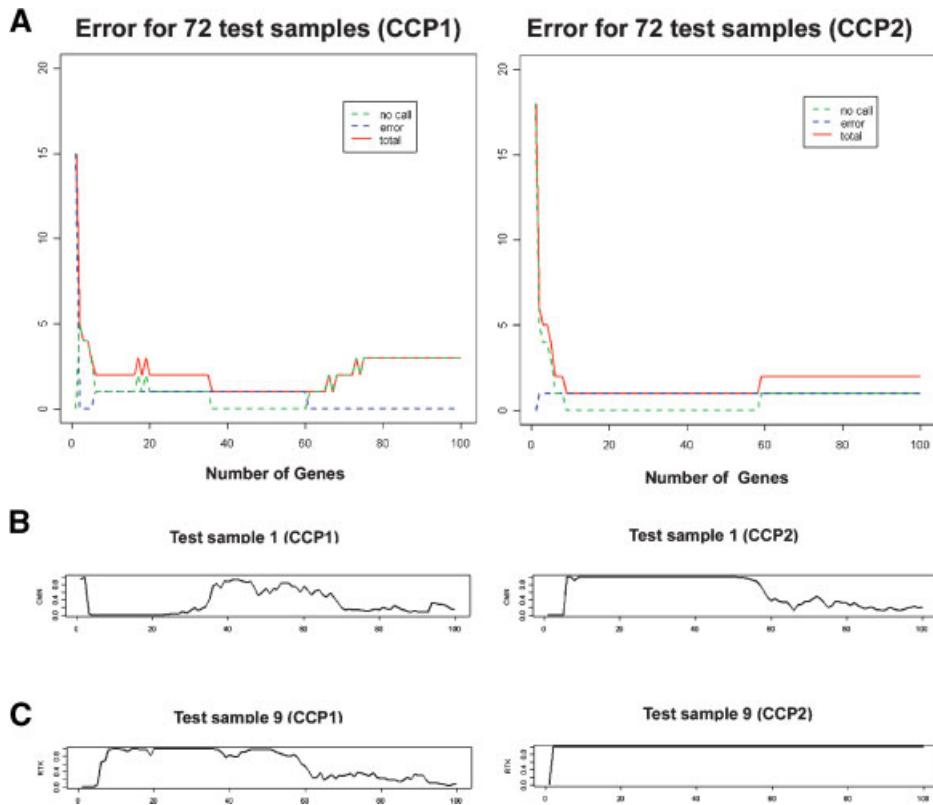


Fig. 4. Analysis of error within the test set. **A:** For the test set of 72 tumors, the total number of errors (red solid line), the number of tumor samples misclassified (blue dashed line), and the number of tumor samples for which no decision was made (green dashed line) on the Y-axis versus the number of genes used in the classifier for CCP-1 and CCP-2 on the X-axis. **B:** Test sample 1: This tumor was classified as CMN when 10–55 genes were utilized in the classifier developed by CCP-2, with a probability of greater than 0.95. However, when greater than 55 genes were utilized, increased errors were made by not classifying it as CMN (right). CCP-1 was less successful, although there remained a significant probability that this tumor was correctly classified as CMN using a variable number of genes from 35 to 60 (left). **C:** Test sample 9: This tumor was originally diagnosed as CCSK. Classification on the basis of gene expression by CCP-1 classified this tumor as RTK using 10–55 genes (left), and CCP-2 consistently classified the tumor as RTK (right).

CCP-2 consistently classified the tumor as RTK (Fig. 4C). Retrospective pathologic analysis, with the greater experience and additional diagnostic tools now available, is consistent with the diagnosis of RTK. Immunohistochemistry for the INI-1 protein demonstrated no nuclear expression in the tumor cells. This is in keeping with the current knowledge that RTKs show consistent genetic deletions or mutations of the INI-1 gene, resulting in loss of protein expression.

In summary, classifiers developed using both CCP-1 and CCP-2 were remarkably successful in correctly classifying pediatric renal tumors. CCP-2 provided calls that were unambiguous using a smaller number of genes; however, CCP-1 more sensitively identified cases such as test sample 1 that were histologically different from other tumors in the group.

Classifiers Based on Gene Expression Provide Increased Diagnostic Confidence When Compared to Pathologic Classification

Sufficient records were available to evaluate both the institutional and central review diagnoses for 39 of the 41

patients evaluated in this study who had CCSK, RTK, and CMN. Of these 39, six were diagnosed incorrectly at the original institution and an additional six were not given a diagnosis locally. Therefore, 12 of 39 (31%) of the tumors were not able to be accurately classified at the local institution. At central pathology review, 28 of 39 tumors were diagnosed with full confidence ($\geq 95\%$), 5 of 39 were diagnosed with 90–95% confidence, 3 of 39 with 75–89% confidence, and 3 of 39 with less than 75% confidence (including test sample 9 which was incorrectly diagnosed). Therefore, even at central pathology review, 29% of the tumors were not able to be diagnosed with the confidence generally expected. This is also the experience of central reviewers of the soft tissue sarcoma protocols due to the current lack of diagnostic tools and criteria for many rare sarcomas.

Summary of Genes Differentially Expressed in Each Tumor Type

Upon development of classifiers of each malignant renal tumor, the genes present in these classifiers were analyzed to gain insight into those genes that distinguish these tumors.

The top eight genes used for classification of each tumor type are identified in Table II. Among these we identified genes known to be upregulated in renal tumors, known oncogenes, regulatory factors of apoptosis, and genes upregulated in other tumor types. It is beyond the scope of this study to provide an in-depth analysis of the gene expression patterns of each tumor type. These will be reported separately.

DISCUSSION

Pediatric renal tumors provide significant diagnostic challenges; approximately 25% are inaccurately classified by the local institution. The majority of these difficulties arise in the evaluation of the sarcomas that arise in the kidney and their distinction from undifferentiated Wilms tumors. Sarcomas at all anatomic sites pose diagnostic problems, largely because their histologic features are not specific, immunohistochemical markers are highly variable within each tumor type, and even genetic markers are unreliable due to variants. This has resulted in tumors that lack defined diagnostic criteria. Lastly, the diagnostic work-up for sarcomas is often quite complex and expensive. Sarcomas that preferentially arise in the kidney (CCSK, CMN,

CMN, and RTK) are no exception and provide a useful model for the role of gene expression profiling in sarcoma diagnosis.

Gene expression profiling is a potentially powerful diagnostic tool for tumors such as pediatric sarcomas that do not have sufficient reliable, clinically available molecular markers. Profiling is also useful for tumors with available molecular markers such as translocations, deletions, and mutations, but markers for which a number of variants are recognized that are not detectable by routine methods. Lastly, gene expression profiling has the potential to uncover new therapeutic targets. Identification of such targets may open the door to more effective therapies, particularly for highly aggressive tumors such as RTK.

In this study, we found that the most common malignant tumors arising in the kidney show distinct and different gene expression profiles despite their frequent histologic similarities. Using a training set, cancer signatures were derived from these profiles, which were then applied successfully to a test set in a blinded fashion. This resulted in complete diagnostic accuracy and it provided much greater confidence when compared to routine pathologic diagnosis. Of significance, classification based on gene expression was

TABLE II. The Top Eight Upregulated Genes in Each Classifier With Fold-change >2 Using CCP-2

CMN			CCSK		
Symbol	Gene	Fold-change	Symbol	Gene	Fold-change
HLX1	H2.O-like homeo box 1	6.7	FOXF1	Forkhead box F1; tumor suppressor	7.3
SPRY4	Sprouty homolog 4; tumor suppressor; oncogene	6.1	HLXB9	Homeo box HB9	3.3
ANGPTL2	Angiopoietin-like 2	4.2	D4S234E	DNA segment chromosome 4	4.9
SRPUL	Sushi-repeat protein; represses apoptosis	3.1	NPTX1	Neuronal pentraxin I	6.2
PTHlh	Parathyroid hormone-like hormone; oncogene	3.3	FOXF2	Forkhead box F2	4.9
ADAM12	A disintegrin and metallo-proteinase domain 12	2.7	PCDH11	Protocadherin 11 X-linked	6.1
EGR3	Early growth response-3; regulates FasL	2.7	EN2	Engrailed homolog 2	3.8
PDLIM2	PDZ and LIM domain 2	2.9	NPTXR	Neuronal pentraxin receptor	3.8
RTK			WT		
Symbol	Gene	Fold-change	Symbol	Gene	Fold-change
GCG	Glucagon	41.1	EYA1	Eyes absent homolog 1; upregulated in WT	9.6
NQO1	NAD(P)H dehydrogenase quinone 1	4.7	PAX2	Paired box gene 2; upregulated in WT	6.3
RSU1	Ras suppressor protein 1; suppresses Ras	5.7	TMEFF1	Transmembrane protein with EGF-like and two follistatin-like domains-1;	2.8
PLP2	Proteolipid protein 2	3.2	SCHIP1	Schwannomin interacting protein 1	3.7
FLJ22662	Hypothetical protein	3.9	MYCN	Viral related oncogene, neuroblastoma derived	3.6
APLP1	Amyloid beta(A4) precursor-like protein 1	2.6	DSP	Desmoplakin; upregulated in variety of tumors	5.0
CBX6	Chromobox homolog 6	2.2	GPR64	G protein-coupled receptor 64	7.1
NIT2	Nit protein 2	2.1	WASF3	WAS protein family, member 3	5.3

The fold-change is calculated as the ratio of the average expression of one type of tumor to the average expression of all the other three types.

accomplished using a single test. This compares with the non-standardized sarcoma work-up that currently requires a large number of tests, each associated with significant cost. Evidence of the reliability of the cancer signatures includes accurate prediction results using both unsupervised and supervised approaches, and the small number of genes necessary in order for each classifier to be accurate. In addition, the fact that our samples were collected from various hospitals at different times indicates that the gene signatures were preserved robustly regardless of sample preparation, collection variables, and other variations. The specific gene expression data generated by this study need to be considered in light of the different pathologic, clinical, and molecular features that characterize each malignant pediatric renal tumor, as briefly summarized below.

Cellular Mesoblastic Nephroma (Infantile Fibrosarcoma)

CMNs are round or spindle-cell lesions that occur almost exclusively in children under 2 years of age [6]. The t(12;15)(p13;q25) is identified in greater than 90% of CMN [7]. This translocation fuses the *ETV6* gene from chromosome 12p13 with the 15q25 neurotrophin-3 receptor gene *NTRK3* [8]. The presence of the same translocation in the histologically identical infantile fibrosarcoma provides support to the current consensus that CMNs represent intrarenal infantile fibrosarcoma [7]. Despite their potential malignancy, CMNs and infantile fibrosarcomas are treated by resection alone, without adjuvant chemotherapy, and are associated with a greater than 95% overall survival [9]. It is the experience of the Renal Tumor Pathology Center that occasional CMNs have prominent nucleoli and may perfectly mimic RTK. Conversely, RTKs may have small nuclei without prominent nucleoli and may mimic CMNs. It is therefore of interest that the hierarchical clustering in this study (Fig. 1A) revealed greater similarities in gene expression patterns between CMN and RTK than any of the other tumors. Despite this similarity, the classifiers were able to perfectly distinguish these tumors. Several genes of particular interest were identified within the CMN classifier (Table II). These include genes implicated in adult cancers (*SPRY4* and *EGFR3*) [10,11]. Of interest, *ADAM12* has been identified as upregulated in a category of related lesions, aggressive fibromatosis [12]. Our data suggest that CMNs or closely related tumors may exist that lack the characteristic fusion gene (as seen in test sample 1). This suggests that there are lesions within the spectrum of pediatric fibroblastic-myofibroblastic tumors that are difficult to categorize using current classification systems but have clinical, morphologic, and now gene expression overlap with other closely related lesions. The ability to classify such rare tumors as CMN using gene expression profiling supports the strength of this technology. However, caution must be exercised until classifiers have been developed for all the rare pediatric

neoplasms. It remains possible that test sample 1 is another type of primitive mesenchymal neoplasm. Until experience with a broad range of sarcomas has been developed, this will continue to be an unanswered question.

Rhabdoid Tumor of the Kidney

Rhabdoid tumors are round-cell tumors characterized pathologically by prominent nucleoli and whorls of cytoplasmic intermediate filaments [13,14]. Rhabdoid tumors are highly aggressive tumors of renal, central nervous system, and soft tissue sites in young children, with an associated overall survival of 25% [15,16]. All rhabdoid tumors appear to contain mutations or deletions that inactivate the *hSNF5/INI1* gene, whose role is to alter the conformation of the DNA-histone complex so that transcription factors have access to target genes [17]. A clinically applicable test for these genetic changes is not yet available, although recently an immunohistochemical stain for the INI-1 protein has demonstrated reliable negativity in RTKs and positivity in other pediatric renal tumors [18]. While our study focuses on genes that are overexpressed in one category of tumors compared with the other categories (due to the greater robustness that such genes provide for classification) we did note that *INI-1* was 2.3-fold underexpressed in RTKs compared with all other histologic types. Somewhat surprising is the presence of glucagon at the top of the classifier list with a greater than 40-fold increase over the other renal tumors. Glucagon is formed by the pancreas and, to a lesser extent, the gastrointestinal tract, and acts predominantly in the regulation of blood glucose levels. It is of interest that some studies have reported that genes in the glycolysis/glucogenesis pathway are upregulated in tumors that are more aggressive and have a poor prognosis. Gene expression profiles of central nervous system rhabdoid tumors and their comparison with other pediatric brain tumors has been reported using Affymetrix HG-U95 arrays [19]. There was no overlap between the 10 genes in the classifier reported for the CNS rhabdoid tumors and the top 20 genes that were identified in our classifier. This is likely because the tumors that were being compared to rhabdoid tumors in each study were quite different, therefore the differentially expressed genes would necessarily be quite different. Of note, we analyzed two extrarenal soft tissue rhabdoid tumors using the same techniques described in this paper, and they clustered cleanly with renal rhabdoid tumors and were confidently classified using the RTK classifier (data not shown).

Clear Cell Sarcoma of Kidney

CCSKs are round and spindled lesions of infancy virtually only recognized in the kidney [20]. Overall survival of patients with CCSKs of all stages is approximately 75% [21]. Despite many attempts to find immunohistochemical markers or recurrent genetic changes that may aid in the

diagnosis of CCSK, none have been identified [20,22]. CCSK remains the most frequently misdiagnosed pediatric renal tumor due to its highly variable histologic appearance and its ability to mimic all the other pediatric renal tumors. Comprehensive analysis of our gene expression data of CCSK has been performed and is reported separately [23]. Briefly, CCSKs demonstrate prominent expression of neural markers and genes involved in the related Akt and Sonic Hedgehog pathways. One measure of the validity of a set of statistically derived classifiers is the extent to which individual genetic markers are associated with recognized and relevant biologic functions. For CCSK, the prominent neural differentiation identified in our in-depth analysis was also represented in our CCSK classifiers. Neuronal pentraxin and its receptor are in the top genes of the classifier; *EN2* and *D4S234E* are also genes involved in neural activities. The increased expression of neural markers suggests that these tumors may be neural or neuroectodermal in origin. The top gene in the CCSK classifier, *FOXF1*, is a Sonic Hedgehog-dependent forkhead transcription factor [24].

Wilms Tumor

Nephroblastoma is a tumor that replicates the histology of the developing kidney. It is the most common genitourinary cancer in children, affecting approximately one in every 8,000 children [25,26]. Approximately 10% of nephroblastomas develop in association with one of several well-characterized dysmorphic syndromes associated with *WT-1* and *WT-2* genes located on 11p [27,28]. Despite the success in defining *WT-1* and *WT-2*, it is clear that additional genetic loci are responsible for the pathogenesis of most nephroblastomas. Recent studies that analyzed Wilms tumor using gene expression profiling found that Wilms tumors over-express genes that correspond to the earliest stage of metanephric development [29]. These include *PAX2* and *EYA1*, genes essential for survival and differentiation of the early metanephros. Of note, *PAX2* and *EYA1* represent the top two genes in our WT classifier. In addition, our classifier includes *TMEFF1*, which contains structural domains that suggest a role in growth factor signaling. However, this gene has also been suggested to function as a tumor suppressor in brain cancers [30].

In summary, this study supports the potential of gene expression profiling to contribute much greater diagnostic confidence for sarcomas when using renal sarcomas as a model. However, gene expression profiling is only able to accurately classify those tumor types that were included in the development of the classifiers. For example, profiling a synovial sarcoma may result in erroneous classification using the classifiers we report for WT, CMN, CCSK, and RTK. Therefore, pathologic analysis remains the important initial investigation in all cases, as this provides the differential diagnosis and ensures that gene expression profiling is being applied in the correct context. This study highlights the added

power and confidence that could be provided by comprehensive gene expression analysis of intermediate and malignant mesenchymal neoplasms. From the perspectives of both the institutional pathologist and the central review pathologist, as well as the treating clinicians, it is extremely difficult to be in the position of directing therapy for a child based on a diagnosis that lacks definitive criteria and is associated with less than 90% confidence. The combined approach of histopathologic and gene expression analysis allows for greater confidence and may allow the recognition and development of better therapeutic strategies for neoplasms that represent unusual variants of defined entities.

ACKNOWLEDGMENT

The authors acknowledge the excellent technical support provided by Chiler Hasan and Donna Kersey. The authors thank Yong Zeng for array data normalization. The authors also acknowledge the effort and support of pathologists, oncologists, radiologists, surgeons, radiation therapists, and clinical research assistants who have contributed the significant efforts required to develop the unique clinical resources comprising the National Wilms Tumor Study Group, led by Dr. Daniel Green and, at its completion, by Dr. Paul Grundy.

REFERENCES

1. Beckwith JB, Palmer NF. Histopathology and prognosis of Wilms tumors: Results from the First National Wilms' Tumor Study. *Cancer* 1978;41:1937–1948.
2. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. *Nat Biotechnol* 2003; 21:818–821.
3. Eisen MB, Spellman PT, Brown PO, et al. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;95:14863–14868.
4. Hedenfalk I, Duggan D, Chen Y, et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001;344:539–548.
5. Simon R, Radmacher MD, Dobbin K, et al. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;95:14–18.
6. Howell CG, Othersen HB, Kiviat NE, et al. Therapy and outcome in 51 children with mesoblastic nephroma: A report of the National Wilms' Tumor Study. *J Pediatr Surg* 1982;17:826–831.
7. Knezevich SR, Garnett MJ, Pysher TJ, et al. ETV6-NTRK3 gene fusions and trisomy 11 establish a histogenetic link between mesoblastic nephroma and congenital fibrosarcoma. *Cancer Res* 1998;58:5046–5048.
8. Knezevich SR, McFadden DE, Tao W, et al. A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nat Genet* 1998;18:184–187.
9. Fitchev P, Beckwith JB, Perlman EJ. Congenital mesoblastic nephroma: Prognostic factors. *Lab Invest* 2003;83:2.
10. Cabrita MA, Christofori G. Sprouty proteins: Antagonists of endothelial cell signaling and more. *Thromb Haemost* 2003; 90:586–590.
11. Inoue A, Omoto Y, Yamaguchi Y, et al. Transcription factor EGR3 is involved in the estrogen-signaling pathway in breast cancer cells. *J Mol Endocrinol* 2004;32:649–661.

12. Skubitz KM, Skubitz AP. Gene expression in aggressive fibromatosis. *J Lab Clin Med* 2004;143:89–98.
13. Beckwith JB. Renal tumors. In: Stocker JT, Askin FB, editors. *Pathology of solid tumors in children*. New York: Chapman & Hall Medical; 1998: pp 1–23.
14. Bonnin JM, Rubinstein LJ, Palmer NF, et al. The association of embryonal tumors originating in the kidney and in the brain. A report of seven cases. *Cancer* 1984;54:2137–2146.
15. Vujanic GM, Sandstedt B, Harms D, et al. Rhabdoid tumour of the kidney: A clinicopathological study of 22 patients from the International Society of Paediatric Oncology (SIOP) nephroblastoma file. *Histopathology* 1996;28:333–340.
16. Weeks DA, Beckwith JB, Mierau GW, et al. Renal neoplasms mimicking rhabdoid tumor of kidney. A report from the National Wilms' Tumor Study Pathology Center. *Am J Surg Pathol* 1991;15:1042–1054.
17. Verstege I, Sevenet N, Lange J, et al. Truncating mutations of hSNF5/INI1 in aggressive pediatric cancer. *Nature* 1998;394:203–206.
18. Hoot AC, Russo P, Judkins AR, et al. Immunohistochemical analysis of hSNF5/INI1 distinguishes renal and extra-renal malignant rhabdoid tumors from other pediatric soft tissue tumors. *Am J Surg Pathol* 2004;28:1485–1491.
19. Pomeroy SL, Tamayo P, Gaasenbeek M, et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 2002;415:436–442.
20. Argani P, Perlman EJ, Breslow NE, et al. Clear cell sarcoma of the kidney: A review of 351 cases from the National Wilms' Tumor Study Group Pathology Center. *Am J Surg Pathol* 2000;24:4–18.
21. Seibel NL, Li S, Breslow NE, et al. Effect of duration of treatment on treatment outcome for patients with clear-cell sarcoma of the kidney: A report from the National Wilms' Tumor Study Group. *J Clin Oncol* 2004;22:468–473.
22. Argani P, Fritsch M, Kadkol SS, et al. Detection of the ETV6-NTRK3 chimeric RNA of infantile fibrosarcoma/cellular congenital mesoblastic nephroma in paraffin-embedded tissue: Application to challenging pediatric renal stromal tumors. *Mod Pathol* 2000;13:29–36.
23. Cutcliffe C, Kersey D, Huang CC, et al. Clear cell sarcoma of the kidney: Upregulation of neural markers with activation of the sonic hedgehog and Akt pathways. *Clin Cancer Res* 2005;11:7986–7990.
24. Mahlapuu M, Enerback S, Carlsson P. Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations. *Development* 2001;128:2397–2406.
25. Breslow N, Olshan A, Beckwith JB, et al. Epidemiology of Wilms tumor. *Med Pediatr Oncol* 1993;21:172–181.
26. Webber BL, Parham DM, Drake LG, et al. Renal tumors in childhood. *Pathol Annu* 1992;27:191–232.
27. Coppes MJ, Haber DA, Grundy PE. Genetic events in the development of Wilms' tumor. *N Engl J Med* 1994;331:586–590.
28. Green DM, Breslow NE, Beckwith JB, et al. Screening of children with hemihypertrophy, aniridia, and Beckwith-Wiedemann syndrome in patients with Wilms tumor: A report from the National Wilms' Tumor Study. *Med Pediatr Oncol* 1993;21:188–192.
29. Li CM, Guo M, Borczuk A, et al. Gene expression in Wilms' tumor mimics the earliest committed stage in the metanephric mesenchymal-epithelial transition. *Am J Pathol* 2002;160:2181–2190.
30. Gery S, Yin D, Xie D, et al. TMEFF1 and brain tumors. *Oncogene* 2003;22:2723–2727.

Predicting Relapse in Favorable Histology Wilms Tumor Using Gene Expression Analysis: A Report from the Renal Tumor Committee of the Children's Oncology Group

Chiang-Ching Huang,¹ Samantha Gadd,² Norman Breslow,³ Colleen Cutcliffe,² Simone T. Sredni,² Irene B. Helenowski,¹ Jeffrey S. Dome,⁴ Paul E. Grundy,⁵ Daniel M. Green,⁶ Michael K. Fritsch,⁷ and Elizabeth J. Perlman²

Abstract **Purpose:** The past two decades has seen significant improvement in the overall survival of patients with favorable histology Wilms tumor (FHWT); however, this progress has reached a plateau. Further improvements may rely on the ability to better stratify patients by risk of relapse. This study determines the feasibility and potential clinical utility of classifiers of relapse based on global gene expression analysis.

Experimental Design: Two hundred fifty FHWT of all stages enriched for relapses treated on National Wilms Tumor Study-5 passed quality variables and were suitable for analysis using oligonucleotide arrays. Relapse risk stratification used support vector machine; 2- and 10-fold cross-validations were applied.

Results: The number of genes associated with relapse was less than that predicted by chance alone for 106 patients (32 relapses) with stages I and II FHWT treated with chemotherapy, and no further analyses were done. This number was greater than expected by chance for 76 local stage III patients. Cross-validation including an additional 68 local stage III patients (total 144 patients, 53 relapses) showed that classifiers for relapse composed of 50 genes were associated with a median sensitivity of 47% and specificity of 70%.

Conclusions: This study shows the feasibility and modest accuracy of stratifying local stage III FHWT using a classifier of <50 genes. Validation using an independent patient population is needed. Analysis of genes differentially expressed in relapse patients revealed apoptosis, Wnt signaling, insulin-like growth factor pathway, and epigenetic modification to be mechanisms important in relapse. Potential therapeutic targets include FRAP/MTOR and CD40.

Wilms tumor is the most common urogenital malignancy in children, with ~500 new cases per year in North America. Several national and international cooperative group clinical trials have optimized the therapy resulting in

an increase in the overall survival rate to ~90%. The current therapeutic approach for Wilms tumor is based on histologic subtype (favorable versus unfavorable histology) and tumor stage (1). The majority of Wilms tumor has favorable histology, defined as the absence of anaplasia, and these represent the focus of the current study. Patients with anaplasia are treated differently than those with favorable histology Wilms tumor (FHWT) and are beyond the scope of this study.

In recent years, the improvement in relapse-free and overall survival for FHWT at each stage has reached a plateau. Some patients are not successfully treated initially, resulting in relapse and less frequently death. Of equal importance, many patients may receive more therapy than needed; this is particularly true for patients with stage III disease (2, 3). Further improvements in outcome will rely in part on the ability to identify markers associated with relapse, with the hope of better stratifying patients. This goal represented a major focus of the National Wilms Tumor Study-5 clinical protocol, which included a large-scale effort aimed at tumor banking and molecular analysis. These efforts showed that loss of heterozygosity (LOH) for both chromosomes 1p and 16q was associated with poor outcome (4). However, LOH is able to detect only a very small subset of FHWT patients who have an increased risk of relapse and death. Additional efforts are therefore required to further define markers of relapse. In this study, we analyzed gene expression

Authors' Affiliations: Departments of ¹Preventive Medicine and ²Pathology, Northwestern University's Feinberg School of Medicine and Robert H. Lurie Cancer Center, Chicago, Illinois; ³Department of Biostatistics, University of Washington and Fred Hutchinson Cancer Research Center, Seattle, Washington; ⁴Division of Oncology, Children's National Medical Center, Washington, District of Columbia; ⁵Departments of Pediatrics and Oncology, Cross Cancer Institute and University of Alberta, Edmonton, Alberta, Canada; ⁶Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee; and ⁷Department of Pathology, University of Wisconsin-Madison, Madison, Wisconsin

Received 5/2/08; revised 7/31/08; accepted 10/12/08; published OnlineFirst 02/10/2009.

Grant support: NIH grants U10CA42326 (N. Breslow, D.M. Green, and E.J. Perlman), U10CA98543 (J.S. Dome, P.E. Grundy, and E.J. Perlman), and U01CA88131 (E.J. Perlman).

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.

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

Requests for reprints: Elizabeth J. Perlman, Children's Memorial Hospital, 2300 Children's Plaza, Box 17, Chicago, IL 60614. E-mail: EPerlman@childrensmemorial.org.

© 2009 American Association for Cancer Research.
doi:10.1158/1078-0432.CCR-08-1030

Translational Relevance

This article evaluates gene expression signatures to predict relapse in patients registered on the National Wilms Tumor Study-5 cooperative group protocol using stage and treatment-specific analyses. This will enable independent validation using samples from patients registered in the ongoing Children's Oncology Group protocols. Successful signatures will be able to be used for therapeutic stratification during protocols estimated to open in 2012. Signatures with 50 genes were associated with relapse in stage III tumors (sensitivity of 47% and specificity of 70%). Existing markers for relapse currently used for stratification (1p and 16q loss of heterozygosity) have a sensitivity of 8% and specificity of 96%. Analysis of specific genes associated with relapse revealed apoptosis, Wnt signaling, and the insulin-like growth factor pathway to be important. These pathways will be separately validated at the protein level within the current protocol. Importantly, all the above-identified pathways have been previously targeted for developmental therapies in the current literature. Two additional potential therapeutic targets identified in this study were FRAP/MTOR and CD40. MTOR inhibitors and anti-CD40 antibodies are currently used in clinical trials. The Children's Oncology Group Renal Tumor Committee has an active Developmental Therapeutics Subcommittee that is able to bring such promising agents forward.

patterns to identify such markers and to investigate the feasibility of developing classifiers able to predict patients at high risk for relapse.

Materials and Methods

Patient samples. On June 25, 2003, all 1,451 patients with FHWT registered on National Wilms Tumor Study-5 from August 1995 to June 2002 who had available pretreatment tumor tissues were identified. Six hundred patients, consisting of all those known to have relapsed and a random sample including ~30% of the remainder, were randomly divided into two groups of 300 patients each (groups A and B), both enriched for relapse.

The National Wilms Tumor Study-5 protocol was approved by the review boards of institutions that registered patients. Parents or guardians signed informed consent for collection and banking of biological samples. Histologic diagnosis and local stage were confirmed by central review. Patients who did not receive the therapy consistent with the diagnosis and stage found on central review were rejected from the analysis. Chest X-ray, abdominal ultrasound, and computed tomography studies were required for staging. Tumor samples were obtained at initial diagnosis before the initiation of therapy and were snap frozen. A frozen section was evaluated to confirm at least 80% viable tumor cellularity.

Gene expression analysis. RNA was extracted and hybridized to Affymetrix U133A arrays, scanned, and subjected to quality-control variables according to the previously described protocol (5).

Quantitative reverse transcription-PCR. RNA levels of five genes showing a range of fold changes and overall expression levels were further analyzed by quantitative reverse transcription-PCR using TaqMan Gold and the ABI Prism 7700 Sequence Detection System (Applied Biosystems). β -Actin was used as the endogenous control.

Primer and probe sequences are provided in Supplementary Table S1.⁸ Each threshold cycle (C_T) was determined and the C_T for β -actin was subtracted from this for normalization.

Data analysis. Positional-dependent nearest-neighbor model software⁹ was used to translate the scanned images into expression analysis files and to normalize the data across all arrays (6). Genes with maximum expression less than a log scale of 6 across all tumors and Affymetrix control genes were excluded, resulting in 20,931 probe sets for analysis. Support vector machine was chosen for relapse risk stratification (7). The following procedure was used to construct 150 different classifiers for each randomly drawn training set of tumors. For each gene, the t statistic comparing expression between cases and controls and the associated P value were calculated using the Welch method. The K genes (1-150 genes) with the smallest P values were selected to construct a n support vector machine model as developed by Chang and Lin¹⁰ and implemented in a R software package, e1071 (7). This model was then applied to categorize tumors within the complementary test set as having high or low risk for relapse.

Validation procedures. Cross-validations (2- and 10-fold) were used to investigate the ability of classifiers established in a training set to predict relapse in an independent test set. For 2-fold cross-validation, the data set was randomly divided 500 times into training and corresponding test sets of equal size, each including half the patients who relapsed. A classifier for relapse was identified for each training set and used to assign tumors in the corresponding test set to low-risk and high-risk categories. The training and test sets were then swapped. The number of top K genes in each classifier evaluated ranged from 1 to 150. Therefore, 150,000 different classifiers were developed, one for each value of K from 1 to 150 for each of the 1,000 (500 \times 2) training sets. For 10-fold cross-validation, the data set was randomly divided 500 times into 10 groups of approximately equal size. Each group included approximately the same number of relapses. For each such group, a classifier was built with the remaining 9 of 10 of the samples and then used to categorize tumors in the group as low or high risk; the process was repeated until all tumor samples were categorized as low or high risk. For all the cross-validation procedures, to avoid gene-selection bias, classifiers were completely rebuilt in each cross-validation iteration (8).

Results

Patient information and quality control

Group A. Ninety of the 300 tumors from group A were rejected. Following verification of the stage, relapse status, clinical follow-up, and diagnosis, 1 tumor (with diffuse hyperplastic perilobar nephroblastomatosis) was excluded for diagnosis, 14 tumors with clinical follow-up <3 years were excluded, 29 tumors were rejected due to <80% viable tumor cellularity, 34 tumors were rejected due to A_{260}/A_{280} ratios < 1.8, and 12 tumors failed to meet Affymetrix-recommended hybridization quality-control standards. Illustrated in Table 1 is the stage and relapse status of group A. The tumors were categorized based on local (abdominal) and overall stage. Stage I tumors (49) were confined to the kidney. Stage II tumors (73) showed infiltration beyond the kidney with negative surgical margins and regional lymph nodes; 6 of these patients had distant metastases (local stage II, overall stage IV). Of 76 local stage III patients, 12 presented with distant metastasis (local stage III, overall stage IV). Six patients presented with bilateral

⁸ Supplementary tables can be found at <http://www.childrensmrc.org/perlman/>.

⁹ <http://odin.mdacc.tmc.edu/~zhangli/PerfectMatch/>

¹⁰ <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

Table 1. Categorization of 300 samples in group A by stage

Local stage/overall stage	No. (relapses)	Rejected (relapses)	Final (relapses)
I/I	69 (10)	20 (0)	49 (10)
II/II	103 (30)	36 (5)	67 (25)
II/IV	7 (1)	1 (0)	6 (1)
III/III	85 (22)	21 (5)	64 (17)
III/IV	16 (9)	4 (2)	12 (7)
V	9 (4)	3 (1)	6 (3)
Biopsy only	11 (6)	5 (3)	6 (3)
Total	300 (82)	90 (16)	210 (66)

disease (stage V) and 6 unilateral tumors were biopsied and treated before the subsequent removal of the kidney following therapy ("biopsy only").

This study grouped together for analysis patients who received the same therapy. First, there were 10 stage I patients (3 relapses) registered on a treatment arm within National Wilms Tumor Study-5 that included patients ages ≤ 24 months with stage I disease whose kidney and tumor had a combined weight of ≤ 550 g; these patients were treated without initial chemotherapy as reported previously (9). These patients are not further considered in this particular study due to their small numbers. The remaining patients with stage I and those with overall stage II disease were treated with nephrectomy, vincristine, and dactinomycin (106 patients, 32 relapses); they formed a second group referred to below as stage I/II tumors. The third group, referred to collectively as stage III tumors, includes those patients with local and overall stage III disease (64 patients) and those with local stage III, overall stage IV disease (12 patients). Careful consideration was given to the inclusion of patients with distant metastasis (stage IV) to this group due to their potential biological differences. Because the goal of the study was to define clinically applicable signatures, it was considered to be optimal to group all patients receiving the same therapy. Patients with stages III/III and III/IV disease both receive nephrectomy, vincristine, actinomycin, doxorubicin, and abdominal radiation therapy. In addition, both stages III and IV tumors have known or presumed residual disease, and one of the most significant causes for relapse is resistance to chemotherapy, which does not change with stage. To ensure that the stage IV patients did not significantly alter the results, all analyses of the stage III tumors were also performed, leaving the patients with overall stage IV disease out. Six patients with local stage II, overall stage IV disease were excluded from the stage I/II group because they received additional adjuvant therapy; they were also excluded from the stage III group because they did not receive abdominal radiation. Six tumors biopsied and treated before removal of the kidney were not included nor were 6 patients with bilateral disease (stage V) due to the individualized nature of their therapy. In total, gene expression analysis from 182 tumors from group A was used in these analyses.

Group B. The only patients analyzed from group B were the 99 appropriately diagnosed and treated stages III/III and III/IV tumors as defined above. The same quality-control variables outlined above were applied. Nine tumors were rejected following histology review, 5 had < 3 years of follow-up, 8 had

poor RNA quality, and 9 failed to meet quality-control variables following hybridization, leaving 68 tumors for analysis. Clinical stage and outcome for all 144 stage III tumors within the entire group of 600 tumors are provided in Table 2.

Classifiers that identify relapse risk: initial feasibility assessment

Group A was used to assess the feasibility of developing classifiers that predict risk of relapse within the above groups. To accomplish this, the expression of each probe set in the tumors that relapsed was compared with the tumors that did not relapse using the *t* test, and a nominal *P* value was established for this association. The distribution of *P* values within the stage I/II tumors failed to show more probe sets with *P* < 0.05 than would be expected by chance alone, indicating that further investigation would be associated with high false discovery rates. Similar results were found with the separate analysis of stage II tumors alone. The distribution of *P* values for stage III tumors, however, showed a higher frequency of probe sets at low *P* values than would be expected by chance alone (1,445 probe sets with *P* < 0.05). This analysis was repeated excluding stage IV patients with similar results. This indicates that a gene signature associated with relapse may exist. The remainder of the study therefore focused on the stage III tumors.

Experimental verification

Five genes with *P* values < 0.05 were analyzed using quantitative reverse transcription-PCR. These included *COL2A1*, *DNMT1*, *ENPP2*, *GCNT*, and *LY6G6D*. In all cases, the gene expression data from array analysis and from quantitative reverse transcription-PCR were comparable (data not shown). Replicate analysis beginning with RNA extraction done on a subgroup of 10 tumors confirmed reproducibility of the expression patterns.

Table 2. Clinical characteristics of 144 stage III tumors

	Cases (n = 53)	Controls (n = 91)	Total
Stage			
III/III	37	77	114
III/IV	16	14	30
Reason for stage III			
Positive margin or rupture/spillage	15	36	51
Lymph node involvement	19	41	60
Positive margin and lymph node involvement	19	14	33
Median (range) time to relapse (d)	284.5 (76-1940)		
Median (range) disease-free follow-up (d)		2,162 (1,124-3,788)	
Site of relapse			
Lung only	26		
Abdomen	13		
Liver	5		
Other	9		
LOH for 1p and 16q	3 of 53	5 of 89	8 of 142

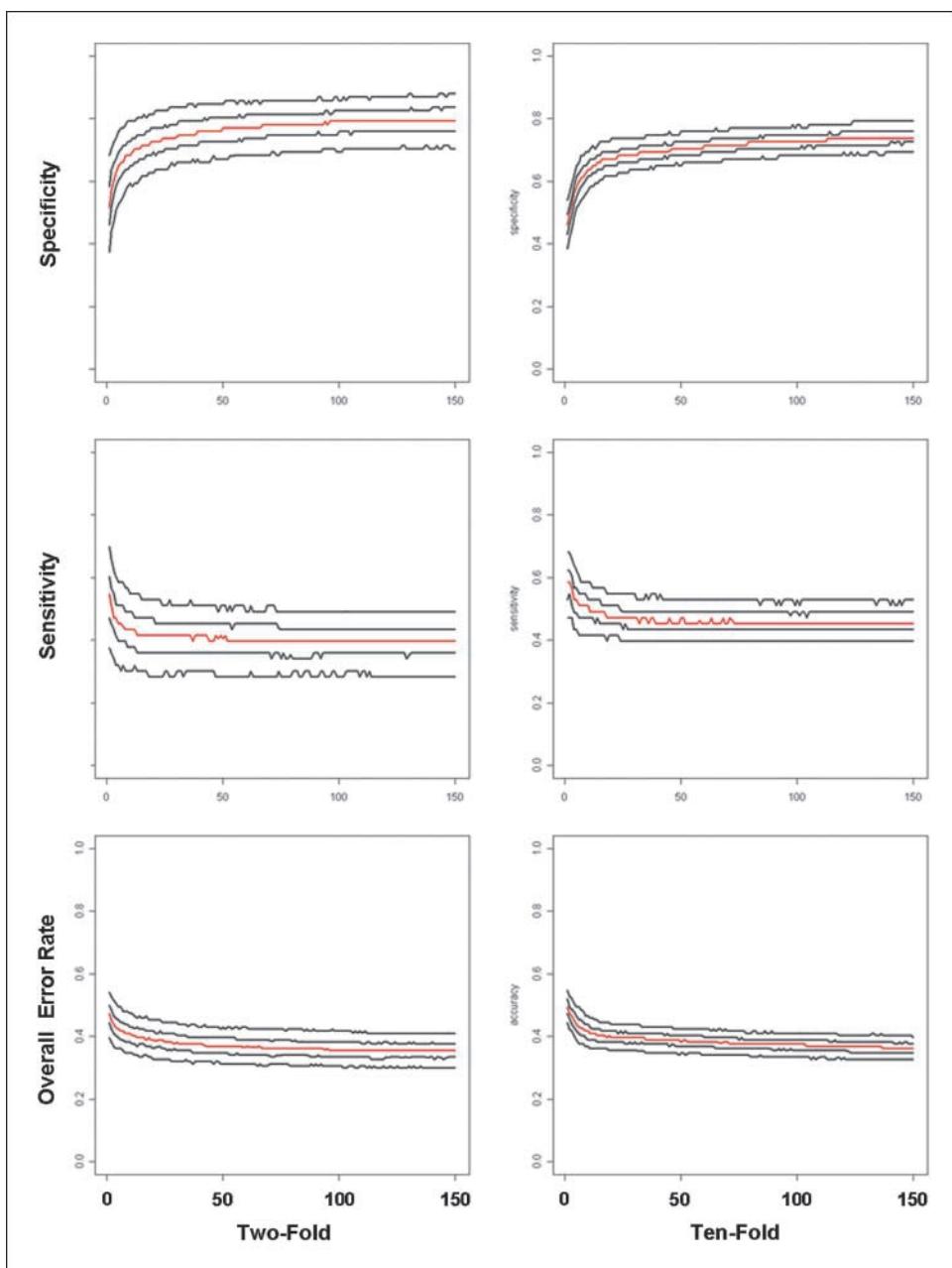


Fig. 1. Cross-validation with 144 stage III FHWT: all 144 stage III FHWT were randomly divided into 2 and 10 groups. For each training set composed of all but one group, a classifier was developed using from 1 to 150 genes; each classifier was then applied to the group withheld to define low-risk and high-risk tumors. This was repeated until all groups were categorized. This entire process was repeated 500 times. *Left* and *right*, results of the 2- and 10-fold cross-validations, respectively. For each, five curves are illustrated representing the 5th, 25th, 50th, 75th, and 95th percentiles. *A*, specificity: among tumors that did not relapse, the percentage categorized using gene expression as low risk for relapse. *B*, sensitivity: among tumors that relapsed, the percentage categorized as high risk for relapse. *C*, error rate/accuracy: percentage of tumors incorrectly classified.

Developing and validating classifiers that predict relapse in stage III FHWT

The original design was to identify classifiers based on the association of gene expression with time to relapse using the 300 tumors in group A and to then apply these classifiers in a blinded fashion to tumors in group B. Following this design, a classifier was developed from stage III tumors in group A using the association between gene expression and time to relapse to select the top genes. This classifier was applied to stage III tumors within group B in a blinded fashion. Six of 10 tumors placed in the high-risk category relapsed and 23 of 63 tumors placed in the low-risk category relapsed, resulting in an odds ratio measure of association of 2.6 ($P = 0.18$, Fisher's exact test). Due to concerns regarding the small number of relapses and the use of time to relapse in addition to fact of relapse to select

genes, we also analyzed the results using a case-control design including all stage III patients. Three years had elapsed during the laboratory analysis, allowing for extended follow-up for the 600 patients. Lastly, it is increasingly recognized that small sample sizes may hinder the reliability of gene classifiers, particularly those for which outcome prediction is the goal (10). Therefore, for the remaining analyses, a case-control design was applied. All stage III patients within both groups A and B who were known to have relapsed as of February 2007 were defined as cases; those who had not relapsed and who had been followed >3 years were defined as controls. The risk of relapse after 3 years is low, with only 3 of 213 (1.4%) patients relapsing beyond 3 years in National Wilms Tumor Study-5 (4).

The 144 stage III tumors from groups A and B include 53 relapses (cases) and 91 nonrelapses (controls; Table 2). The

distribution of P values for all 144 tumors again showed a higher frequency of probe sets at low P values than would be expected by chance alone (2896 probe sets with $P < 0.05$). (When the stage IV tumors were left out, 2,259 probe sets had a $P < 0.05$.) Cross-validations (2- and 10-fold) were repeated 500 times, varying the number of genes in the classifiers from 1 to 150. The specificity, sensitivity, and overall error rate were determined for each classifier and are illustrated in Fig. 1. The error rate improved markedly when the number of genes in the classifier increased from 5 to 50 but only very slowly with increasing number of genes thereafter. With 10-fold cross-validation, when 50 genes were used in the classifier, the median sensitivity was 47% (range, 35–58%), the median specificity was 70% (range, 60–78%), and the median overall error rate was 38% (range, 31–45.8%) in this mix of cases and controls. The 10-fold cross-validation analysis was repeated leaving the stage IV patients out, and the median sensitivity was 35% and the median specificity was 78%.

Performance of classifiers within individual tumors

To evaluate the robustness of the classifiers when applied to individual tumor samples, we observed the 10-fold cross-validation analyses, selecting only the 50-gene classifiers. Every tumor was assessed 500 times for relapse risk using a 50-gene classifier and the success of every classification was tracked for each individual tumor. The percentage of classifiers that categorized each tumor as high risk is illustrated in Fig. 2. Fifteen (28%) relapsed tumors were consistently and correctly classified as high risk using a stringent cut-point requiring >90% of the classifiers to predict relapse (*top dashed line*). Similarly, 47 (52%) nonrelapsed tumors were consistently and correctly classified as low risk using a cut-point requiring <10% of the classifiers to predict relapse (*bottom dashed line*). However, a large number of tumors fell in between these cut-points, and 18 (34%) of the tumors that relapsed were

Table 3. Categorization of selected genes of interest

		Chromosomal location	Unigene ID
Metastasis and tumorigenesis			
CD40	Up	chr20q12-q13.2	Hs.118832
CDC42BPA	Up	chr1q42.11	Hs.326691
CHC1	Up	chr1p36.1	Hs.84746
EWSR1	Up	chr22q12.2	Hs.129953
FGF18	Up	chr5q34	Hs.49585
FRAP1/MTOR	Up	chr1p36.2	Hs.250723
IGF-I	Down	chr12q22-q23	Hs.85112
Ly6D	Up	chr8q24-qter	Hs.3185
MARK2	Up	chr11q12-q13	Hs.157199
MMP26	Up	chr11p15	Hs.204732
PDE3A	Up	chr12p12	Hs.777
POSTN	Down	chr13q13.3	Hs.136348
PPFIA1	Up	chr11q13.3	Hs.183648
PTGER3	Up	chr1p31.2	Hs.170917
PTPRH	Up	chr19q13.4	Hs.179770
TRIM17	Up	chr1q42	Hs.121748
UBE4B	Up	chr1p36.3	Hs.24594
UQCRH	Up	chr9q31	Hs.152601
Tumor suppressor			
DIABLO	Down	chr12q24.31	Hs.169611
PEG3	Down	chr19q13.4	Hs.139033
RECK	Down	chr9p13-p12	Hs.29640
RECQL	Down	chr12p12	Hs.235069
TUSC3	Down	chr8p22	Hs.71119.0

NOTE: Among the top 109 genes (sorted by P value and provided in Supplementary Table S2), approximately one-quarter were selected as genes of interest based on their known genetic and molecular functions.

consistently but incorrectly placed in the low-risk group even when using the least stringent cut-point of 10%.

Genes associated with relapse in stage III FHWT

Using the 144 stage III tumors, 109 genes were identified with $P < 0.001$ and are listed in Supplementary Table S2¹ (false discovery rate of 11.6%; ref. 11). The entire gene expression data and description of the experiment using the MIAME format are available.¹¹ Of the 20,931 genes analyzed, a total of 1,626 genes were represented in at least 1 of the 5,000 50-gene classifiers used in the 10-fold cross-validation: 1,004 genes were selected <10 times, 579 genes were selected >50 times, 244 genes were selected >500 times, and 55 genes were selected >1,000 times. The number of times each of the top 100 genes was used and its ranking within the 5,000 50-gene classifiers is provided within Supplementary Table S2. The P value correlated highly with the number of times each gene was selected.

Among the top 109 genes with $P < 0.001$, approximately one-quarter are known to function in metastasis, tumorigenesis, tumor progression, or tumor suppression; these are listed in Table 3. The chromosomal location of the 109 top genes (provided in Supplementary Table S2) was also analyzed to determine consistent areas of gain or loss of expression. No genes located on chromosomes 1p or 16q were in the top gene list and down-regulated. However, 16 of the top genes were

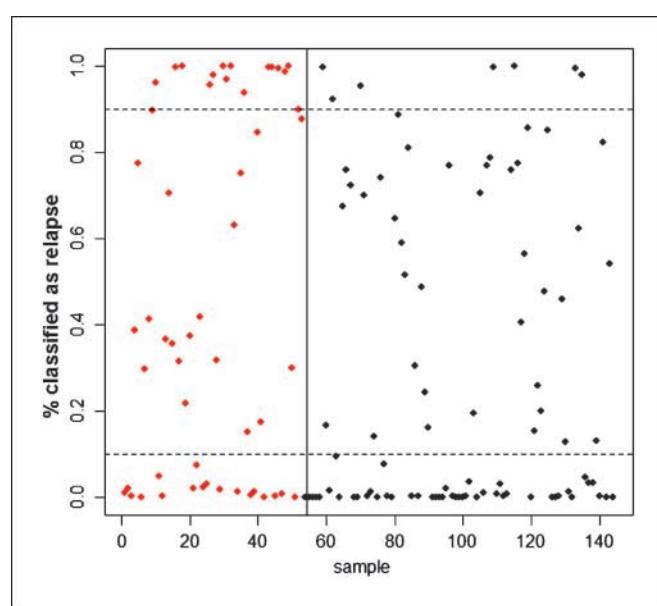


Fig. 2. Success of the 10-fold cross-validation method using 50-gene classifiers for each individual tumor. Y-axis, percentage of the 500 50-gene classifiers that categorized each tumor as high risk for relapse; X axis, arbitrary tumor number, with relapses (red dots to the left of the vertical line) having low numbers and nonrelapses having higher numbers (black dots).

¹¹ <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=pnknysekesyuly&acc=GSE10320>

located in the region 1p31-1pter and were up-regulated and 6 genes were located on distal 1q and were up-regulated in relapsed patients. These include CHC1, FRAP/MTOR, PTGER, and UBE4B from 1p and CDC42BPA and TRIM17 from 1q (Table 3).

Discussion

We report the gene expression analysis of a large group of FHWT consistently treated within a cooperative group setting. The goals were to determine the feasibility of using gene expression classifiers to identify patients who could benefit from increased or decreased adjuvant chemotherapy and to identify specific biological markers associated with relapse in patients with FHWT.

Ability to predict relapse in FHWT using gene expression varies with tumor stage. Efforts to establish an association between gene expression and relapse were not successful for stages I and II tumors yet were moderately successful for stage III tumors. The clinical endpoint of relapse in FHWT relies on two distinctly different features, the tumor's invasiveness (the ability to achieve lymphovascular access and distant growth, resulting in occult or overt residual disease) and the responsiveness of any residual disease to the therapy provided. The outcome of stage I FHWT <550 g in patients <24 months would suggest that most low-stage FHWT do not achieve lymphovascular access and are able to be completely treated by surgery alone. The majority of FHWT at all stages are exceedingly responsive to chemotherapy. As a result, our failure to detect gene expression signatures correlating with relapse in stage I to II patients treated with chemotherapy is not surprising. The majority of such tumors that are biologically "high risk" due to unresponsiveness to therapy are completely resected and the majority of those that are not completely resected (due to occult metastases, for example) are responsive to therapy. Conversely, stage III tumors by definition have evidence of residual disease. Therefore, the gene expression profile of the residual disease would be expected to correlate with responsiveness to therapy. It should be noted that the current study analyzes a random sample of the tumor, which is only an imperfect approximation of the residual disease itself.

Best classifiers in stage III FHWT will only be able to detect a subset of high-risk tumors. Our analysis indicates that not all stage III tumors that will go on to relapse can be correctly and consistently classified as high risk by global gene expression using Affymetrix U133A arrays. There are several possible explanations, all of which likely play a role:

- Global gene expression analysis performs well when detecting large differences in genes expressed at moderate to high levels. It is more difficult to detect reliable expression differences with genes that are tightly regulated with small differences in expression or genes that are expressed and regulated at very low levels.
- Drug resistance may develop during the administration of chemotherapy and the causes of this may not be detectable in the original sample.
- Clonal evolution may result in a more invasive or less responsive tumor clone. If this clone is not sampled, the tumor will not be correctly classified.

- Significant heterogeneity for the intrinsic causes of relapse likely exists, requiring a very large number of samples to define each cause.
- Wilms tumors have a wide spectrum of histologic appearances and varied lines of differentiation. The resulting spectrum of gene expression may obscure the identification of genes specifically associated with relapse.
- Some stage III tumors will have margins that are technically positive, but with no residual viable disease, or will have positive regional lymph nodes, with completely resected tumor. Therefore, it is expected that some stage III tumors would be identified based on gene expression as high risk that would not relapse.

Clinical utility of classifiers predicting relapse in stage III FHWT. In order for a biological marker to be clinically useful, it must be sufficiently robust and must have acceptable sensitivity and specificity to address the particular clinical needs of the patients being addressed. In this study, the classifiers developed using 50 genes resulted in a median sensitivity of 47% (25 of 53) and specificity of 70% (64 of 91). This compares with a sensitivity of 8% and specificity of 96% for LOH of both 1p and 16q in stages III and IV FHWT (4). To put this information into further perspective, the sensitivity and specificity of anaplasia for predicting death in Wilms tumor are 30% and 96%, respectively (12). Although the data suggest that the reliability of classifiers based on gene expression will be limited, strategies can be envisioned that propose either increased adjuvant therapy for those tumors at high risk or decreased adjuvant therapy for those tumors at low risk.

This study represents the first of a multistep process. The second step is the identification of a specific set of genes and the establishment of a mathematical rule that will be optimal for prospective prediction of individual tumors. This gene selection process may use the sensitivity and specificity of each individual gene, the overall absolute expression levels and differences in median expression levels (variables that take into account robustness of expression), the biological functions of the genes, and the interactions of expression within the top genes. The final and most critical step is the validation of the proposed classifier when applied in a blinded fashion to an independent patient population. This process is not currently possible and will require the analysis of patients in the current therapeutic protocols. In the absence of an ability to perform an independent validation currently, the proposal of a specific classifier may result in erroneous and premature conclusions. Therefore, we have chosen not to propose a specific classifier.

Comparison of gene expression data with prior publications. Three publications have reported differences in gene expression associated with clinical outcome in Wilms tumor (13–15). Zirn et al. analyzed 67 samples of all stages and of favorable and unfavorable histology collected following preoperative chemotherapy (13). Of their top genes associated with relapse, 3 were concordantly regulated and found in our Supplementary Table S2: IL11RA, POSTN, and SMPDL3A. IL11RA and SMPDL3A have been shown to be up-regulated in common adult cancers (16, 17). Both genes were down-regulated in FHWT that relapsed. POSTN, periostin, is up-regulated in many cancers and may activate the Akt/protein kinase B pathway and promote anchorage-independent growth (18, 19). Conversely,

Table 4. Genes concordantly differentially expressed within reported genetic regions of gain or loss in relapses

Probe set ID	Unigene	Gene symbol	Chromosomal location	P
Gains at 1q				
202337_at	Hs.94446	PMF1	chr1q12	0.0095375
214197_s_at	Hs.20991	SETDB1	chr1q21	0.0016913
217778_at	Hs.7854	SLC39A1	chr1q21	0.0044798
212147_at	Hs.4990	SMG5	chr1q21.2	0.0024339
208100_x_at	Hs.516316	SEMA6C	chr1q21.2	0.0029472
202596_at	Hs.111680	ENSA	chr1q21.2	0.0051370
217007_s_at	Hs.92208	ADAM15	chr1q21.3	0.0029510
219373_at	Hs.110477	DPM3	chr1q22	0.0049546
219125_s_at	Hs.292154	RAG1AP1	chr1q22	0.0085894
205789_at	Hs.1799	CD1D	chr1q22-q23	0.0023229
216885_s_at	Hs.110707	WDR42A	chr1q22-q23	0.0096002
205253_at	Hs.155691	PBX1	chr1q23	0.0079297
217797_at	Hs.177507	Ufc1	chr1q23.3	0.0063011
201482_at	Hs.77266	QSCN6	chr1q24	0.0011956
207350_s_at	Hs.102664	VAMP4	chr1q24-q25	0.0017361
215109_at	Hs.127338	KIAA0492	chr1q25.1	0.0074330
218253_s_at	Hs.274151	LGTN	chr1q31-q32	0.000072
208270_s_at	Hs.283667	RNPEP	chr1q32	0.0078793
212165_at	Hs.17481	TMEM183A	chr1q32.1	0.0064554
202186_x_at	Hs.155079	PPP2R5A	chr1q32.2-q32.3	0.0048414
212922_s_at	Hs.66170	SMYD2	chr1q41	0.0082410
220406_at	Hs.169300.1	TGFBI2	chr1q41	0.000893
207025_at	Hs.100072	GJA12	chr1q41-q42	0.0042464
220279_at	Hs.121748	TRIM17	chr1q42	0.0000805
213595_s_at	Hs.326691	CDC42BPA	chr1q42.11	0.000582
215583_at	Hs.287529	TMEM63A	chr1q42.12	0.0042900
221582_at	Hs.28777	HIST3H2A	chr1q42.13	0.0000185
212155_at	Hs.144904	RNF187	chr1q42.13	0.0052750
205689_at	Hs.31438	PCNXL2	chr1q42.2	0.000767
208934_s_at	Hs.4082	LGALS8	chr1q42-q43	0.0063052
203581_at	Hs.119007	RAB4A	chr1q42-q43	0.0074064
204678_s_at	Hs.79351	KCNK1	chr1q42-q43	0.0094796
220661_s_at	Hs.23617	ZNF692	chr1q44	0.0026939
Gains at 16p13.3-16p13.2				
222009_at	Hs.274361	CEMP1	chr16p13.3	0.0017931
206416_at	Hs.592088	ZNF205	chr16p13.3	0.0022295
218969_at	Hs.16089	Magmas	chr16p13.3	0.0052089
207134_x_at	Hs.294158	TPS2	chr16p13.3	0.0087975
213177_at	Hs.88500	MAPK8IP3	chr16p13.3	0.0094881
Losses at 2q36.3-2q37.1				
201370_s_at	Hs.78946	CUL3	chr2q36.3	0.0072559
218091_at	Hs.171545	HRB	chr2q36.3	0.0097753
Losses at 12q24.13-12q24.31				
219350_s_at	Hs.169611	DIABLO	chr12q24.31	0.000952
213153_at	Hs.154525	SETD1B	chr12q24.31	0.0042973
201632_at	Hs.78592	EIF2B1	chr12q24.31	0.0081670
218175_at	Hs.288909	CCDC92	chr12q24.31	0.0099578

NOTE: Listed are genes differentially expressed ($P < 0.01$) in the current study corresponding within the genetic regions showing loss or gain in association with relapse in the study of Natrajan et al. (21).

other studies have shown *POSTN* to be down-regulated in several human tumors and to inhibit metastasis and anchorage-independent growth (20). *POSTN* was down-regulated in relapses in FHWT. Williams et al. compared 27 pre-chemotherapy FHWT samples of all stages, of which 13 relapsed (14). None of their 15 top genes were identified in our Supplementary Table S2. Li et al. used a custom cDNA array and analyzed 26 pre-chemotherapy tumor samples (12 relapses) of all stages and of both favorable and unfavorable histology (15). Four top genes were reported that were capable of identifying tumors that relapsed: *C/EBPB*, *p21*, *H4FG*, and cDNA CF542255

(*CLK1*). The P values of the associations of these genes with relapse in our study were all >0.1 . In summary, the available current literature associating gene expression with relapse in Wilms tumor includes studies with small sample sizes and significant clinical heterogeneity. Our study has the advantage of increased numbers as well as the analysis of only stage III tumors of favorable histology treated using a single chemotherapeutic regimen.

Natrajan et al. performed array-based comparative genomic hybridization analysis to define regions of genetic gain or loss that correlate with relapse in 76 FHWT of all stages, half of

which relapsed (21). Genes within these chromosomal regions that were concordantly differentially expressed with $P < 0.01$ in our study are listed in Table 4. (No genes meeting these criteria were found within regions 13q31–13q33, 14q32, 18q21, 21q22, and 1q32.) Several genes were found to be both gained and overexpressed on chromosome 1q. *ADAM15* may function in metastasis (22), *TGFB2* has an established role in tumorigenesis and cancer progression (23), *CDC42BPA* (*MRCK*) is a protein kinase implicated in tumor cell invasion (24), and *SETDB1* and *SMYD2* are methyltransferases up-regulated in relapses, and the latter suppresses the proapoptotic action of p53 (25). Genes gained and overexpressed in the 16p13.2–13.2 region include *MAPK8IP3* whose up-regulation is associated with cell invasion and brain tumor malignancy (26). Lastly, 4 genes in 12q24.13–24.31 were down-regulated in our analysis including *DIABLO*, involved in apoptosis as described further below.

Specific genes involved in relapse of stage III FHWT. Those genes within Supplementary Table S2 with known or speculated biological function are provided in Table 3. There are several genes that merit special attention. *PEG3* (down-regulated in FHWT relapse) is induced after DNA damage by a p53-dependent mechanism and contributes to the Bax translocation necessary for apoptosis (27). *PEG3* is silenced by methylation in several gynecologic cancer cell lines (28). *DIABLO*, likewise down-regulated in FHWT relapse, is released from the mitochondrial membrane into the cytosol during apoptotic signaling, where it contributes to caspase activation (29). Loss of function mutations and down-regulation of *DIABLO* are associated with tumor aggressiveness and drug resistance (30, 31). Because *DIABLO*, *PEG3*, and *SMYD2* (mentioned earlier) all operate in the p53/Bax apoptotic pathway, this pathway may be of significance in relapse of FHWT. *TUSC3* is a potential tumor suppressor gene with decreased expression in advanced ovarian cancer (32). As with *PEG3*, down-regulation of *TUSC3* occurs via methylation in glioblastoma (33). *RECK* acts as a suppressor of tumor invasion and metastasis by inhibiting matrix metalloproteinases (34). Down-regulation of *RECK* in colon cancer via promoter methylation results in increased cell invasion, and these actions are reversed by DMNT inhibition (35). Therefore, down-regulation by methylation of several top genes in our study and up-regulation of two methyltransferases (*SETDB1* and *SMYD2*) suggest that epigenetic changes may play a role in outcome within stage III FHWT.

Several biologically significant genes were up-regulated with relapse in our study. *FRAP1/MTOR* is a serine/threonine kinase involved in signal transduction, translation initiation, and elongation. It exerts cell growth and survival effects via the Akt/PTEN pathway. *FRAP1* has a well-documented role in cancer (reviewed in ref. 36) and is a promising therapeutic target (37). Another up-regulated gene that is a potential therapeutic target is *CD40*, a member of the tumor necrosis factor family. *CD40* induces antiapoptotic genes (including *Bcl-2* and its family members), early angiogenesis, and activation of cell proliferation signaling pathways (38). Both *BCL2* and *BLC2L11*, a *Bcl-2*-like gene, were significantly up-regulated in our data ($P = 0.008$ and 0.0001, respectively). The up-regulation of *CD40* and *Bcl-2* provides further support to the importance of apoptosis in FHWT relapse. *LY6D*, a GPI-anchored protein involved in cell-cell adhesion, is associated with high-risk lung cancer and with colorectal cancer invasion (39, 40).

FGF18, up-regulated in FHWT relapse, is also up-regulated in colon cancer (41). *FGF18* is regulated at the transcriptional level by TCF-CTNNB1 responsive elements in its promoter region (41). *FGF18* induces nuclear accumulation of CTNNB1 by down-regulation of GSK3B activity at the protein level (42). CTNNB1 has been shown to undergo mutational activation in 10% to 15% of Wilms tumor, and there is a strong association between WT-1 mutation or deletion and CTNNB1 mutation (43). In addition, the putative tumor suppressor gene, *WTX*, which is mutated in ~30% of Wilms tumors, was recently shown to induce degradation of CTNNB1 (44, 45). Our data therefore suggest that aberrant Wnt signaling may be involved in FHWT relapse.

Insulin-like growth factor (IGF)-I and its receptor, *IGFR1*, have been inversely associated with WT-1 in studies showing repression of *IGFR1* by WT-1 and repression of WT-1 by IGF-I (46). Copy number gain at chromosome 15q23.5 (the location of *IGFR1*) has been shown to correlate with FHWT relapse, and gain of expression of *IGFR1* was confirmed by reverse transcription-PCR (47). Analysis of our expression data showed no significant increases within relapse samples of *IGFR1*, *IGFR2*, or *IGF2*, whereas IGF-I was significantly down-regulated in association with relapse. Although high levels of IGF-I have been implicated in oncogenesis in several different cancers, exogenous IGF-I was reported to induce apoptosis in a Wilms tumor cell line (48), suggesting that IGF-I may in certain situations induce cell death. Therefore down-regulation of IGF-I in relapses, or up-regulation in nonrelapses, may affect cell survival.

LOH of 1p and of 16q have been associated with relapse in FHWT (4). Only 3 of the 53 stage III relapses in our study were predicted as high risk by LOH of both 1p and 16q, consistent with the reported data that LOH detects only a small subset of relapses in FHWT. Rather than provide support for the association of loss of expression on 1p with relapse, our data show that 16 of the top 100 genes were located on 1p and were overexpressed, including several biologically important genes such as *PTGER3*, *UQCRH*, and *FRAP1*. It has been suggested that the critical event resulting in 1p LOH may involve gain of 1q rather than loss of 1p (49). This hypothesis is supported by Natrajan et al., as discussed above, which showed increased copy number of 1q to be associated with relapse (21). Our data provide additional support of this hypothesis by showing 6 of the top genes to be located on 1q, all of which were overexpressed.

In conclusion, this study provides some evidence supporting an association between gene expression and relapse in tumors of patients with stage III FHWT. The sensitivity of 47% and specificity of 70% moderate enthusiasm for direct clinical translation of a classifier based on gene expression, although this remains a potential interest if validation on an independent patient population is successful. Several different hypotheses are raised or further supported by these data and merit further investigation in the attempt to find potentially more robust methods of stratifying patients with FHWT. These include the relationship between relapse and apoptotic pathways, IGF-I signaling, the Wnt/β-catenin pathway, 1q gain, and epigenetic modification.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Green DM. The treatment of stages I-IV favorable histology Wilms tumor. *J Clin Oncol* 2004;22:1366–72.
2. Breslow NE, Ou SS, Beckwith JB, et al. Doxorubicin for favorable histology, stage II-III Wilms tumor: results from the National Wilms Tumor Studies. *Cancer* 2004;101:1072–80.
3. Grundy RG, Hutton C, Middleton H, et al. Outcome of patients with stage III or inoperable WT treated on the second United Kingdom WT protocol (UKWT2); a United Kingdom Children's Cancer Study Group (UKCCSG) study. *Pediatr Blood Cancer* 2004;42:311–9.
4. Grundy PE, Breslow NE, Li S, et al. Loss of heterozygosity for chromosomes 1p and 16q is an adverse prognostic factor in favorable-histology Wilms tumor: a report from the National Wilms Tumor Study Group. *J Clin Oncol* 2005;23:7312–21.
5. Huang CC, Cutcliffe C, Coffin C, Sorensen PH, Beckwith JB, Perlman EJ. Classification of malignant pediatric renal tumors by gene expression. *Pediatr Blood Cancer* 2006;46:728–38.
6. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. *Nat Biotechnol* 2003;21:818–21.
7. Vapnik VN. Statistical learning theory. New York (NY): John Wiley and Sons; 1998.
8. Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;95:14–8.
9. Green DM, Breslow NE, Beckwith JB, et al. Treatment with nephrectomy only for small, stage I/favorable histology Wilms tumor: a report from the National Wilms Tumor Study Group. *J Clin Oncol* 2001;19:3719–24.
10. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;365:488–92.
11. Storey JD. A direct approach of false discovery rates. *J R Stat Soc Ser B* 2002;64:479–98.
12. Beckwith JB, Palmer NF. Histopathology and prognosis of Wilms tumors: results from the First National Wilms Tumor Study. *Cancer* 1978;41:1937–48.
13. Zirn B, Hartmann O, Samans B, et al. Expression profiling of Wilms tumors reveals new candidate genes for different clinical parameters. *Int J Cancer* 2006;118:1954–62.
14. Williams RD, Hing SN, Greer BT, et al. Prognostic classification of relapsing favorable histology Wilms tumor using cDNA microarray expression profiling and support vector machines. *Genes Chromosomes Cancer* 2004;41:65–79.
15. Li W, Kessler P, Yeger H, et al. A gene expression signature for relapse of primary Wilms tumors. *Cancer Res* 2005;65:2592–601.
16. Hanavadi S, Martin TA, Watkins G, Mansel RE, Jiang WG. Expression of interleukin 11 and its receptor and their prognostic value in human breast cancer. *Ann Surg Oncol* 2006;13:802–8.
17. Wright KO, Messing EM, Reeder JE. Increased expression of the acid sphingomyelinase-like protein ASML3a in bladder tumors. *J Urol* 2002;168:2645–9.
18. Kudo Y, Ogawa I, Kitajima S, et al. Periostin promotes invasion and anchorage-independent growth in the metastatic process of head and neck cancer. *Cancer Res* 2006;66:6928–35.
19. Bao S, Ouyang G, Bai X, et al. Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. *Cancer Cell* 2004;5:329–39.
20. Kim CJ, Yoshioka N, Tambe Y, Fukushima R, Okada Y, Inoue H. Periostin is down-regulated in high grade human bladder cancers and suppresses *in vitro* cell invasiveness and *in vivo* metastasis of cancer cells. *Int J Cancer* 2005;117:51–8.
21. Natrajan R, Williams RD, Hing SN, et al. Array CGH profiling of favourable histology Wilms tumours reveals novel gains and losses associated with relapse. *J Pathol* 2006;210:49–58.
22. Kuefer R, Day KC, Kleer CG, et al. ADAM15 disintegrin is associated with aggressive prostate and breast cancer disease. *Neoplasia* 2006;8:319–29.
23. Bellone G, Carbone A, Tibaudi D, et al. Differential expression of transforming growth factors- β 1, - β 2 and - β 3 in human colon carcinoma. *Eur J Cancer* 2001;37:224–33.
24. Wilkinson S, Paterson HF, Marshall CJ, Cdc42-MRCK and Rho-ROCK signalling cooperate in myosin phosphorylation and cell invasion. *Nat Cell Biol* 2005;7:255–61.
25. Huang J, Perez-Burgos L, Placek BJ, et al. Repression of p53 activity by Smad2-mediated methylation. *Nature* 2006;444:629–32.
26. Takino T, Nakada M, Miyamori H, et al. JSAP1/JIP3 cooperates with focal adhesion kinase to regulate c-Jun N-terminal kinase and cell migration. *J Biol Chem* 2005;280:37772–81.
27. Deng Y, Wu X, Peg3/Pw1 promotes p53-mediated apoptosis by inducing Bax translocation from cytosol to mitochondria. *Proc Natl Acad Sci U S A* 2000;97:12050–5.
28. Dowdy SC, Gostout BS, Shridhar V, et al. Biallelic methylation and silencing of paternally expressed gene 3 (PEG3) in gynecologic cancer cell lines. *Gynecol Oncol* 2005;99:126–34.
29. Kim R. Recent advances in understanding the cell death pathways activated by anticancer therapy. *Cancer* 2005;103:1551–60.
30. Bonavida B, Huerta-Yépez S, Goodlick L, Mizutani Y, Miki T. Can we develop biomarkers that predict response of cancer patients to immunotherapy? *Biomarkers* 2005;10 Suppl 1:S69–76.
31. Tirro E, Consoli ML, Massimino M, et al. Altered expression of c-IAP1, survivin, and Smac contributes to chemotherapy resistance in thyroid cancer cells. *Cancer Res* 2006;66:4263–72.
32. Pils D, Horak P, Gleiss A, et al. Five genes from chromosomal band 8p22 are significantly downregulated in ovarian carcinoma: N33 and EFA6R have a potential impact on overall survival. *Cancer* 2005;104:2417–29.
33. Li Q, Jedlicka A, Ahuja N, et al. Concordant methylation of the ER and N33 genes in glioblastoma multiforme. *Oncogene* 1998;16:3197–202.
34. Noda M, Oh J, Takahashi R, Kondo S, Kitayama H, Takahashi C. RECK: a novel suppressor of malignancy linking oncogenic signaling to extracellular matrix remodeling. *Cancer Metastasis Rev* 2003;22:167–75.
35. Cho CY, Wang JH, Chang HC, Chang CK, Hung WC. Epigenetic inactivation of the metastasis suppressor RECK enhances invasion of human colon cancer cells. *J Cell Physiol* 2007;213:65–9.
36. Hay N, Sonenberg N. Upstream and downstream of mTOR. *Genes Dev* 2004;18:1926–45.
37. Faivre S, Kroemer G, Raymond E. Current development of mTOR inhibitors as anticancer agents. *Nat Rev Drug Discov* 2006;5:671–88.
38. Dallman C, Johnson PW, Packham G. Differential regulation of cell survival by CD40. *Apoptosis* 2003;8:45–53.
39. Lu Y, Lemon W, Liu PY, et al. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med* 2006;3:e467.
40. Reichling T, Goss KH, Carson DJ, et al. Transcriptional profiles of intestinal tumors in Apc(Min) mice are unique from those of embryonic intestine and identify novel gene targets dysregulated in human colorectal tumors. *Cancer Res* 2005;65:166–76.
41. Shimokawa T, Furukawa Y, Sakai M, et al. Involvement of the FGF18 gene in colorectal carcinogenesis, as a novel downstream target of the β -catenin/T-cell factor complex. *Cancer Res* 2003;63:6116–20.
42. Katoh M, Katoh M. Cross-talk of WNT and FGF signaling pathways at GSK3 β to regulate β -catenin and SNAIL signaling cascades. *Cancer Biol Ther* 2006;5:1059–64.
43. Maiti S, Alam R, Amos CI, Huff V. Frequent association of β -catenin and WT1 mutations in Wilms tumors. *Cancer Res* 2000;60:6288–92.
44. Rivera MN, Kim WJ, Wells J, et al. An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. *Science* 2007;315:642–5.
45. Major MB, Camp ND, Berndt JD, et al. Wilms tumor suppressor WTX negatively regulates WNT/ β -catenin signaling. *Science* 2007;316:1043–6.
46. Bentov I, LeRoith D, Werner H. The WT1 Wilms tumor suppressor gene: a novel target for insulin-like growth factor-I action. *Endocrinology* 2003;144:4276–9.
47. Natrajan R, Reis-Filho JS, Little SE, et al. Blastemal expression of type I insulin-like growth factor receptor in Wilms tumors is driven by increased copy number and correlates with relapse. *Cancer Res* 2006;66:11148–55.
48. Granerus M, Johannsson A, Ekblom P, Engstrom W. Insulin-like growth factors I and II induce cell death in Wilms tumor cells. *Mol Pathol* 2001;54:30–5.
49. Lu YJ, Hing S, Williams R, Pinkerton R, Shipley J, Pritchard-Jones K. Chromosome 1q expression profiling and relapse in Wilms tumor. *Lancet* 2002;360:385–6.

Activated TLR Signaling in Atherosclerosis among Women with Lower Framingham Risk Score: The Multi-Ethnic Study of Atherosclerosis

Chiang-Ching Huang^{1*}, Kiang Liu¹, Richard M. Pope², Pan Du³, Simon Lin³, Nalini M. Rajamannan⁴, Qi-Quan Huang², Nadereh Jafari⁵, Gregory L. Burke⁶, Wendy Post⁷, Karol E. Watson⁸, Craig Johnson⁹, Martha L. Daviglus¹, Donald M. Lloyd-Jones^{1,4}

1 Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States of America, **2** Rheumatology Division, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States of America, **3** Bioinformatics Core, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States of America, **4** Cardiology Division, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States of America, **5** Genomics Core, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States of America, **6** Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, North Carolina, United States of America, **7** Johns Hopkins University, Baltimore, Maryland, United States of America, **8** Division of Cardiology, UCLA School of Medicine, Los Angeles, California, United States of America, **9** Collaborative Health Studies Coordinating Center, University of Washington, Seattle, Washington, United States of America

Abstract

Background: Atherosclerosis is the leading cause of cardiovascular disease (CVD). Traditional risk factors can be used to identify individuals at high risk for developing CVD and are generally associated with the extent of atherosclerosis; however, substantial numbers of individuals at low or intermediate risk still develop atherosclerosis.

Results: A case-control study was performed using microarray gene expression profiling of peripheral blood from 119 healthy women in the Multi-Ethnic Study of Atherosclerosis cohort aged 50 or above. All participants had low (<10%) to intermediate (10% to 20%) predicted Framingham risk; cases (N=48) had coronary artery calcium (CAC) score >100 and carotid intima-media thickness (IMT) >1.0 mm, whereas controls (N=71) had CAC<10 and IMT <0.65 mm. We identified two major expression profiles significantly associated with significant atherosclerosis (odds ratio 4.85; P<0.001); among those with Framingham risk score <10%, the odds ratio was 5.30 (P<0.001). Ontology analysis of the gene signature reveals activation of a major innate immune pathway, toll-like receptors and IL-1R signaling, in individuals with significant atherosclerosis.

Conclusion: Gene expression profiles of peripheral blood may be a useful tool to identify individuals with significant burden of atherosclerosis, even among those with low predicted risk by clinical factors. Furthermore, our data suggest an intimate connection between atherosclerosis and the innate immune system and inflammation via TLR signaling in lower risk individuals.

Citation: Huang C-C, Liu K, Pope RM, Du P, Lin S, et al. (2011) Activated TLR Signaling in Atherosclerosis among Women with Lower Framingham Risk Score: The Multi-Ethnic Study of Atherosclerosis. PLoS ONE 6(6): e21067. doi:10.1371/journal.pone.0021067

Editor: Sunil K. Ahuja, South Texas Veterans Health Care System, United States of America

Received January 13, 2011; **Accepted** May 19, 2011; **Published** June 16, 2011

Copyright: © 2011 Huang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by the National Institutes of Health (R01 HL086678; C.C.H. and R01 AR055240; R.M.P.). The CHAS was supported by grant R01 HL081141 (M.L.D.). The MESA was supported by contracts N01-HC-95159 through N01-HC-95166 and N01-HC-95169 from the National Heart, Lung, and Blood Institute, National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: huangcc@northwestern.edu

Introduction

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality in the United States and the developed world, and it will soon be the leading cause in the developing world. Atherosclerosis accounts for the vast majority of fatal and non-fatal CVD events. Multivariable (MV) risk equations [1], such as the Framingham risk score (FRS), that include traditional CVD risk factors are generally correlated with presence and extent of subclinical atherosclerosis; however, substantial numbers of individuals with low (<10%) to intermediate (10% to 20%) FRS still develop atherosclerosis [2,3], and the majority of CVD events

occur in those with <20% 10-year predicted risk [4]. Refining risk assessment by identifying those apparent outliers who have significant atherosclerotic burden despite low or intermediate predicted risk, and are therefore at higher risk for events than average [5], could help identify novel mechanisms of atherosclerosis development and progression and target preventive therapies to those who are most likely to benefit.

Atherosclerosis is a disease of inflammation characterized by interactions among platelets, leukocytes, and endothelial cells [6,7,8]. Dysfunctional endothelium associated with the presence of atheroma expresses surface molecules such as chemokines, chemoattractants and adhesion molecules that are associated with



alterations of expression of cell-surface and secretory proteins in circulating leukocytes and platelets. Therefore, peripheral blood can be useful to measure the burden and further understand the molecular mechanisms of atherosclerosis.

In the present study, we performed microarray gene expression profiling of whole blood from a population of healthy women with low to intermediate FRS to test whether gene expression profiles could distinguish those with and without significant atherosclerosis. We further sought to identify genes and pathways associated with significant burden of atherosclerosis among these women predicted to be at lower risk for CVD events.

Results

Characteristics of participants

Among all 119 MESA women, 48 had significant subclinical atherosclerosis (cases) and 71 had no evident subclinical atherosclerosis (controls); among those with $\text{FRS} < 10\%$, there were 39 cases and 69 controls. The characteristics of all 119 participants at the fourth examination with and without evident subclinical atherosclerosis are shown in Table 1. As expected, a greater burden of CVD risk factors was observed among participants with significant subclinical atherosclerosis. Age, smoking, systolic blood pressure, and the use of cholesterol lowering medications were significantly different between the two groups ($p < 0.05$).

Overview of the gene expression profiles

The gene expression patterns differ greatly between cases and controls. The sample dendrogram generated by hierarchical clustering shows that the samples separate into two major branches indicated by BL and BR (Figure 1). The largest cluster of genes (G4) was responsible for separating the samples into these two groups. The G4 cluster contains a large number of genes encoding ribosomal proteins (RPS, RPL, MRPS, MRPL genes) and genes related to oxidative phosphorylation (ATP5J, NDUFA1, A6, A9, B8, S4, COX7C, 7A2, 17), suggesting a role for

mitochondrial function and energy metabolism. More African Americans and fewer whites were found in the BR branch than in the BL branch (BR: 38% vs BL: 15% for African American, BR: 40% vs BL: 68% for whites, $P = 0.002$). More cases were found in the BL branch, however, the association of these two branches with atherosclerosis status did not reach statistical significance (BL vs BR branch: 46% vs 33%, $P = 0.19$). Each of these two branches further clustered into two sub-branches (BL1, BL2, BR1, and BR2). Two major gene clusters (G1 and G3) distinguish between BL1 and BL2. The majority of genes in G1 are related to immune and inflammatory response (TLR1, 2, 4, 5, 6, 8, IL1 β , IL1R2, IL1RN, NAMPT, FCGR2A, PTGS2) while cell cycle and apoptosis-related genes are found in G3 (BCL2, BAG3, BLK, ATM, MYC, CD27, TNFRSF25). An inverse correlation of gene expression between these two gene clusters is evident. In contrast, the distinct feature between BR1 and BR2 was indicated by gene cluster G2. Several genes involved in lipid and reactive oxygen species metabolic process were discovered (PRDX2, GPX1, GPX4, GLRX5, HAGA, ADIPOR1, OSBP2, CHPT1, PNPLA2). An examination of the prevalence of atherosclerosis among these four sub-branches shows the highest rate in BL2 (22/43, 51%) and lowest rate in BR2 (5/20, 25%), while both the other two branches, BR1 and BL1, had the same rate of 37.5% (9/24 and 12/32). Of note, two subsets indicated by purple bars within the BL2 had the highest rate (both 9/13, 69%) of atherosclerosis. Both subsets are characterized by higher expression of G1 and lower expression of G3, while the subset on the right also had lower expression of G2.

Association of gene expression profiles with atherosclerosis

The distribution of the 2,057 P-values (Figure S1) using t tests comparing all 48 cases and 71 controls showed a higher frequency of probes at low P values than would be expected by chance alone (229 probes, or 11%, with $P < 0.05$). One hundred and eighty genes were found to have an $\text{FDR} < 24\%$ (215 genes with

Table 1. Characteristics of 119 MESA non-diabetic women with low or intermediate Framingham risk scores.

Variable	With Atherosclerosis (n=48)	Without Atherosclerosis (n=71)	P Value*
Framingham Risk Score (%)	5.8 \pm 4.5	2.6 \pm 2.9	<0.001
Age, years	69.4 \pm 6.8	64.3 \pm 7.3	<0.001
Race† (%)	54/27/13/6	58/24/14/4	0.93
Body-Mass Index (Kg/m ²)	28.2 \pm 6.0	27.2 \pm 5.8	0.38
Systolic BP (mm Hg)	130.6 \pm 21.0	116.3 \pm 18.4	<0.001
Diastolic BP (mm Hg)	68.7 \pm 9.0	66.9 \pm 10.5	0.30
Total cholesterol (mg/dL)	200.2 \pm 34.3	193.4 \pm 31.5	0.28
LDL cholesterol (mg/dL)	117.6 \pm 35.5	108.3 \pm 28.5	0.14
HDL cholesterol (mg/dL)	60.0 \pm 13.1	72.9 \pm 16.0	0.28
Triglycerides (mg/dL)	113.1 \pm 69.8	112.1 \pm 73.4	0.72
hs-CRP‡ (mg/L)	5.6 \pm 9.4	2.3 \pm 3.2	0.02
Current smoker (%)	50.0	26.7	0.02
Lipid-lowering medication use (%)	47.9	14.1	<0.001
Anti-hypertension medication use (%)	47.9	32.4	0.13

*p-value by t-test or chi-square test as appropriate.

†non-Hispanic white /African American/Hispanic/Chinese.

‡hs-CRP levels were measured at MESA baseline examination during 2000~2002.

doi:10.1371/journal.pone.0021067.t001

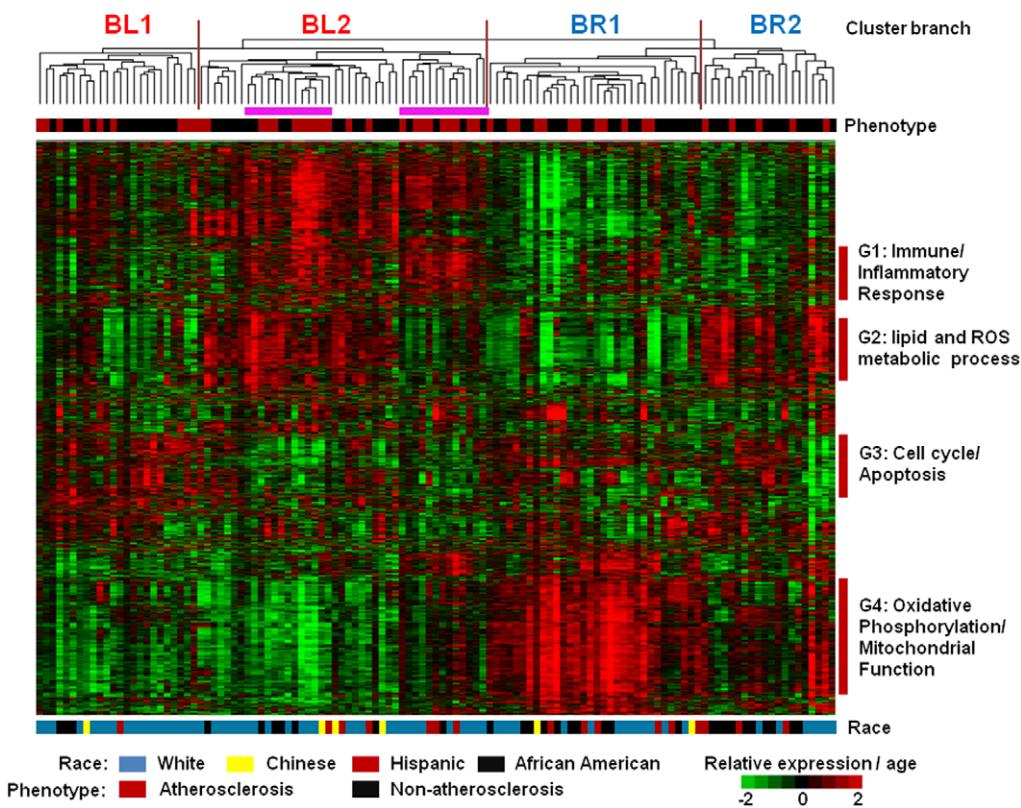


Figure 1. Gene expression patterns among individuals with and without atherosclerosis. 2,057 probes were selected for which expression coefficient of variation was >0.3 and for which at least a half of the 119 microarray samples were detected for the expression values. The unsupervised hierarchical clustering dendrogram shows the relationship among the samples and corresponding gene clusters. Samples have been color-coded by their phenotype (atherosclerosis in red, non-atherosclerosis in black), race (white in blue, Chinese in yellow, Hispanic in red, and African American in black), and age. Age and each probe have been centered on their mean expression values across all samples analyzed. Measurements that are above the mean are colored red and those below the mean are colored green. Groups of genes on the right hand side indicated with colored bars are shown in greater detail and labeled G1–G4. The sample dendrogram shows 4 branches (BL1, BL2, BR1, and BR2). Two subsets (purple bar) under the branch BL2 are enriched with atherosclerosis (both 9/13).

FDR<20% by SAM) and 20% of genes were estimated to be differentially expressed. This result suggests that a gene signature associated with atherosclerosis may exist; however, the moderate FDR of differentially expressed genes implicates that heterogeneous expression profiles may be present within the case and/or control groups.

The multiple random validation procedure using top 50–60 genes with smallest p-values identified two major gene expression profiling (or molecular) subtypes with an odds ratio of 4.85 (sensitivity 60%, specificity 76%) for the association with atherosclerosis (see Figure S2 and S3). We termed these molecular profiles as Associated with Atherosclerosis (AWA) or not AWA. The odds ratio remained high at 5.3 (sensitivity 62%, specificity 77%) when only the 108 women with FRS <10% were considered. As expected, there was a significantly greater number of differentially expressed genes (1,254, 61%) between these two molecular profiles. As shown in Table 2, age and systolic blood pressure were two risk factors significantly different between those with the two molecular profiles, but less of a difference compared with Table 1. Using multiple logistic regression with adjustment for FRS, the gene expression profiling subtypes remained significantly associated with atherosclerosis with an odds ratio of 3.8 (95% CI = [1.6, 8.9], P = 0.002) for all 119 participants and 3.7 (95% CI = [1.5, 9.3], P = 0.004) for the 108 low-FRS participants.

A sensitivity analysis was further performed by excluding those on lipid lowering medication, leaving 25 cases and 61 controls for comparison. The multiple random validation procedure on this subset of samples resulted in an association similar to the results from the entire 119 samples (data not shown). Similarly, the association did not change substantially by excluding smokers or those taking medications for hypertension.

Ontology analysis of differentially expressed genes

In order to better characterize pathways associated with the AWA profile, we performed ontology analysis on the top 344 probes representing 325 unique genes that had an FDR <0.1% from the comparison between the two molecular profiles. Information on these 325 genes is provided in Table S1. The genes that were identified as associated with the AWA profile strongly represented by members of the innate immune signaling pathway. Genes representing several canonical pathways including Toll-like receptors (TLRs), NF- κ B, p38 MAPK, and IL-10 signaling were identified (Table 3). All of these pathways have extensive cross talk in inflammatory signaling. The key genes involved in these pathways include six TLRs (TLR1,2,4,5,6,8), IL1 β , IL1RN, IL1R2, IRAK3, and MAPK14. A representation of this immune and inflammatory network from a subset of the 325 differentially expressed genes is shown in Figure 2, in which several genes are also involved in apoptosis and mobilization of calcium. This highly

Table 2. Characteristics of 119 MESA non-diabetic women between two major molecular profiles (e.g., Associate with Atherosclerosis (AWA) or not AWA).

Variable	With AWA profile (n=46)	Without AWA profile (n=73)	P Value*
Framingham Risk Score (%)	5.0±4.0	3.1±3.8	0.01
Age, years	68.8±6.5	64.8±7.7	0.003
Race [†] (%)	59/20/17/4	55/29/11/5	0.58
Body-Mass Index (Kg/m ²)	27.3±5.8	27.8±5.9	0.62
Systolic BP (mm Hg)	128.2±21.2	118.1±19.4	0.01
Diastolic BP (mm Hg)	69.0±10.3	66.6±9.6	0.21
Total cholesterol (mg/dL)	196.4±32.4	195.9±33.1	0.94
LDL cholesterol (mg/dL)	111.8±33.4	112.2±30.9	0.95
HDL cholesterol (mg/dL)	63.4±14.1	60.6±15.4	0.31
Triglycerides (mg/dL)	105.4±62.6	116.9±77.0	0.38
hs-CRP [‡] (mg/L)	5.3±10.1	2.7±2.4	0.08
Current smoker (%)	45.6	30.1	0.13
Lipid-lowering medication use (%)	36.9	21.9	0.12
Anti-hypertension medication use (%)	47.8	31.5	0.11

*p-value by t-test or chi-square test as appropriate.

[†]non-Hispanic white /African American/Hispanic/Chinese.[‡]hs-CRP levels were measured at MESA baseline examination during 2000~2002.

doi:10.1371/journal.pone.0021067.t002

connected network is characterized by activation of the innate immune system, including the TLR and IL-1 signaling pathways (on the right side) and downstream signaling pathways (PTGS2, MAPK14, DUSP1 and DUSP6). Also upregulated are genes that promote myeloid cell growth and survival (CSF2RA, CSF2RB, CSF3R, C5aR1, GBP1, in addition to the TLR and IL-1 signaling molecules). These data suggest that the AWA profile was an Activated Innate Immune Gene (AIIG) signature. The network also shows decreased expression in genes related to adaptive immunity, including the genes important in T cell activation and maturation (CD3, CD3 epsilon, CD27, CD6, and ZAP70) and CCR7 which is important in the recruitment of T cells and myeloid dendritic cells. Consistent with the result from IPA, DAVID bioinformatics analysis also showed enrichment in response to wounding, immune/inflammatory response, leukocyte activation, and TLR signaling. Other than immune response and cytokine and chemokine mediated signaling pathways, PANTHER analysis also revealed 8 genes involved in the apoptosis pathway. Information regarding the enriched functional categories and genes from these bioinformatics analyses are provided in Table S2.

Comparison with gene expression pattern of peripheral blood leukocytes after endotoxin injection

To further validate the observation that TLR pathway is activated in AWA profile, we compared our gene expression data with Calvano's study[9], which examined the gene expression profiles in blood leukocytes of human subjects who received bacteria endotoxin injection – an established model of acute inflammatory response via TLR activation. For Calvano's study, fold changes and P values for genes were calculated by comparing expression levels at two or six hours after endotoxin injection with baseline levels. A total of 501 and 583 probes, respectively, were found to have P values <0.05 in both our data and Calvano's data at two and six hours after endotoxin injection ($P<10^{-5}$ and $P<10^{-15}$). More importantly, a significant concordance of fold changes was observed between both

studies (Figure 3), with a stronger correlation of fold change between the two studies at six hours after endotoxin injection than at two hours (Pearson correlation 0.59 vs 0.47, $P<10^{-15}$).

Confirmation of gene expression data

Quantitative real-time RT-PCR was performed to confirm the RNA expression levels of eight selected genes involved in various functions including lipid metabolism, immune and inflammatory response, apoptosis, cell proliferation, and hypoxia signaling (TNFSF14, TLR8, TLR4, CREB5, IL1β, IL1RN, EGLN1, TGFA). All eight genes showed high correlation (0.47~0.65, $P = 5.6 \cdot 10^{-8} \sim 1.8 \cdot 10^{-15}$, see Table S3) between the microarray expression and RT-PCR measures.

The association of AIIG signature with atherosclerosis among men

We further characterized the association of atherosclerosis with the AIIG signature described above, employing peripheral blood from 16 white men aged 65–80. The clinical characteristics of these 16 men are shown in Table S4. Most of CVD risk factors are comparable between the two CAC groups. However, lipid-lowering medication use is significantly higher among those with high CAC (75% vs 25%, $P = 0.05$) while their total cholesterol levels were lower (154 mg/dL vs 193 mg/dL, $P<0.001$). Cluster analysis was performed using the top 50 genes with smallest P-values from the AIIG signature. As shown in Figure 4, two major branches were identified that were significantly associated with CAC status ($p = 0.009$, χ^2 test). The left branch consisted of all eight men with high CAC and two men with low CAC, while the right branch consisted of six men, and all had CAC<1. Of note, the two men with low CAC on the left branch had the highest CAC scores (1.2 and 0.9) among the eight men with low-CACs. Because CAC score is highly correlated with the sums of histological artery plaque areas[10], this result suggests that the AIIG signature-atherosclerosis association exists in both genders.

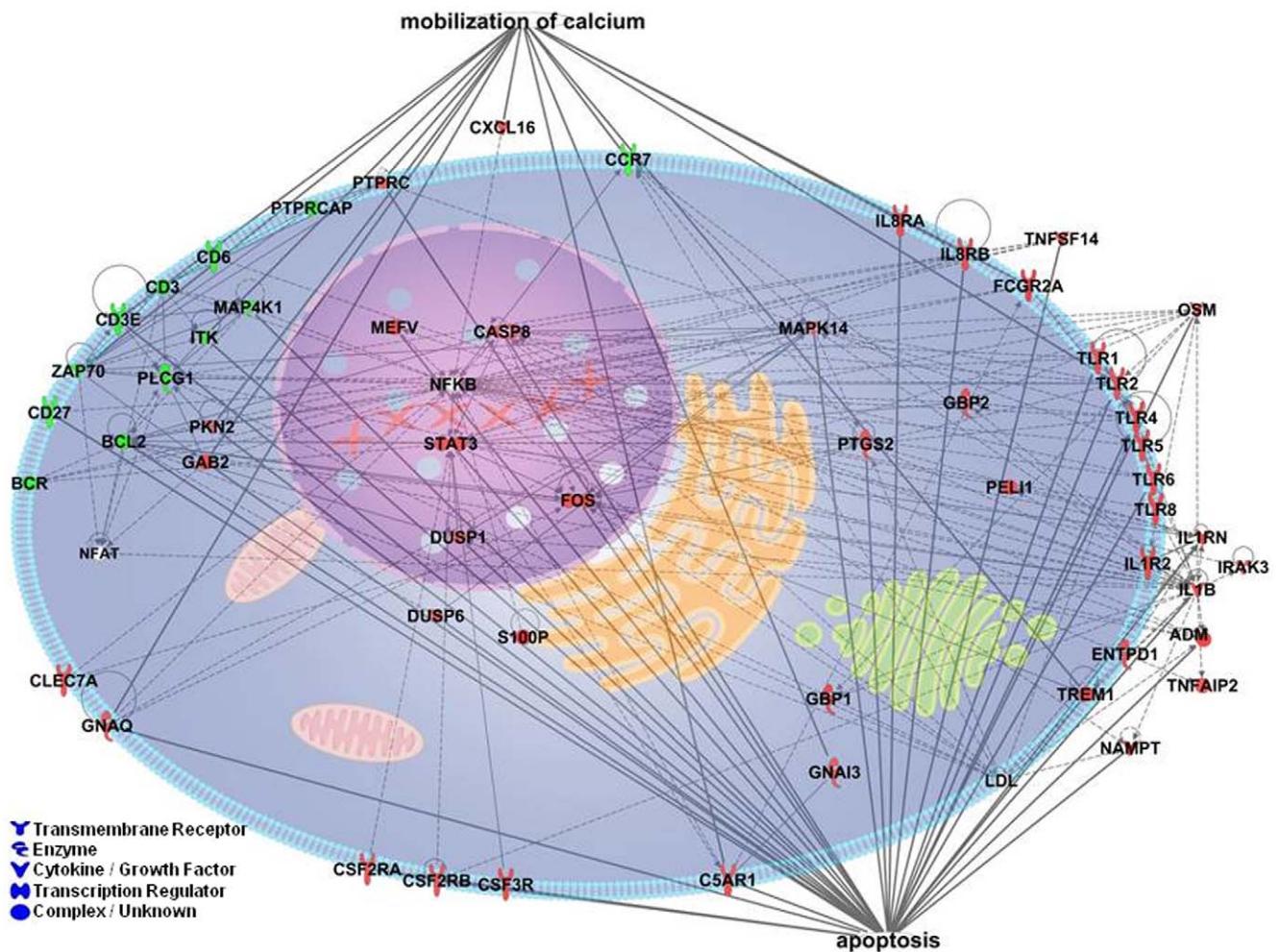


Figure 2. Network representation of the atherosclerosis-associated immune and inflammatory response of peripheral blood. The network consists of 55 genes showing perturbed expression (green, decreased; red, increased expression). This highly connected network is characterized by activation of toll-like receptor and IL-1 signaling (on the right side).
doi:10.1371/journal.pone.0021067.g002

Discussion

Atherosclerosis is the main cause of clinical CVD events. Inflammation plays a central role at all stages of this disease [6]. However, the agents and associated signaling pathways responsible for initiating and perpetuating atherogenesis remain to be

elucidated. The present study used global gene expression profiles of peripheral blood which identified abundant genes that are involved in innate immunity and host defense (AIIG signature) which were strongly associated with significant atherosclerosis among women with low to intermediate FRS. A similar association was also observed among men with advanced atherosclerosis who did not have history of CVD. The finding that increased expression of genes from a number of signaling pathways was associated with atherosclerosis suggests a systemic immune activation mediated through the TLR and IL-1 receptor (IL-1R) signaling pathways, and by the pro-inflammatory mediators MAP kinases (MAPK14 and FOS) and STAT3. This observation is further supported by the striking consistency of expression patterns between our and Calvano's study, suggesting that a chronic low grade inflammation through activation of TLR/IL-1R signaling may be important in the pathogenesis of atherosclerosis, prior to the appearance of any symptoms. Thus, low grade activation of the TLR/IL-1R signaling pathway may represent a potential biomarker for use in measuring presence and/or burden of atherosclerosis.

Prior studies have demonstrated the increased expression of TLR2 and TLR4 on monocytes from patients with angina and acute coronary syndrome [11,12]. Further, patients with recurrent

Table 3. Atherosclerosis-associated genes and the corresponding enriched canonical pathways identified by Ingenuity Pathway Analysis.

Signaling Pathways	Molecule
Toll-like Receptor	TLR1, TLR2, TLR4, TLR5, TLR6, TLR8, FOS, IRAK3, MAPK14
P38 MAPK	MAPK4K1, MAPK14, IL1B, IL1RN, IL1R2, IRAK3, DUSP1, CREB5
IL-10	FCGR2A, FOS, IL1B, IL1RN, IL1R2, MAPK14, STAT3
NF-κB	IL1B, IL1RN, IL1R2, IRAK3, TGFA, ZAP70, TLR1, TLR2, TLR4, TLR5, TLR6, TLR8

doi:10.1371/journal.pone.0021067.t003

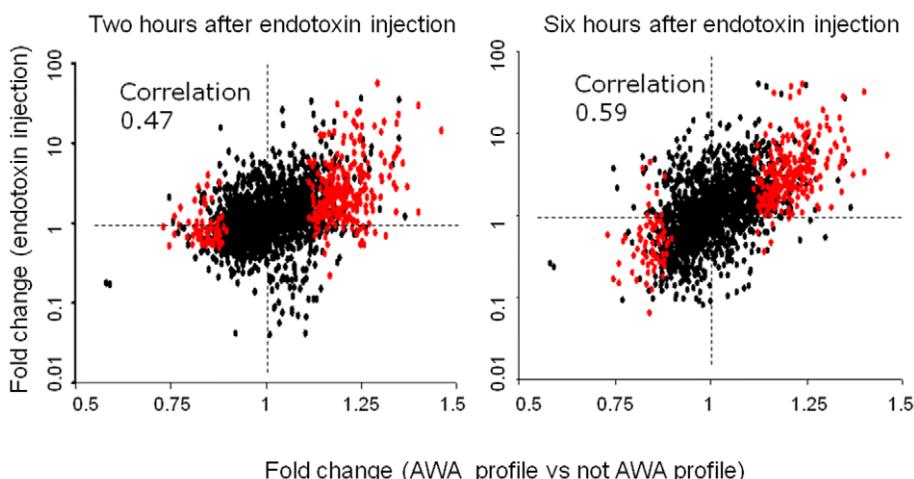


Figure 3. Expression changes of 2,057 probes between acute and chronic inflammation. X-axis: fold changes of gene expression comparing AWA profile with not AWA profile. Y-axis: fold changes of gene expression comparing expression levels before and after endotoxin (LPS) injection. A stronger correlation of fold changes between chronic inflammation and acute inflammation six hours after endotoxin injection was observed. Red dots: 325 genes in the activated innate immune gene signature. AWA: associated with atherosclerosis.
doi:10.1371/journal.pone.0021067.g003

unstable angina demonstrated an increased response to the TLR4 ligand lipopolysaccharide [13]. In contrast, our data suggests that the increased expression of TLR2 and TLR4, and other members of the TLR family, is observed in peripheral blood prior to the onset of clinical symptoms. Supporting the role of TLR2 and TLR4, response for peripheral blood monocytes to TLR2 and TLR4 ligands was significantly associated with the degree of coronary artery stenosis [14]. Additionally, in human atheroma, TLRs 1, 2, 4, 5 and 6 are highly up-regulated when compared to tissue from healthy arteries [15]. Further supporting the role of TLR2, employing cells isolated from carotid endarterectomies of patients with atherosclerosis, inhibition of TLR2 and MyD88, which mediates signals for all TLRs except TLR3, suppressed the expression of spontaneously expressed inflammatory mediators and matrix metalloproteinases [16].

Experimental studies have consistently demonstrated the significant role of innate immunity in atherogenesis. Genetic deletion of TLR4 or TLR2 resulted in significant reduction of arterial plaques in atherosclerosis-prone apoE or LDLR deficient mice [17,18,19]. Additionally, both IL1 β and IL1 α contribute to

the pathogenesis of experimental atherosclerosis [20,21]. Toll-like receptors have also been implicated in several autoimmune and infectious diseases [22,23]. Epidemiological studies have demonstrated that individuals with systemic lupus erythematosus, rheumatoid arthritis, chronic infectious diseases or high levels of circulating bacterial endotoxin have a substantially increased risk of atherosclerosis [24,25,26,27]. These associations may be mediated by signaling through TLR2 (which serves as a co-receptor with TLR1 or TLR6), 4, 7, 8 and 9 [28]. Each of these TLRs except TLR 7 and 9 was increased in our population with atherosclerosis.

T cell activation has been implicated in the pathogenesis of atherosclerosis. T cells are enriched in atherosclerotic lesions [29]. Further, T cell activation has been implicated in patients with acute coronary syndromes [30,31,32]. In contrast, the individuals with atherosclerosis in our study demonstrated a reduction in the expression of genes which contribute to T cell activation and maturation (CD3, CD3E, CD6, CD27). This apparent discrepancy may be explained by the fact that increased T cell activation has been described in patients with acute coronary syndrome.

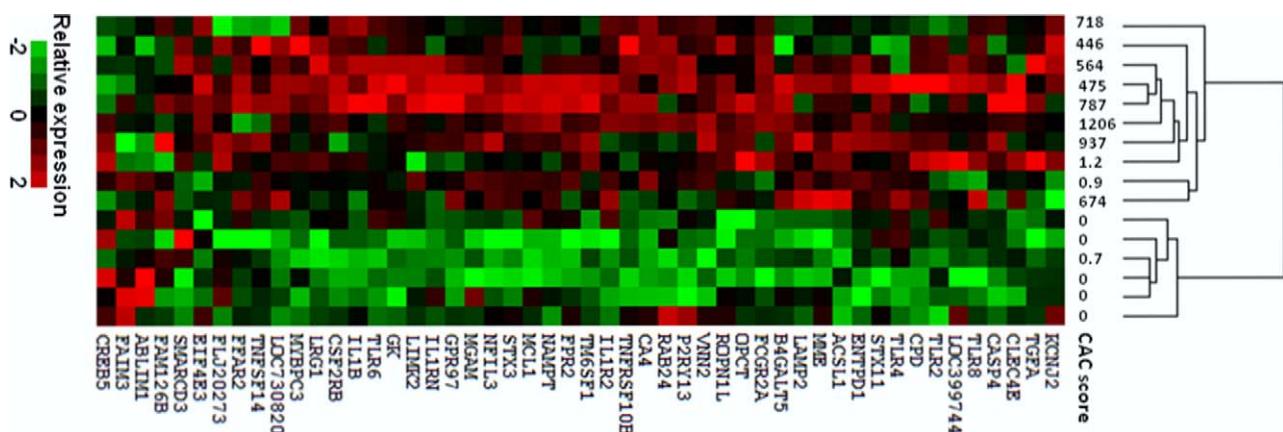


Figure 4. Association of expression pattern with coronary artery calcium (CAC) among men. Cluster analysis of 16 selected men from the CHAS cohort was performed using top 50 genes from the activated innate immune gene (AIIG) signature.
doi:10.1371/journal.pone.0021067.g004

Those in our study had no angina. In fact, these observations are reminiscent of rheumatoid arthritis, which is driven mediated in large part through activation of the innate immune system [33]. Patients with rheumatoid arthritis demonstrate activated macrophages and T lymphocytes in the synovial tissue. However, these patients demonstrate decreased T cell mediated adoptive immunity in the circulation which improves following therapy [33,34]. These observations suggest that activation of the innate immune system may result in reduction of adaptive immunity as measured peripherally.

Recently, two microarray gene expression studies [35,36] of peripheral blood leukocytes in patients with angiographically-documented coronary stenosis were reported by Wingrove et al. and Sinnaeve et al. We applied our innate immune gene signature to their expression data, but failed to verify the associations observed in either study. We found our 325 atherosclerosis-associated genes overlapped little (2 genes) with Sinnaeve's 160 genes associated with presence of coronary stenosis >50%. We observed somewhat greater overlap (14 genes) between our gene signature and Wingrove's 106 differentially-expressed gene that were identified from their literature search or their microarray results in patients referred for angiography and with $\geq 75\%$ stenosis in one major coronary artery or $\geq 50\%$ in two major arteries. Most of these overlapping genes belong to the TNF superfamily, IL1R signaling, and receptors for IL8 or colony stimulating factor. There are several potential reasons for the lack of substantial overlap between our results and those of Sinnaeve and Wingrove. These could include use of different measures to define the presence of atherosclerosis, since CAC presence and increased IMT may represent different (earlier) phases of atherosclerosis progression than presence of severe coronary stenosis. Likewise, there were clear differences in methods for selection of participants between studies. Our study participants were at lower risk for CVD events and are all asymptomatic, while cases in the Sinnaeve and Wingrove studies were referred for angiography on clinical grounds and some had recent acute coronary syndromes; furthermore, control participants in these two studies may have had atherosclerosis, just not severe enough to meet the case definition, whereas our control participants appear to be substantially free of atherosclerosis. Finally, there were differences in statistical analyses used to define differentially expressed genes. Sinnaeve et al used correlation analysis as opposed to two-group comparisons used in our and the Wingrove studies. All of these factors may have contributed to the minimal gene overlap observed between these 3 studies.

The findings from the present study have several potential clinical implications. Higher burden of atherosclerosis, as measured by CAC or carotid IMT, is associated with significantly greater risk for CVD events in the short term, independent of traditional risk factors [37,38,39,40]. Our data suggest that several genes, especially those in the TLR/IL-1R signaling pathway, may serve as potential markers of significant atherosclerosis burden even for those predicted to be at lower risk by traditional risk equations. Toll-like receptors have been suggested as therapeutic target for several inflammatory diseases, including atherosclerosis [41]. Perhaps this signal or similar data indicating TLR/innate immune activation could be used to identify individuals who would benefit from more intensive prevention efforts. An approach that combines gene expression changes in TLR signaling and conventional CVD risk factors may enhance risk prediction for atherosclerosis. Such a strategy would require validation in larger samples.

Some limitations of the current study should be acknowledged. First, because of the cross-sectional nature of our study, the causal

relationship of TLR signaling with atherosclerosis cannot be established in the present data. Second, our study only included 119 women aged 50~86 and 16 men aged 65~80. It is unknown whether our gene signature is applicable to identify atherosclerosis among younger adults. Larger and prospective studies are warranted to examine expression alterations of members of TLR and IL-1R signaling pathways in association with developing subclinical atherosclerosis. Nonetheless, our translational finding is unique in revealing an AIIG signature that was associated with significant atherosclerotic burden, even among a group predicted to be at low risk for atherosclerosis and clinical CVD.

Conclusions

We used global gene expression profiles of whole blood to systematically examine the dynamic transcriptional alterations associated with atherosclerosis among women with low to intermediate predicted risk for the disease. Our results suggest the involvement of the innate immune system in atherogenesis and that whole blood provides an accessible and informative source of transcriptomic information regarding the inflammatory status of atherosclerosis in the preclinical phase of the disease. Because of the easy access of blood samples, expression profiles of whole blood may be a useful tool to assess the burden of atherosclerosis and could potentially be used to enhance the risk prediction of CVD beyond the traditional risk factors. Future studies will be needed to validate our findings.

Methods

Ethics statement

The Institutional Review Boards at all Multi-Ethnic Study of Atherosclerosis (MESA) sites (Northwestern University, Wake Forest University, Johns Hopkins University, Columbia University, University of Minnesota, and UCLA) as well as the Chicago Health Aging Study (CHAS) center (Northwestern University) approved the study, and all participants gave written informed consent.

Participants' blood samples

A subset of 119 study participants was selected from the MESA cohort. MESA is designed to study the prevalence, risk factors, and progression of subclinical cardiovascular disease in a multiethnic cohort. A detailed description of the study design and methods has been published previously [42]. Briefly, 6814 participants aged 45 to 84 years and free of known clinical CVD were recruited from 6 US communities – Baltimore, MD, Chicago, IL, Forsyth County, NC, Los Angeles, CA, New York, NY, and St. Paul, MN. The baseline examination was performed between July 2000 and September 2002. Participants included white (38%), black (28%), Hispanic (22%), and Chinese (12%) Americans. MESA conducted three subsequent examinations of the cohort between 2002 and 2007.

In our case-control study, cases were selected as having a coronary artery calcium (CAC) score ≥ 100 and common carotid intima-media thickness (CC-IMT) ≥ 1.0 mm. Controls were selected as having a CAC score ≤ 10 and CC-IMT ≤ 0.65 mm. Additionally, both cases and controls were selected from those without diabetes or history of CVD who had an FRS less than 10% at baseline examination (2000~2002). Among all MESA women, only 48 women agreed to participate and met the selection criteria as cases with significant atherosclerosis burden despite low predicted risk for CHD. We also selected potential control participants from among MESA women based on similar age and race – 71 women were identified as controls and agreed to

participate, giving us a ratio of 1.5 controls for each case. Out of these 119 participants, the FRS of 11 participants had increased from 11% to 20% when the MESA data for the fourth examination (2005~2007; concurrent with the present study) became available. Carotid IMT was measured for all MESA participants only at the baseline examination. Measurement of CAC was performed for all MESA participants at the baseline examination. Follow-up testing for CAC was performed in two stages: 50% of participants returned from September 2002 to January 2004 (examination 2), and the remainder returned from March 2004 to July 2005 (examination 3). Therefore, we used the CAC score from examination 2 or 3 for these 119 participants. All these 119 participants were called back during April-July 2007 to have 5 ml whole blood drawn into two PAXgene tubes for RNA isolation.

PAXgene blood samples of additional 16 white men without diabetes or history of CVD from the CHAS cohort were selected for gene expression profiling. The blood samples of these men were collected during their clinical visits between April 2009 and July 2009. Because IMT data were not available in CHAS, cases and controls were identified based on their CAC scores. Eight of these 16 men had CAC scores 446~1,206. The other eight had CAC scores <2. The Institutional Review Boards at each MESA site as well as the CHAS center (Chicago, IL) approved the study, and all participants gave informed consent.

CAC and IMT Assessment

Scanning centers assessed CAC by chest-computed tomography using either a cardiac-gated electron-beam computed tomography scanner at three MESA field centers or a multi-detector computed tomography system at the other three MESA field centers. Each participant was scanned twice, and these scans were read independently at a central reading center. The amount of calcium was quantified with the Agatston scoring method. The average Agatston score was used for data analysis. Carotid IMT was measured using high-resolution B-mode ultrasonography of the right and left near and far walls of the internal carotid and common carotid arteries. Images were acquired bilaterally for both the common and internal carotid arteries. The CHAS used a similar method to assess CAC. The methodology for acquisition and interpretation of the scans has been documented previously [43,44].

RNA extraction and microarray experiment

Five ml whole blood from each participant was drawn into two PAXgene tubes and incubated at room temperature for three hours before being frozen at -70°C . RNAs were extracted using the PAXgene blood RNA extraction kit according to the manufacturer's protocol. The concentration of the extracted total RNA and their quality were measured by NanoDrop (Thermo-fisher) and Bioanalyzer 2100 (Agilent), respectively. Total RNA preparations with 260/280 ratio between 1.98–2.22 and RIN number >7.4 with enough quantity were used for the expression analysis. Globin reduction was performed using the Ambion GLOBIN clear kit. The quality of the globin-reduced RNA samples was assessed by Agilent 2100 Bioanalyzer. High quality samples were used to make first and second strand DNA followed by an IVT reaction. The size distribution of the resulting biotin labeled cRNA and the yield was checked by Agilent 2100 Bioanalyzer and NanoDrop, respectively. A normalized amount of labeled cRNA was hybridized for 18 hours at 55°C to the Sentrix Human Ref-8 v2 Expression BeadChip for the 119 MESA samples and HumanHT-12 V4 Expression BeadChip for the 16 CHAS samples (Illumina, Inc, San Diego). After washing and staining with Cy3, the chips were scanned on the Illumina iScan.

The Illumina Human Ref-8 v2 BeadChip allows genome-wide expression profiling of more than 22,000 transcripts of genes and known alternative splice variants from the RefSeq database while the HumanHT-12 V4 Expression BeadChip targets more than 47,000 transcripts of genes, gene candidates, and splice variants. The microarray analysis was performed at the Northwestern Genomic Core Facility at the Center for Genetic Medicine. All the 119 MESA RNA samples were run on 5 different days within a week. Samples collected from six MESA field centers were randomly assigned to different chips and days. All case and control samples in each MESA field center were also randomly assigned to different chips and days. The 16 CHAS RNA samples were also randomly assigned to two different chips and run on the same day.

Quantitative RT-PCR

RNA levels of eight selected genes were analyzed by quantitative real time RT-PCR using the TaqMan chemistry on the 7300 Real-Time PCR System from Life Technologies. GAPDH was used as the endogenous control. Relative expression levels of these genes were determined by the delta-delta threshold cycle (DDC_T) method.

Data analysis

BeadStudio software was used to translate the scanned images into expression data, which were further log transformed and normalized by the quantile normalization procedure using the Bioconductor package: affy. Gene expression data from microarray experiments are known to have substantial background noise to signal ratios for genes expressed at the lower levels. To reduce the chance of identifying such genes, genes with a coefficient of variation less than 0.03 across all 119 MESA samples or a P value for expression detection call less than 0.01 for 60 samples or fewer were excluded, resulting in 2,057 probes for data analysis. Average linkage hierarchical clustering was performed for these 2,057 probes in the 119 MESA samples using Cluster 3.0 software (<http://rana.lbl.gov/EisenSoftware.htm>) using Pearson correlation as a distance metric. Significance analysis of microarray (SAM) [45] and the method developed by Storey and Tibshirani [46] were used to estimate the false discovery rate (FDR) and the number of differentially expressed genes for comparisons between case and control groups and between the two major gene expression profiling subtypes. Three widely-used applications (Ingenuity Pathway Analysis (IPA), PANTHER ontology analysis [47], and DAVID bioinformatics software [48]) were applied to identify pathways or biological functions of differentially expressed genes associated with atherosclerosis. The full MIAME-compliant microarray data were submitted to the NCBI Gene Expression Omnibus data repository (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE20129.

Identification of major gene expression profiling subtypes

We used a multiple random validation procedure [49] to identify major gene expression profiling subtypes associated with atherosclerosis among the 119 MESA participants. Specifically, the MESA data set was first randomly divided into ten groups of approximately equal size. Each group included approximately the same number of samples from participants with significant subclinical atherosclerosis. For each group (test set), a classifier was built with the remaining nine of ten groups (training set) and then used to categorize samples in the group as atherosclerosis or non-atherosclerosis. The classification was performed across all ten groups. For each randomly drawn training set of samples, a multi-gene support vector machine classifier was constructed. For each

gene, the t statistic comparing expression between cases and controls within a training set was calculated. Genes (1–100 genes) with the largest absolute values of t statistics were selected to construct a support vector machine model using an R software package, e1071. This model was then applied to classify samples within the complementary test set as atherosclerosis or non-atherosclerosis. This whole procedure was repeated 1,000 times, and the final classification status was determined by the majority vote of the 1,000 classification results (e.g. >500 votes). The samples were separated into two major gene expression profiling subtypes based on their final classification status.

Gene selection for comparison between the present study and Calvano's study

We used the Northwestern Unigene system to map our 2,057 Illumina probes to those in the Affymetrix U133 +2 chip used in Calvano's study. 1,841 unique genes were found in both microarray chips.

Supporting Information

Figure S1 Distribution of 2,057 P-values from t tests comparing all 71 controls and 48 cases among the MESA women. The dashed black line is the uniform distribution under the null hypothesis of no differential expression. If there were no differential expression observed, all the bars will be approximately at the height of the dashed black line. For these data, the observed P-value distribution is skewed to the right. The dashed red line indicates that a proportion of non-differentially expressed genes (80.4%).
(TIF)

Figure S2 Association of atherosclerosis with gene expression profiles of peripheral blood. The optimal association from the multiple random validation procedure was achieved using a set of 50–60 genes, resulting in an odds ratio of 4.85. When only the 108 women with FRS<10% were considered, the odds ratio was 5.3. SVM: support vector machine. FRS: Framingham risk score.
(TIF)

References

- Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, et al. (1998) Prediction of coronary heart disease using risk factor categories. *Circulation* 97: 1837–1847.
- Postey JE, Perez A, Wong ND, Gardin JM (2009) Prevalence and distribution of sub-clinical atherosclerosis by screening vascular ultrasound in low and intermediate risk adults: the New York physicians study. *J Am Soc Echocardiogr* 22: 1145–1151.
- Nucifora G, Schuijff JD, van Werkhoven JM, Jukema JW, Djaberi R, et al. (2009) Prevalence of coronary artery disease across the Framingham risk categories: coronary artery calcium scoring and MSCT coronary angiography. *J Nucl Cardiol* 16: 368–375.
- Ajani UA, Ford ES (2006) Has the risk for coronary heart disease changed among U.S. adults? *J Am Coll Cardiol* 48: 1177–1182.
- Lakoski SG, Greenland P, Wong ND, Schreiner PJ, Herrington DM, et al. (2007) Coronary artery calcium scores and risk for cardiovascular events in women classified as "low risk" based on Framingham risk score: the multi-ethnic study of atherosclerosis (MESA). *Arch Intern Med* 167: 2437–2442.
- Libby P (2002) Inflammation in atherosclerosis. *Nature* 420: 868–874.
- Gawaz M, Langer H, May AE (2005) Platelets in inflammation and atherogenesis. *J Clin Invest* 115: 3378–3384.
- May AE, Langer H, Seizer P, Bigalke B, Lindemann S, et al. (2007) Platelet-leukocyte interactions in inflammation and atherothrombosis. *Semin Thromb Hemost* 33: 123–127.
- Calvano SE, Xiao W, Richards DR, Feliciano RM, Baker HV, et al. (2005) A network-based analysis of systemic inflammation in humans. *Nature* 437: 1032–1037.
- Rumberger JA, Simons DB, Fitzpatrick LA, Sheedy PF, Schwartz RS (1995) Coronary artery calcium area by electron-beam computed tomography and coronary atherosclerotic plaque area. A histopathologic correlative study. *Circulation* 92: 2157–2162.
- Methe H, Kim JO, Kofler S, Weis M, Nabauer M, et al. (2005) Expansion of circulating Toll-like receptor 4-positive monocytes in patients with acute coronary syndrome. *Circulation* 111: 2654–2661.
- Ashida K, Miyazaki K, Takayama E, Tsujimoto H, Ayaori M, et al. (2005) Characterization of the expression of TLR2 (toll-like receptor 2) and TLR4 on circulating monocytes in coronary artery disease. *J Atheroscler Thromb* 12: 53–60.
- Liuzzo G, Angiolillo DJ, Buffon A, Rizzello V, Colizzi C, et al. (2001) Enhanced response of blood monocytes to in vitro lipopolysaccharide-challenge in patients with recurrent unstable angina. *Circulation* 103: 2236–2241.
- Versteeg D, Hoefer IE, Schoneveld AH, de Kleijn DP, Busser E, et al. (2008) Monocyte toll-like receptor 2 and 4 responses and expression following percutaneous coronary intervention: association with lesion stenosis and fractional flow reserve. *Heart* 94: 770–776.
- Edfeldt K, Swedenborg J, Hansson GK, Yan ZQ (2002) Expression of toll-like receptors in human atherosclerotic lesions: a possible pathway for plaque activation. *Circulation* 105: 1158–1161.
- Monaco C, Gregan SM, Navin TJ, Foxwell BM, Davies AH, et al. (2009) Toll-like receptor-2 mediates inflammation and matrix degradation in human atherosclerosis. *Circulation* 120: 2462–2469.
- Liu X, Ukai T, Yumoto H, Davey M, Goswami S, et al. (2008) Toll-like receptor 2 plays a critical role in the progression of atherosclerosis that is independent of dietary lipids. *Atherosclerosis* 196: 146–154.
- Michelsen KS, Wong MH, Shah PK, Zhang W, Yano J, et al. (2004) Lack of Toll-like receptor 4 or myeloid differentiation factor 88 reduces atherosclerosis and alters plaque phenotype in mice deficient in apolipoprotein E. *Proc Natl Acad Sci U S A* 101: 10679–10684.
- Mullik AE, Tobias PS, Curtiss LK (2005) Modulation of atherosclerosis in mice by Toll-like receptor 2. *J Clin Invest* 115: 3149–3156.
- Kirii H, Niwa T, Yamada Y, Wada H, Saito K, et al. (2003) Lack of interleukin-1beta decreases the severity of atherosclerosis in ApoE-deficient mice. *Arterioscler Thromb Vasc Biol* 23: 656–660.
- Kamari Y, Werman-Venkert R, Shaish A, Werman A, Harari A, et al. (2007) Differential role and tissue specificity of interleukin-1alpha gene expression in atherogenesis and lipid metabolism. *Atherosclerosis* 195: 31–38.

22. Li M, Zhou Y, Feng G, Su SB (2009) The critical role of Toll-like receptor signaling pathways in the induction and progression of autoimmune diseases. *Curr Mol Med* 9: 365–374.
23. Hornef MW, Henriques-Normark B, Normark S (2008) The function and biological role of toll-like receptors in infectious diseases: an update. *Curr Opin Infect Dis* 21: 304–312.
24. Roman MJ, Moeller E, Davis A, Paget SA, Crow MK, et al. (2006) Preclinical carotid atherosclerosis in patients with rheumatoid arthritis. *Ann Intern Med* 144: 249–256.
25. Roman MJ, Shanker BA, Davis A, Lockshin MD, Sammaritano L, et al. (2003) Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. *N Engl J Med* 349: 2399–2406.
26. Kiechl S, Egger G, Mayr M, Wiedermann CJ, Bonora E, et al. (2001) Chronic infections and the risk of carotid atherosclerosis: prospective results from a large population study. *Circulation* 103: 1064–1070.
27. Wiedermann CJ, Kiechl S, Dunzendorfer S, Schratzberger P, Egger G, et al. (1999) Association of endotoxemia with carotid atherosclerosis and cardiovascular disease: prospective results from the Bruneck Study. *J Am Coll Cardiol* 34: 1975–1981.
28. Huang Q, Pope RM (2010) Toll-like receptor signaling: a potential link among rheumatoid arthritis, systemic lupus, and atherosclerosis. *J Leukoc Biol* 88: 253–262.
29. Ross R (1999) Atherosclerosis--an inflammatory disease. *N Engl J Med* 340: 115–126.
30. Liuzzo G, Giubilato G, Pinnelli M (2005) T cells and cytokines in atherogenesis. *Lupus* 14: 732–735.
31. Weyand CM, Goronzy JJ, Liuzzo G, Kopecky SL, Holmes DR, Jr., et al. (2001) T-cell immunity in acute coronary syndromes. *Mayo Clin Proc* 76: 1011–1020.
32. Ammirati E, Vermi AC, Cianflone D, Banfi M, Foglieni C, et al. (2008) Expansion of T-cell receptor zeta dim effector T cells in acute coronary syndromes. *Arterioscler Thromb Vasc Biol* 28: 2305–2311.
33. Pope RM, Krieger WT, Talal N, Dauphinee M (1993) Delayed type hypersensitivity in patients with rheumatoid arthritis. *J Rheumatol* 20: 17–20.
34. Smith MD, Smith A, O'Donnell J, Ahern MJ, Roberts-Thomson PJ (1989) Impaired delayed type cutaneous hypersensitivity in rheumatoid arthritis reversed by chrysotherapy. *Ann Rheum Dis* 48: 108–113.
35. Sinnaeve PR, Donahue MP, Grass P, Seo D, Vonderscher J, et al. (2009) Gene expression patterns in peripheral blood correlate with the extent of coronary artery disease. *PLoS One* 4: e7037.
36. Wingrove JA, Daniels SE, Schnert AJ, Tingley W, Elashoff MR, et al. (2008) Correlation of peripheral-blood gene expression with the extent of coronary artery stenosis. *Circ Cardiovasc Genet* 1: 31–38.
37. Detrano R, Guerci AD, Carr JJ, Bild DE, Burke G, et al. (2008) Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N Engl J Med* 358: 1336–1345.
38. Folsom AR, Krommal RA, Detrano RC, O'Leary DH, Bild DE, et al. (2008) Coronary artery calcification compared with carotid intima-media thickness in the prediction of cardiovascular disease incidence: the Multi-Ethnic Study of Atherosclerosis (MESA). *Arch Intern Med* 168: 1333–1339.
39. Polonsky TS, McClelland RL, Jorgensen NW, Bild DE, Burke GL, et al. (2010) Coronary artery calcium score and risk classification for coronary heart disease prediction. *JAMA* 303: 1610–1616.
40. Nambi V, Chambliss L, Folsom AR, He M, Hu Y, et al. (2010) Carotid intima-media thickness and presence or absence of plaque improves prediction of coronary heart disease risk: the ARIC (Atherosclerosis Risk In Communities) study. *J Am Coll Cardiol* 55: 1600–1607.
41. O'Neill LA, Bryant CE, Doyle SL (2009) Therapeutic targeting of toll-like receptors for infectious and inflammatory diseases and cancer. *Pharmacol Rev* 61: 177–197.
42. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, et al. (2002) Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 156: 871–881.
43. Carr JJ, Nelson JC, Wong ND, McNitt-Gray M, Arad Y, et al. (2005) Calcified coronary artery plaque measurement with cardiac CT in population-based studies: standardized protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) study. *Radiology* 234: 35–43.
44. O'Leary DH, Polak JF, Wolfson SK, Jr., Bond MG, Bommer W, et al. (1991) Use of sonography to evaluate carotid atherosclerosis in the elderly. The Cardiovascular Health Study. CHS Collaborative Research Group. *Stroke* 22: 1155–1163.
45. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98: 5116–5121.
46. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100: 9440–9445.
47. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, et al. (2003) PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 31: 334–341.
48. Huang da W, Sherman BT, Tan Q, Kir J, Liu D, et al. (2007) DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 35: W169–175.
49. Michiels S, Koscielny S, Hill C (2005) Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 365: 488–492.

Gene expression variation between African Americans and whites is associated with coronary artery calcification: the multiethnic study of atherosclerosis

Chiang-Ching Huang,¹ Donald M. Lloyd-Jones,^{1,2} Xiuqing Guo,⁵ Nalini M. Rajamannan,² Simon Lin,³ Pan Du,³ Qiquan Huang,⁴ Lifang Hou,¹ and Kiang Liu¹

¹Department of Preventive Medicine, ²Cardiology Division, Department of Medicine, ³Bioinformatics Core, ⁴Division of Rheumatoid Arthritis, Department of Medicine, Northwestern University, Chicago, Illinois; and ⁵Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California

Submitted 9 December 2010; accepted in final form 21 April 2011

Huang CC, Lloyd-Jones DM, Guo X, Rajamannan NM, Lin S, Du P, Huang Q, Hou L, Liu K. Gene expression variation between African Americans and whites is associated with coronary artery calcification: the multiethnic study of atherosclerosis. *Physiol Genomics* 43: 836–843, 2011. First published April 26, 2011; doi:10.1152/physiolgenomics.00243.2010.—Coronary artery calcium (CAC) is a strong indicator of total atherosclerosis burden. Epidemiological data have shown substantial differences in CAC prevalence and severity between African Americans and whites. However, little is known about the molecular mechanisms underlying initiation and progression of CAC. Microarray gene expression profiling of peripheral blood leucocytes was performed from 119 healthy women aged 50 yr or above in the Multi-Ethnic Study of Atherosclerosis cohort; 48 women had CAC score >100 and carotid intima-media thickness (IMT) >1 mm, while 71 had CAC <10 and IMT <0.65 mm. When 17 African Americans were compared with 41 whites in the low-CAC group, 409 differentially expressed genes (false discovery rate <5%) were identified. In addition, 316 differentially expressed genes were identified between the high- and low-CAC groups. A substantial overlap between these two gene lists was observed (148 genes, $P < 10^{-6}$). Furthermore, genes expressed lower in African Americans also tend to express lower in individuals with low CAC (correlation 0.69, $P = 0.002$). Ontology analysis of the 409 race-associated genes revealed significant enrichment in mobilization of calcium and immune/inflammatory response ($P < 10^{-9}$). Of note, 25 of 30 calcium mobilization genes were involved in immune/inflammatory response ($P < 10^{-10}$). Our data suggest a connection between immune response and vascular calcification and the result provides a potential mechanistic explanation for the lower prevalence and severity of CAC in African Americans compared with whites. ethnicity; coronary artery calcium; immune response

CORONARY ARTERY CALCIUM (CAC) is highly correlated with total atherosclerosis burden (33). Epidemiological studies have shown that CAC can improve prediction of cardiovascular disease (CVD) events independently of traditional coronary risk factors among various race/ethnic groups (8, 13, 17, 31). Several studies have shown racial/ethnic differences in CAC prevalence and severity that cannot be fully explained by conventional coronary risk factors (3, 10, 24, 28). In all studies, African Americans have a significantly lower prevalence of CAC and average CAC score than whites, despite a higher prevalence of hypertension and diabetes. Nevertheless, the

mechanisms accounting for the differential CAC burden in these two populations are poorly understood. These observations suggest that underlying biological processes and genetic predisposition are likely to play a role in the modulation of CAC. Identifying the mechanisms underlying differential manifestations of CAC between these two populations could potentially identify novel targets for prevention of CVD.

Recent studies have demonstrated that vascular calcification is an active biological process similar to bone formation (7). Several signaling pathways, including RANKL and Wnt signaling, have been shown to be crucial in vascular calcification and bone-related diseases. Because these pathways also play an important role in the immune response (14, 25, 35), these findings suggest a potential mechanistic link between atherosclerosis, which is promoted by immune and inflammatory response, and vascular calcification. Whole blood consists of several immune cells such as B cells, T cells, and monocytes that are in frequent contact with endothelial cells, transit into the subendothelial vascular wall, and also participate in osteogenesis. Therefore, whole blood may represent a useful surrogate tissue to investigate the pathophysiology of vascular calcification. Studying gene expression patterns of peripheral blood may reveal novel insights into the development of CAC.

In the present study, we performed gene expression profiling of whole blood in a selected sample of women without diabetes or history of clinical cardiovascular disease from the Multi-Ethnic Study of Atherosclerosis (MESA). With its well-documented phenotype data and multiethnic populations, MESA provides a unique opportunity to compare gene expression patterns between African Americans and whites and investigate whether differentially expressed genes between African Americans and whites and genetic pathways identified therein are associated with CAC.

METHODS AND MATERIALS

Participants. Participants in the present study were selected from the MESA cohort. The MESA was designed to study the prevalence, risk factors, and progression of subclinical cardiovascular disease in a multiethnic cohort. A detailed description of the study design and methods has been published previously (2). In brief, 6,814 participants aged 45–84 yr and free of known clinical CVD were recruited from six US communities (Baltimore, MD; Chicago, IL; Forsyth County, NC; Los Angeles, CA; New York, NY; and St. Paul, MN). The baseline examination was performed between July 2000 and September 2002. Participants included white (38%), African American (28%), Hispanic (22%), and Chinese (12%) individuals. MESA con-

Address for reprint requests and other correspondence: C.-C. Huang, 680 N. Lake Shore Dr., Ste. 1400, Chicago, IL 60611 (e-mail: huangcc@northwestern.edu).

ducted three subsequent examinations of the cohort between 2002 and 2007, and the fifth examination is ongoing from 2010.

Participants in the present study included 67 white and 30 African American women without diabetes or a history of CVD; 26 whites and 13 African Americans had a CAC score ≥ 100 and common carotid intima-media thickness (CC-IMT) ≥ 1.0 mm, while 41 whites and 17 African Americans had a CAC score ≤ 10 and CC-IMT ≤ 0.65 mm. Additionally, 16 Hispanic and 6 Chinese women were included. Three Chinese and six Hispanic women were among the high-CAC and high-IMT group. These 16 women were only included in analyses for gene expression comparison between high-CAC and low-CAC groups. The majority of the 119 women (108) had a low Framingham risk score (i.e., FRS < 10%). Men were not included in the current study because very few MESA male participants had low FRS and high CAC. Carotid IMT was measured for all MESA participants only at the baseline examination. Measurement of CAC was first performed for all MESA participants at the baseline examination. A second CAC measurement was performed during the follow-up: 50% of all participants at the second MESA examination (2002–2003), and the remaining 50% of participants at the third MESA examination (2003–2004). We used the most recent CAC score available for each of these 119 participants. These participants were called back during April–July 2007 to have 5 ml whole blood drawn into two PAXgene tubes for RNA isolation. The Institutional Review Boards at each site approved the study, and all participants gave informed consent.

CAC and IMT assessment. Scanning centers assessed CAC by chest-computed tomography using either a cardiac-gated electron-beam computed tomography scanner at three field centers or a multidetector computed tomography system at the other three field centers. Each participant was scanned twice, and these scans were read independently at a central reading center, as described previously (5). Carotid IMT was measured using high-resolution B-mode ultrasonography of the right and left near and far walls of the internal carotid and common carotid arteries (29). Given the large number of missing internal carotid IMT data, we elected to use CC-IMT as the selection criterion.

RNA extraction and microarray experiment. Five milliliters of whole blood from each participant were drawn into two PAXgene tubes and incubated at room temperature for 3 h before being frozen at -70°C . RNAs were extracted using the PAXgene blood RNA extraction kit according to the manufacturer's protocol. The concentration of the extracted total RNA and their quality were measured by NanoDrop (Thermofisher) and Bioanalyzer 2100 (Agilent), respectively. Total RNA preparations with 260/280 ratio between 1.98–2.22 and RNA integrity number > 7.4 with sufficient quantity were used for the mRNA expression analysis. Globin reduction was performed using the Ambion GLOBINclear kit. The quality of the globin-reduced RNA samples was assessed by the Bioanalyzer 2100. High-quality samples were used to make first- and second-strand DNA followed by an IVT reaction. The size distribution of the resulting biotin-labeled cRNA, and the yield was checked by Agilent 2100 and NanoDrop, respectively. A normalized amount of labeled cRNA was hybridized to the Human Ref-8 bead chips for 18 h at 55°C . After being washed and stained with Cy3, the chips were scanned on the Illumina iScan. The Ref-8 BeadChip allows genome-wide expression profiling of $>22,000$ gene transcripts and known alternative splice variants from the RefSeq database.

The microarray analysis was performed at the Northwestern Genomic Core Facility at the Center for Genetic Medicine. All the 119 RNA samples were run on 5 different days within 1 wk. Samples collected from six MESA field centers were randomly assigned to different chips and days. All samples with low and high CAC in each MESA field center were also randomly assigned to different chips and days. Correlations of gene expression between two replicates run in different days were >0.99 . Cluster analysis of normalized gene expression data revealed no significant chip-to-chip or day-to-day variation.

Quantitative RT-PCR. To validate the expression levels in microarrays, total RNA samples from the low-CAC white and low-CAC African American women were analyzed for mRNA expression by quantitative real-time RT-PCR (QRT-PCR) for GAB2, NOTCH1, MMP9, QPCT, TNFSF14, TLR8, and IL1RN genes. Reverse transcription was performed by the SuperScript III VILO cDNA synthesis kit (Invitrogen). Real-time PCR was performed by the Taqman gene expression assay system (Applied Biosystems, Emeryville, CA). The primer and probe mixes were purchased from Applied Biosystems referenced to the probe location for each gene applied in the microarray assay. GAPDH was used as the endogenous control. The relative gene expression was determined by the $\Delta\Delta\text{-}C_{\text{T}}$ method. All experiments were performed in duplicate.

Data analysis. BeadStudio software was used to translate the scanned images into expression data, which were further log-transformed and normalized by the quantile normalization procedure using the Bioconductor package affy. All data analyses were based on 2,057 probes representing 1,967 genes that met two criteria: 1) coefficient of variation > 0.03 across all samples, and 2) *P* value for expression detection call < 0.01 for 60 samples or more. Significance analysis of microarrays (40) (SAM) was used to identify differentially expressed genes for a two-group comparison. Standardized fold changes were calculated as mean expression differences divided by pooled standard deviation in the log scale. To account for potential confounding factors, body mass index (BMI), triglyceride, and insulin resistance measure using the homeostatic model assessment (HOMA) were included in the regression models, and FDR for differentially expressed genes (16) were calculated. To account for European admixture in African Americans, we further adjusted for the first principle component derived from the MESA genome-wide association data among African Americans that are available from the National Center for Biotechnology Information (NCBI) dbGaP database (26). The MESA SHARe project (study accession phs000209, http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000209.v5.p1) includes 8,227 individuals (2,686 Caucasians, 777 Chinese, 2,590 non-Hispanic African-Americans, and 2,174 Hispanics). All of them were genotyped using the Affymetrix Human SNP Array 6.0. We excluded monomorphic single nucleotide polymorphisms (SNPs); SNPs with missing rate $> 5\%$ or observed heterozygosity $> 53\%$; SNPs that were not in Hardy-Weinberg equilibrium ($P < 10^{-5}$); as well as SNP pairs in high linkage disequilibrium (LD) ($r^2 \geq 0.8$). We further removed 6,849 SNPs in genomic regions that have been shown to harbor long-range LD. This left a final sample of 718,707 autosomal SNPs from which we performed Principal Component Analysis as implemented in the program SMARTPCA (30, 32) from the software package EIGENSTRAT to compute principal components of ancestry. Ingenuity Pathway Analysis (IPA, <http://www.ingenuity.com>) and PANTHER ontology analysis (39) were used to identify enriched pathways or biological functions among differentially expressed genes. All microarray data were submitted to the NCBI Gene Expression Omnibus data repository under accession number GSE20129.

RESULTS

Characteristics of the study sample. Characteristics of the 97 white and African American participants in the study sample are shown in Table 1 stratified by CAC status and race. Among all CVD risk factors, triglyceride levels were significantly different between white and African American women in both high-CAC and low-CAC groups, while BMI was significantly different between white and African American women only in the low-CAC group. Among the low-CAC group, lipid-lowering medication use was substantially higher among whites, while antihypertension medication use was higher among African Americans, although neither of these differences was statistically significant. No significant differences were ob-

Table 1. Characteristics of 97 MESA nondiabetic women by race and by coronary artery calcification status

Variable	CAC <10		CAC >100	
	African American	White	African American	White
n	17	41	13	26
Age, yr	64.1 ± 7.7	62.9 ± 6.8	68.3 ± 8.5	69.0 ± 6.5
Body mass index, kg/m ² †	32.0 ± 6.4	25.5 ± 4.2	30.9 ± 7.1	27.4 ± 4.6
Systolic BP, mmHg	119.1 ± 17.3	114.4 ± 18.7	130.6 ± 18.2	130.0 ± 22.6
Diastolic BP, mmHg	68.6 ± 7.8	65.9 ± 11.1	71.4 ± 9.5	67.1 ± 8.7
Total cholesterol, mg/dl	191.8 ± 32.4	197.5 ± 31.0	199.7 ± 29.7	204.0 ± 37.0
LDL cholesterol, mg/dl	113.5 ± 27.7	110.2 ± 27.5	122.6 ± 30.4	119.0 ± 39.1
HDL cholesterol, mg/dl	61.6 ± 13.1	65.0 ± 15.6	61.3 ± 13.5	59.1 ± 12.6
Triglycerides, mg/dl‡§	83.7 ± 32.8	111.3 ± 67.5	78.6 ± 29.8	129.2 ± 72.3
Fasting glucose, mg/dl	91.5 ± 8.0	89.7 ± 8.5	90.6 ± 7.5	91.0 ± 8.7
HOMA insulin resistance	0.92 ± 0.45	0.74 ± 0.39	1.04 ± 0.55	1.06 ± 0.70
Insulin, uIU/ml	4.4 ± 2.0	3.6 ± 1.8	4.8 ± 2.4	4.9 ± 3.1
Current smoker, %	23.5	37.5	53.8	48.0
Lipid-lowering medication use, %	5.8	15.3	38.4	52.0
Antihypertension medication use, %	47.0	25.6	53.8	48.0

MESA, Multi-Ethnic Study of Atherosclerosis. †P < 0.05 comparing African American with white among women with coronary artery calcification (CAC) <10. §P < 0.05 comparing African American with white among women with CAC >100. Note: P value by t-test or χ² test as appropriate.

served for blood pressure, total, LDL, or HDL cholesterol, smoking status, fasting glucose, insulin, and insulin resistance measure of HOMA between these two ethnic groups in both high-CAC and low-CAC groups.

Identifying differentially expressed genes between African Americans and whites. The gene expression pattern was distinct between 41 whites and 17 African Americans in the low-CAC group. SAM identified 431 probes representing 409 genes [false discovery rate (FDR) < 5%] differentially expressed between these two groups (Supplemental Fig. S1A and Supplemental Table S1).¹ In contrast, only 30 differentially expressed genes were identified at the FDR < 5% level between the 26 whites and 13 African Americans in the high-CAC group (Supplemental Fig. S1B). Because BMI and triglyceride levels are significantly different between African American and white women in our data, we fitted a linear regression model adjusted for BMI and triglyceride for each of the 2,057 probes from the low-CAC group comparison. Only 86 probes were associated with BMI and 37 probes associated with triglyceride with nominal P values < 0.05 (FDR 51 and 99%, respectively). Race was associated with 484 probes (nominal P value < 0.05, FDR 14%), with a majority of these being (409) race-associated genes (333, 81.4%). Regression analysis with additional adjustment for insulin resistance or adjusted only for BMI showed a similar result. When the same regression analysis was performed with additional adjustment for African ancestry in African American women, 28 probes were associated with BMI and 191 probes associated with triglyceride with nominal P values < 0.05 (FDR 99 and 36%, respectively). Race was associated with 524 probes (nominal P value < 0.05, FDR 11%), among which the majority (342, 84%) were race-associated genes. These results suggest that the association between expression levels of race-associated genes and race are not explained by BMI and triglyceride.

We further compared our 409 race-associated genes with those reported by Schisler et al. (34), who examined the expression profiles of whole blood from 17 African American and 30 Caucasian men and women without coronary artery

disease. Schisler et al. reported 151 geo-ancestral genes differentially expressed between these two populations. These geo-ancestral genes were defined as containing at least one SNP within 10 kb of the untranslated regions that has significant difference in allele frequency between the two HapMap populations, the Yoruba of Ibadan, Nigeria (YRI), and the CEPH population (Utah residents with ancestry in northern and western Europe). Only 56 of Schisler's geo-ancestral genes were found among the 1,967 genes in our study because the microarray platform, gene filtering processes, and criteria for selecting differentially expressed genes used in Schisler's study differed from those in our study. Nevertheless, a striking overlap (36 genes, P < 10⁻¹⁶) was observed between the 56 genes and our 409 race-associated genes. Specifically, 4 of 11 genes expressed at a higher level in African Americans and 32 of the 45 genes expressed at a lower level in African Americans in Schisler's list appeared in our gene list (see Table 2). Of note, all of the 36 genes had concordant expression directions between Schisler's and our study. More importantly, among the 32 genes expressed at a lower level in American Americans, 9 of the 10 genes with the largest fold changes in Schisler's list (S100P, MMP9, C20orf3, QPCT, KRT23, USP10, NOV, HK2, and PGD) appeared in our race-associated gene list. The significant overlap between our and Schisler's gene lists suggests a high true positive rate of the identified race-associated genes. Several studies have linked genetic variations to gene expression among various HapMap populations and other Caucasian populations. Comparing the compilation of those studies available in the University of Chicago expression quantitative trait locus (eQTL) database (<http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/>) identified 18 genes (50%) with at least one eQTL (or exon-QTL) (see Table 2), suggesting that genetic variation is attributable to the expression differences between African Americans and whites.

Comparison with the CAC-associated gene expression pattern. A high correlation of global expression changes was observed between the CAC and race-associated gene expression patterns. We calculated the standardized fold changes of all 2,057 probes between 71 low-CAC and 48 high-CAC participants. Interestingly, a high correlation (Pearson correla-

¹ The online version of this article contains supplemental material.

Table 2. 36 common race-associated genes found in both MESA and Schisler's studies

Gene Symbol	Gene Name	Entrez Gene ID	Fold Change	Fold Change (Schisler's study)	With eQTL or exon-QTL
<i>Genes expressed lower in African American</i>					
S100P	S100 calcium binding protein P	6286	2.8	2.79	Y
MMP9	matrix metalloproteinase 9	4318	2.18	1.98	
C20orf3	chromosome 20 open reading frame 3	57136	1.21	1.87	Y
QPCT	glutaminyl-peptide cyclotransferase	25797	1.3	1.68	Y
KRT23	keratin 23	25984	1.19	1.51	Y
USP10	ubiquitin specific peptidase 10	9100	1.25	1.5	
NOV	nephroblastoma overexpressed gene	4856	1.43	1.5	
HK2	hexokinase 2	3099	1.17	1.49	
PGD	phosphogluconate dehydrogenase	5226	1.28	1.49	Y
SULF2	sulfatase 2	55959	1.27	1.48	Y
LAMP2	lysosomal-associated membrane protein 2	3920	1.22	1.46	
GPR97	G protein-coupled receptor 97	222487	1.19	1.42	
KIAA0319L	KIAA0319-like	79932	1.2	1.42	Y
NADK	NAD kinase	65220	1.25	1.41	
MME	membrane metallo-endopeptidase	4311	1.34	1.41	
HEBP2	heme binding protein 2	23593	1.37	1.39	Y
TMEM45B	transmembrane protein 45B	120224	1.21	1.39	
PLAUR	plasminogen activator, urokinase receptor	5329	1.31	1.36	Y
ACOX1	acyl-Coenzyme A oxidase 1, palmitoyl	51	1.25	1.36	Y
STX3	syntaxin 3	6809	1.21	1.35	
RNF135	ring finger protein 135	84282	1.26	1.35	
SRPK1	SFRS protein kinase 1	6732	1.16	1.35	
UBN1	ubinuclein 1	29855	1.3	1.35	Y
FLOT1	flotillin 1	10211	1.2	1.34	Y
TMEM55A	transmembrane protein 55A	55529	1.24	1.34	
REPS2	RALBP1 associated Eps domain containing 2	9185	1.24	1.33	
ANPEP	alanyl (membrane) aminopeptidase	290	1.31	1.33	Y
PYGL	phosphorylase, glycogen, liver	5836	1.25	1.33	
GAB2	GRB2-associated binding protein 2	9846	1.34	1.33	Y
ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminy-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	10610	1.31	1.32	Y
CDA	cytidine deaminase	978	1.5	1.31	Y
AGPAT9	1-acylglycerol-3-phosphate O-acyltransferase 9	84803	1.36	1.31	Y
<i>Genes expressed higher in African American</i>					
CRIP1	cysteine-rich protein 1 (intestinal)	1396	1.27	1.3	
NKG7	natural killer cell group 7 sequence	4818	1.33	1.36	Y
GZMH	granzyme H (cathepsin G-like 2, protein h-CCPX)	2999	1.67	1.38	
IGJ	immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	3512	1.99	1.46	

eQTL, expression quantitative trait locus.

tion 0.57, $P = 0.009$, permutation test) of fold changes was observed between the two low-CAC race group (i.e., low-CAC white vs. low-CAC African American) comparison and the two CAC group (i.e., all low-CAC vs. all high-CAC) comparison (Fig. 1). Among the 409 race-associated genes, the correlation was 0.69 ($P = 0.002$, permutation test) between these two comparisons and 88% of these genes had concordant fold changes. This correlation (0.39) was attenuated when the two high-CAC race groups were used for comparison. SAM identified 337 differentially expressed probes representing 317 genes ($FDR < 22\%$) between the high- and low-CAC groups. A highly significant overlap of these two gene lists was observed (148 genes, $P < 10^{-6}$). More than half of the race-associated genes (145/284 genes) expressed at a lower level in African American women were found among the CAC associated genes ($P < 10^{-15}$). The result was similar when 16 Hispanic and 6 Chinese women were excluded from the analysis. Furthermore, we fit a linear regression model with CAC, race, and the CAC-race interaction term for each of the 2,057 probes from the 97 African American and white women. A

total of 110 of the 409 race-associated genes were associated with CAC (nominal P value < 0.05 , FDR $< 4\%$). In contrast, only 98 of the remaining 1,626 probes were associated with CAC. Comparing the regression coefficients of race and CAC, a higher correlation (0.79) and a higher concordance (94%) of coefficient signs were observed for the 409 race-associated genes. Taken together, these results suggest that some race-associated genes may contribute to the development of high CAC ($P = 0.0057$). The results also suggest that genes expressed at a lower (or higher) level in African American women tend to be expressed at a lower (or higher) level in those with low CAC.

Ontology enrichment analysis of race-associated genes between two low-CAC groups. Among the 409 race-associated genes, IPA identified several enriched functional categories related to immune and inflammatory response and apoptosis. Simultaneous survey and evaluation of these functional categories enabled us to identify molecular mechanisms important to the process of vascular calcification. Examples of the diversity in these mechanisms include 1) increased expression (in

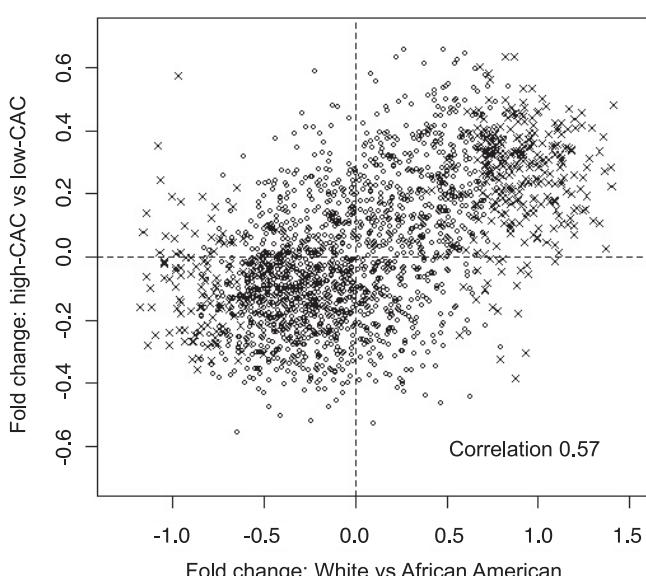


Fig. 1. Correlation of standardized fold changes (log-scale) of 1,967 genes between the low-coronary artery calcium (CAC) race comparison (x-axis) and high-/low-CAC comparison (y-axis). A high correlation ($0.57, P = 0.009$) of standardized fold changes was observed between these 2 group comparisons. Among the 409 race-associated genes (\times), the correlation was 0.69 ($P = 0.002$).

whites) of members of Toll/IL-1R signaling (TLR1, TLR8, IL1RN, IL1R2), a key pathway critical for the activation of innate immunity and development of adaptive immunity; 2) increased expression of ATP-binding cassette transporter family genes (ABCG1, ABCC5) that control cholesterol efflux; 3) increased expression of matrix metalloproteinases and their endogenous inhibitors (MMP9, MMP25, and TIMP2), key regulators of bone formation and extracellular matrix degradation; 4) increased expression of FRAT1 and LRP10, critical for regulation of the Wnt/ β -catenin signaling pathway; and 5) increased expression of cytosolic regulatory components (NCF1, NCF4) of NADPH oxidase pathway, a major source of generating reactive oxygen species in the artery. Of note, mobilization of calcium was one of the most enriched functional categories (30 genes, $P = 8 \cdot 10^{-10}$). Systemic examination of these 30 genes revealed an involvement of the majority (25/30) of these genes in immune and inflammatory responses (see Table 3 and Supplemental Fig. S2). For example, IPA identified an inflammatory and immune response network that included seven calcium-related genes ($P = 10^{-7}$, Supplemental Fig. S3). A literature search for the functions of these 409 genes further identified several additional genes crucial to bone metabolism or bone-related diseases, including TNFSF14 (12), NOTCH1 (15), MMP9 (41), and QPCT (20). Consistent with the IPA results, PANTHER ontology analysis also revealed significant enrichment of B-cell and T-cell-mediated immunity and defense, inflammation mediated by chemokine and cytokine signaling pathway, and apoptosis. Taken together, these observations suggest that many of the race-associated genes involved in multiple biological processes synergistically contribute to the development of CAC.

Validation of gene expression data. Quantitative real-time RT-PCR was performed on the low-CAC African Americans and whites to confirm the microarray expression levels of

seven selected genes with potential roles in RANK pathway, lipid metabolism, immune and inflammatory response, and vascular calcification (GAB2, TNFSF14, TLR8, IL1RN, MMP9, NOTCH1, QPCT). We confirmed the expression results for all these seven genes except NOTCH1 ($P \leq 0.05$) (Fig. 2).

DISCUSSION

African Americans are less likely to develop CAC and have lower CAC scores on average than whites. However, the molecular mechanism responsible for this difference is largely unknown. In the present study, we compared gene expression profiles of whole blood from African American and white women. We found a large number of differentially expressed genes between these two ethnic groups among those with low CAC scores. Interestingly, the expression differences of these genes significantly correlated with the expression differences between groups with negligible or no CAC and high CAC. Investigation of the functional modules of the race-associated genes revealed enrichment and substantial overlap of two major functional axes: calcium mobilization and immune response. The finding of these molecular modules is consistent with that by Storey et al. (36), who identified between the HapMap CEPH and YRI samples a large number of differentially expressed genes that were strongly enriched in inflammatory and immune-related pathways. Studies have shown that RANKL and Wnt signaling pathways are crucial in vascular calcification and bone-related diseases and play an important role in the immune response (14, 25, 35). Thus, these pathways provide a potential mechanistic link between immune and inflammatory response and vascular calcification. Our result not only supports the emerging concept of cross talk between immunity and osteogenic pathways but also suggests that a large proportion of the race-associated genes identified in the present study may be associated with the development of CAC. Although other pathways could possibly explain the difference in CAC prevalence and severity between African Americans and whites, we found no major evidence for this in our gene expression data. Taken together, these results provide a potential and compelling mechanistic explanation for the greater CAC prevalence and severity among whites compared with African Americans, seen even after adjustment for traditional CVD risk factor levels. Our study benefited from using a sample of participants from a well-phenotyped cohort study specifically designed to explore differences in subclinical CVD between race/ethnic groups, thus minimizing the effect of potential confounding factors on gene expression patterns associated with CAC.

Vascular calcification, once regarded as a passive degenerative disease, involves a complex mineralization process similar to bone formation. Remodeling of bone involves synthesis of bone matrix by osteoblasts and coordinates bone resorption by osteoclasts. Recent efforts to improve our knowledge of this disease have led to several findings and provided novel research directions.

An emerging area is the study of key regulators of the immune and bone systems including receptor activator of NF- κ B (RANK), its ligand (RANKL), and osteoprotegerin (OPG). However, the molecular mechanism and coupling molecules that control RANKL signaling have not been fully

Table 3. 30 race-associated genes involved in mobilization of calcium

Gene Symbol	Gene Name	Entrez Gene ID	Fold Change*	Involved in Immune Response	Involved in Inflammatory Response
ANXA1	annexin A1	301	-1.39	X	X
C5AR1	complement component 5a receptor 1	728	1.26	X	X
CAMP	cathelicidin antimicrobial peptide	820	1.5	X	X
CCL5	chemokine (C-C motif) ligand 5	6352	-1.3	X	X
CCR3	chemokine (C-C motif) receptor 3	1232	1.36		X
CD2	CD2 molecule	914	-1.25	X	
CX3CR1	chemokine (C-X3-C motif) receptor 1	1524	-1.2	X	X
CXCL1	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	2919	1.26	X	X
F2R	coagulation factor II (thrombin) receptor	2149	-1.21	X	X
F2RL1	coagulation factor II (thrombin) receptor-like 1	2150	1.33	X	X
FCRL3	Fc receptor-like 3	115352	-1.31		
FPR2	formyl peptide receptor 2	2358	1.38	X	X
GAB2	GRB2-associated binding protein 2	9846	1.34		
GNAQ	guanine nucleotide binding protein (G protein), q polypeptide	2776	1.18		
HEBP1	heme binding protein 1	50865	1.24		X
IL8	interleukin 8	3576	1.36	X	X
CXCR1	chemokine (C-X-C motif) receptor 1	3577	1.21		X
CXCR2	chemokine (C-X-C motif) receptor 2	3579	1.21	X	X
ITGB1	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	3688	-1.32		X
ITK	IL2-inducible T-cell kinase	3702	-1.24	X	
KLRD1	killer cell lectin-like receptor subfamily D, member 1	3824	-1.52	X	
KLRF1	killer cell lectin-like receptor subfamily F, member 1	51348	-1.27	X	
KLRK1	killer cell lectin-like receptor subfamily K, member 1	22914	-1.35	X	
LAT2	linker for activation of T cells family, member 2	7462	1.24	X	
LPAR2	lysophosphatidic acid receptor 2	9170	1.2		
LTB4R	leukotriene B4 receptor	1241	1.21	X	X
PILRA	paired immunoglobulin-like type 2 receptor alpha	29992	1.29		
PLAUR	plasminogen activator, urokinase receptor	5329	1.31		X
PTAFR	platelet-activating factor receptor	5724	1.2	X	X
PTGDR	prostaglandin D2 receptor (DP)	5729	-1.22		X

*Compare white with African American.

characterized. Our study has identified several novel genes that might act as coupling molecules between the immune and bone systems. For example, we found a novel gene, GAB2, that was expressed at a lower level in African Americans and those with low CAC scores. The gene, GAB2, was recently shown to be a key regulatory scaffold molecule that controls RANK signaling pathways and to have a crucial role in the differen-

tiation of human progenitor cells into osteoclasts (42). We also identified TNFSF14, a member of the TNF superfamily that promotes a RANKL-mediated osteoclastogenesis and can induce osteoclast formation independent of RANKL (12, 21).

Another area of active research area is the Wnt3a/β-catenin and LDLR-related protein (LRP5)-dependent activation of the canonical Wnt signaling cascade in calcifying human aortic valves (4). We also observed FRAT1 and LRP10, two molecules essential for the regulation of Wnt3a-induced Wnt signaling and canonical Wnt/β-catenin signaling pathways (18, 22). Further investigation of these genes may deepen our understanding of the complex mechanism of vascular calcification regulated by these signaling pathways.

Gene expression is influenced by genetic variation (6, 37, 38). Storey et al. (36) analyzed gene expression profiles of B lymphoblastoid cells from the CEPH and YRI samples used in the International HapMap project. Among the genes differentially expressed between these two populations, they identified molecular functions and signaling pathways including inflammation mediated by chemokine and cytokine, T-cell and B-cell activation, VEGF signaling, and Toll-like receptor signaling. Their gene ontology enrichment analysis results are similar to ours even though they analyzed different cell types and lifestyle differences exist between African Americans and YRI. Additionally, the geo-ancestral genes identified by Schisler et al. (34) contained at least one SNP that has significantly different allele frequency between the HapMap YRI and CEPH

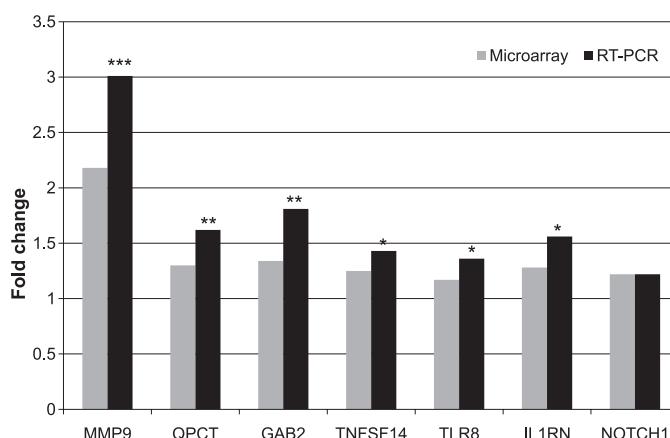


Fig. 2. QRT-PCR validation of 7 race-associated genes identified from the microarray experiment. Bars represent fold changes comparing expression between white and African American women in the low-CAC group. A fold change >1 represents upregulation of gene expression in white women. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$.

populations and significantly overlapped with our gene list. Therefore, underlying genetic variation between these two populations is likely to influence gene expression to a large extent. Future studies combining genetic and expression data may identify genetic variants that influence the initiation and progression of vascular calcification.

Although genetic variation influences an individual's susceptibility to diseases, environmental factors, and lifestyle can contribute substantially to changes in gene expression (9, 11, 23, 43). We observed that differences in gene expression profiles between asymptomatic African American and white women were attenuated considerably when the comparison was made between two high-CAC race groups. This suggests that reductions in the number of differentially expressed genes and the fold change magnitudes are due to the constellation of CVD risk factors influencing gene expression levels. This finding may have an important public health implication because individuals at higher risk for vascular calcification may be able to reduce their susceptibility to the diseases by changing their lifestyle.

The present study included only women for gene expression profile comparisons. It is unknown whether our race-associated genes are differentially expressed between men in these two populations. The significant overlap between our and Schisler's gene lists, however, suggests a lack of sex-related bias. In addition, the race-associated genes may correlate with IMT because participants in the high-CAC group also had higher IMT. Nevertheless, epidemiological data have shown that African Americans tend to have thicker CC-IMT than whites (19, 27). Therefore, the observed positive correlation between race-associated genes and CAC-associated genes is more likely to reflect the difference in CAC rather than IMT. Lastly, the present study has a small sample size in the low-CAC group that allowed us to adjust only a limited number of confounding factors. Therefore, potential residual confounding may exist. Future studies with larger samples and a different design are needed to address the specific confounding issues.

Since CAC is a strong indicator of atherosclerosis, some race-associated genes may also be involved in atherogenesis. For example, our gene list contains several crucial atherogenic genes including PLAUR, MMP9, and several chemokines and cytokines (IL8, CXCL1, CCL5, IL6R). This result is not surprising because several molecular mechanisms are shared between CAC and atherosclerosis (1). To our knowledge, CAC has not been used as a sole phenotype in any gene expression profile studies. Further studies are needed to verify our finding.

In summary, we have reported the first correlation of CAC with gene expression profiles among African Americans and whites. Our data provide a potential explanation for the lower prevalence of CAC among African Americans compared with whites. Functional enrichment analysis of our race-associated genes supports an important role of the immune system in vascular calcification. Joint epidemiological and laboratory investigation of several novel genes discovered in the present study may elucidate the determinants of CAC and other bone-related diseases such as valvular calcification and osteoporosis. Several race-associated genes like cytokines and chemokines could be used, in combination with conventional CVD risk factors, to identify individuals at high risk for developing vascular calcification.

GRANTS

This research was supported by National Institutes of Health (NIH) Grant R01 HL-086678 (C.-C. Huang). The MESA was supported by NIH contracts N01-HC-95159 through N01-HC-95166 and N01-HC-95169.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

1. Abedin M, Tintut Y, Demer LL. Vascular calcification: mechanisms and clinical ramifications. *Arterioscler Thromb Vasc Biol* 24: 1161–1170, 2004.
2. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR ,Jr, Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 156: 871–881, 2002.
3. Bild DE, Detrano R, Peterson D, Guerci A, Liu K, Shahar E, Ouyang P, Jackson S, Saad MF. Ethnic differences in coronary calcification: the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation* 111: 1313–1320, 2005.
4. Caira FC, Stock SR, Gleason TG, McGee EC, Huang J, Bonow RO, Spelsberg TC, McCarthy PM, Rahimtoola SH, Rajamannan NM. Human degenerative valve disease is associated with up-regulation of low-density lipoprotein receptor-related protein 5 receptor-mediated bone formation. *J Am Coll Cardiol* 47: 1707–1712, 2006.
5. Carr JJ, Nelson JC, Wong ND, McNitt-Gray M, Arad Y, Jacobs DR ,Jr, Sidney S, Bild DE, Williams OD, Detrano RC. Calcified coronary artery plaque measurement with cardiac CT in population-based studies: standardized protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) study. *Radiology* 234: 35–43, 2005.
6. Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, Morley M, Spielman RS. Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet* 33: 422–425, 2003.
7. Demer LL, Tintut Y. Vascular calcification: pathobiology of a multifaceted disease. *Circulation* 117: 2938–2948, 2008.
8. Detrano R, Guerci AD, Carr JJ, Bild DE, Burke G, Folsom AR, Liu K, Shea S, Szklo M, Bluemke DA, O'Leary DH, Tracy R, Watson K, Wong ND, Kronmal RA. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N Engl J Med* 358: 1336–1345, 2008.
9. Diez Roux AV, Detrano R, Jackson S, Jacobs DR ,Jr, Schreiner PJ, Shea S, Szklo M. Acculturation and socioeconomic position as predictors of coronary calcification in a multiethnic sample. *Circulation* 112: 1557–1565, 2005.
10. Doherty TM, Tang W, Detrano RC. Racial differences in the significance of coronary calcium in asymptomatic black and white subjects with coronary risk factors. *J Am Coll Cardiol* 34: 787–794, 1999.
11. Dumeaux V, Olsen KS, Nuel G, Paulsen RH, Borresen-Dale AL, Lund E. Deciphering normal blood gene expression variation—The NOWAC postgenome study. *PLoS Genet* 6: e1000873, 2010.
12. Edwards JR, Sun SG, Locklin R, Shipman CM, Adamopoulos IE, Athanasou NA, Sabokbar A. LIGHT (TNFSF14), a novel mediator of bone resorption, is elevated in rheumatoid arthritis. *Arthritis Rheum* 54: 1451–1462, 2006.
13. Folsom AR, Kronmal RA, Detrano RC, O'Leary DH, Bild DE, Bluemke DA, Budoff MJ, Liu K, Shea S, Szklo M, Tracy RP, Watson KE, Burke GL. Coronary artery calcification compared with carotid intima-media thickness in the prediction of cardiovascular disease incidence: the Multi-Ethnic Study of Atherosclerosis (MESA). *Arch Intern Med* 168: 1333–1339, 2008.
14. Fouque-Aubert A, Chapurlat R. Influence of RANKL inhibition on immune system in the treatment of bone diseases. *Joint Bone Spine* 75: 5–10, 2008.
15. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, Grossfeld PD, Srivastava D. Mutations in NOTCH1 cause aortic valve disease. *Nature* 437: 270–274, 2005.
16. Ghosh D, Chinaiyan AM. Covariate adjustment in the analysis of microarray data from clinical studies. *Funct Integr Genomics* 5: 18–27, 2005.

17. Greenland P, LaBree L, Azen SP, Doherty TM, Detrano RC. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *JAMA* 291: 210–215, 2004.
18. Hino S, Michiue T, Asashima M, Kikuchi A. Casein kinase I epsilon enhances the binding of Dvl-1 to Frat-1 and is essential for Wnt-3a-induced accumulation of beta-catenin. *J Biol Chem* 278: 14066–14073, 2003.
19. Huang CC, Fornage M, Lloyd-Jones DM, Wei GS, Boerwinkle E, Liu K. Longitudinal association of PCSK9 sequence variations with low-density lipoprotein cholesterol levels: the Coronary Artery Risk Development in Young Adults Study. *Circ Cardiovasc Genet* 2: 354–361, 2009.
20. Huang QY, Kung AW. The association of common polymorphisms in the QPCT gene with bone mineral density in the Chinese population. *J Hum Genet* 52: 757–762, 2007.
21. Ishida S, Yamane S, Nakano S, Yanagimoto T, Hanamoto Y, Maeda-Tanumura M, Toyosaki-Maeda T, Ishizaki J, Matsuo Y, Fukui N, Itoh T, Ochi T, Suzuki R. The interaction of monocytes with rheumatoid synovial cells is a key step in LIGHT-mediated inflammatory bone destruction. *Immunology* 128: e315–e324, 2009.
22. Jeong YH, Sekiya M, Hirata M, Ye M, Yamagishi A, Lee SM, Kang MJ, Hosoda A, Fukumura T, Kim DH, Saeki S. The low-density lipoprotein receptor-related protein 10 is a negative regulator of the canonical Wnt/beta-catenin signaling pathway. *Biochem Biophys Res Commun* 392: 495–499, 2010.
23. Lampe JW, Stepaniants SB, Mao M, Radich JP, Dai H, Linsley PS, Friend SH, Potter JD. Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke. *Cancer Epidemiol Biomarkers Prev* 13: 445–453, 2004.
24. Lee TC, O'Malley PG, Feuerstein I, Taylor AJ. The prevalence and severity of coronary artery calcification on coronary artery computed tomography in black and white subjects. *J Am Coll Cardiol* 41: 39–44, 2003.
25. Leibbrandt A, Penninger JM. RANK/RANKL: regulators of immune responses and bone physiology. *Ann NY Acad Sci* 1143: 123–150, 2008.
26. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotnik K, Ward M, Kholodov M, Zbicic K, Beck J, Kimelman M, Shevelev S, Preuss D, Yaschenko E, Graeff A, Ostell J, Sherry ST. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* 39: 1181–1186, 2007.
27. Manolio TA, Arnold AM, Post W, Bertoni AG, Schreiner PJ, Sacco RL, Saad MF, Detrano RL, Szklar M. Ethnic differences in the relationship of carotid atherosclerosis to coronary calcification: the Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis* 197: 132–138, 2008.
28. McClelland RL, Chung H, Detrano R, Post W, Kronmal RA. Distribution of coronary artery calcium by race, gender, and age: results from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation* 113: 30–37, 2006.
29. O'Leary DH, Polak JF, Wolfson SK Jr, Bond MG, Bommer W, Sheth S, Psaty BM, Sharrett AR, Manolio TA. Use of sonography to evaluate carotid atherosclerosis in the elderly. The Cardiovascular Health Study CHS Collaborative Research Group. *Stroke* 22: 1155–1163, 1991.
30. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet* 2: e190, 2006.
31. Polonsky TS, McClelland RL, Jorgensen NW, Bild DE, Burke GL, Guerci AD, Greenland P. Coronary artery calcium score and risk classification for coronary heart disease prediction. *JAMA* 303: 1610–1616, 2010.
32. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909, 2006.
33. Rumberger JA, Simons DB, Fitzpatrick LA, Sheedy PF, Schwartz RS. Coronary artery calcium area by electron-beam computed tomography and coronary atherosclerotic plaque area. A histopathologic correlative study. *Circulation* 92: 2157–2162, 1995.
34. Schisler JC, Charles PC, Parker JS, Hilliard EG, Mapara S, Meredith D, Lineberger RE, Wu SS, Alder BD, Stouffer GA, Patterson C. Stable patterns of gene expression regulating carbohydrate metabolism determined by geographic ancestry. *PLoS One* 4: e8183, 2009.
35. Staal FJ, Luis TC, Tiemessen MM. WNT signalling in the immune system: WNT is spreading its wings. *Nat Rev Immunol* 8: 581–593, 2008.
36. Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J, Akey JM. Gene-expression variation within and among human populations. *Am J Hum Genet* 80: 502–509, 2007.
37. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flieck P, Koller D, Montgomery S, Tavare S, Deloukas P, Dermitzakis ET. Population genomics of human gene expression. *Nat Genet* 39: 1217–1224, 2007.
38. Tan Q, Christensen K, Christiansen L, Frederiksen H, Bathum L, Dahlgaard J, Kruse TA. Genetic dissection of gene expression observed in whole blood samples of elderly Danish twins. *Hum Genet* 117: 267–274, 2005.
39. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, Doremioux O. PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucl Acids Res* 31: 334–341, 2003.
40. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98: 5116–5121, 2001.
41. Vu TH, Shipley JM, Bergers G, Berger JE, Helms JA, Hanahan D, Shapiro SD, Senior RM, Werb Z. MMP-9/gelatinase B is a key regulator of growth plate angiogenesis and apoptosis of hypertrophic chondrocytes. *Cell* 93: 411–422, 1998.
42. Wada T, Nakashima T, Oliveira-dos-Santos AJ, Gasser J, Hara H, Schett G, Penninger JM. The molecular scaffold Gab2 is a crucial component of RANK signaling and osteoclastogenesis. *Nat Med* 11: 394–399, 2005.
43. Wang Z, Neuburg D, Li C, Su L, Kim JY, Chen JC, Christiani DC. Global gene expression profiling in whole-blood samples from individuals exposed to metal fumes. *Environ Health Perspect* 113: 233–241, 2005.

Plasma metabolomic profiles predict near-term death among individuals with lower extremity peripheral arterial disease

Chiang-Ching Huang, PhD,^a Mary M. McDermott, MD,^b Kiang Liu, PhD,^a Ching-Hua Kuo, PhD,^c San-Yuan Wang, MS,^d Huimin Tao, MS,^a and Yufeng Jane Tseng, PhD,^{c,d,e} Chicago, Ill; and Taipei, Taiwan

Background: Individuals with peripheral arterial disease (PAD) have a nearly two-fold increased risk of all-cause and cardiovascular disease mortality compared to those without PAD. This pilot study determined whether metabolomic profiling can accurately identify patients with PAD who are at increased risk of near-term mortality.

Method: We completed a case-control study using ¹H NMR metabolomic profiling of plasma from 20 decedents with PAD, without critical limb ischemia, who had blood drawn within 8 months prior to death (index blood draw) and within 10 to 28 months prior to death (preindex blood draw). Twenty-one PAD participants who survived more than 30 months after their index blood draw served as a control population.

Results: Results showed distinct metabolomic patterns between preindex decedent, index decedent, and survivor samples. The major chemical signals contributing to the differential pattern (between survivors and decedents) arose from the fatty acyl chain protons of lipoproteins and the choline head group protons of phospholipids. Using the top 40 chemical signals for which the intensity was most distinct between survivor and preindex decedent samples, classification models predicted near-term all-cause death with overall accuracy of 78% (32/41), a sensitivity of 85% (17/20), and a specificity of 71% (15/21). When comparing survivor with index decedent samples, the overall classification accuracy was optimal at 83% (34/41) with a sensitivity of 80% (16/20) and a specificity of 86% (18/21), using as few as the top 10 to 20 chemical signals.

Conclusion: Our results suggest that metabolomic profiling of plasma may be useful for identifying PAD patients at increased risk for near-term death. Larger studies using more sensitive metabolomic techniques are needed to identify specific metabolic pathways associated with increased risk of near-term all-cause mortality among PAD patients. (J Vasc Surg 2013;■:1-8.)

Lower extremity peripheral arterial disease (PAD) affects eight million men and women in the United States and is expected to increase in prevalence as the U.S. population survives longer with chronic disease.¹ Men and women with PAD have a higher risk for all-cause and cardiovascular disease (CVD) mortality than men and women without PAD, even after adjusting for CVD risk factors and comorbidities.^{2,3} Despite optimal atherosclerotic disease risk factor control, many patients with PAD suffer from cardiovascular events. Conventional

cardiovascular risk factors are useful to predict long-term coronary events and mortality, but are less useful for predicting near-term events.⁴⁻⁷

Many acute coronary events result from plaque rupture and lumen thrombosis at arterial sites with minimally occlusive atherosclerosis.⁸ Circulating inflammatory and hemostatic biomarkers may contribute to plaque instability and rupture.⁹⁻¹¹ However, these associations are not clearly established. Identifying biomarkers that are elevated immediately prior to an acute coronary event could provide important short-term prognostic information and may elucidate mechanisms of acute coronary events.

Metabolites are the end products of complex interactions between the host genome and environmental stimuli. Thus, metabolic changes are the most proximal reporters of an individual's physiological status (such as an acute coronary event). Identifying novel metabolic pathways that are activated immediately prior to death may help establish new mechanisms of mortality that may be targeted with novel therapies to prevent adverse outcomes in patients with PAD. Metabolomics,¹² a relatively new "omics" system, provides an unbiased means to probe the full spectrum of the metabolic pattern. Metabolomics holds great potential to discover major chemical alterations in response to physiological changes and guide the selection of therapeutic targets.^{13,14} Thus, metabolomics represents a potentially powerful tool to capture molecular events that define distinct populations of patients and to attain a high prognostic accuracy for near-term events such as death.

From the Department of Preventive Medicine,^a and Department of Medicine,^b Feinberg School of Medicine, Northwestern University, Chicago; and the School of Pharmacy, College of Medicine,^c Department of Computer Science and Information Engineering,^d and Graduate Institute of Biomedical Electronics and Bioinformatics,^e National Taiwan University, Taipei.

This research was supported by the National Institutes of Health R01HL58099, R01HL64739, R01-HL073912, R01-HL083064, R01-HL071223, R01-HL073551 (MM) and Taiwan National Science Council, grants number 99-2627-B-002-018 and 99-2321-B-002-042 (JT).

Author conflict of interest: none.

Additional material for this article may be found online at www.jvacsurg.org. Reprint requests: Yufeng Jane Tseng, PhD, No. 1, Sec. 4, Roosevelt Rd, Taipei, 10617 Taiwan (e-mail: yjtseng@csie.ntu.edu.tw).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

0741-5214/\$36.00

Copyright © 2013 by the Society for Vascular Surgery.
<http://dx.doi.org/10.1016/j.jvs.2013.04.022>

We performed a pilot study to compare the proton nuclear magnetic resonance (^1H NMR) metabolomic profiles of plasma between PAD patients who died within 8 months of a blood draw and PAD patients who survived for at least 30 months after a blood draw. We hypothesized that metabolomic profiles of plasma can differentiate PAD patients with vs without increased risk of near-term death.

METHODS

Study participants and blood sample selections. The study protocol was approved by the institutional review boards of Northwestern University's Feinberg School of Medicine and participating study sites. Participants gave written informed consent. PAD participants were selected from existing cohorts of outpatients with PAD at Northwestern University including the Walking and Leg Circulation Study (WALCS), II, III cohorts, the Study to Improve Leg Circulation (SILC), and the Reducing Risk Factors in Peripheral Arterial Disease (RRF) study. Detailed descriptions of the WALCS, SILC, and RRF study designs and methods have been published previously.¹⁵⁻¹⁸ Across these studies, PAD was defined as an ankle-brachial index <0.95 , documented evidence of lower extremity revascularization, or documented clinical evidence of lower extremity atherosclerosis.

The current study consists of 20 PAD participants who died within 8 months of their last study blood draws (decedents) and 21 gender- and age-matched (within 3 years) PAD controls (survivors). Decedents died between May 1, 2000 and December 31, 2008. We selected all decedents from the WALCS, SILC, and RRF cohorts who met the following criteria: (a) presence of two blood draws (eg, index and preindex blood draws) prior to death; (b) the final blood draw (index blood draw) was within 1 year before death; and (c) the presence of a matched control survivor PAD participant. The index blood draw is defined as the blood draw that occurred closest to death (2-8 months preceding death). The preindex blood draw is defined as the blood draw at the study visit immediately prior to the index blood draw. The preindex blood draws occurred 7 to 25 months prior to the index blood draws. Each control PAD participant (survivor) was also selected from one of the same cohorts (WALCS I, II, or III, SILC, or RRF) and met the following criteria for inclusion in the study: a) presence of one blood sample that was matched for date of blood draw to within 3 months of the date of the decedent's index blood draw date; survival to at least 30 months (mean \pm standard deviation [sd], 67 \pm 28 months) after the date of their blood draws. There were a total of 61 blood samples for the metabolomics experiment, two for each decedent and one for each survivor.

Ascertainment of mortality. Deaths were identified using the Social Security Administration death database and by contact with participants' family members, proxies, and primary care physicians. Cause of death was determined using death certificates and by a certified nosologist. Cardiovascular deaths were those with International Classification of Disease-10 codes in the ranges I01.0 through

I99.9, including deaths due to coronary heart disease, stroke, peripheral vascular disease, and other CVD.

Plasma sample preparation. Blood was collected into Vacutainer (Becton Dickinson, Franklin Lakes, NJ) tubes containing EDTA and sodium citrate and immediately placed on ice. The plasma was separated immediately by centrifugation at 3000 rpm for 20 minutes at 4°C and samples were frozen and stored at -70°C.

Measurement of clinical variables. The clinical variables and blood biomarkers for the current study include ankle-brachial index, total and high density lipoprotein cholesterol levels, D-dimer, CRP, cigarette smoking, body mass index, and medical history. CRP and D-dimer data were only available for 16 survivors and 16 decedents from the WALCS and WALCS II cohorts. The details for calculation and acquisition of all these measures for each study have been reported previously¹⁵⁻²⁰ and are described briefly in the *Appendix* (online only).

Plasma ^1H NMR spectroscopy protocol and quality control. For each plasma sample, 550 μL of plasma and saline (0.9%) mixture was used for the ^1H NMR spectrum measurement at 600.13 MHz on a Bruker DRX-600 spectrometer (Bruker Corporation, Billerica, Mass). The water resonance was suppressed. The experimental temperature was set at 298 k.

^1H NMR spectra of plasma were acquired using both Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence and diffusion-edited sequence.²¹ Diffusion-edited sequence of the ^1H NMR spectrum characterized the lipid molecules of lipoproteins, such as the $-\text{CH}_3$ group of triglycerides, cholesterols, phospholipids, and glycerophospholipids to generate broad resonances in the spectrum. To characterize the small molecules such as creatinine, lactate, and glucose, we used the spectrometer with CPMG pulse sequence to detect the low molecular weight molecules by suppressing most of the broad resonances. The spectrum from both small molecules and lipoproteins generates an overall chemical structural profile.

Metabolomics data preprocessing, normalization, and statistical analysis. All spectra were phase-corrected and referenced to the chemical shift of alpha-glucose anomeric doublet at 5223 parts per million (ppm). Baseline correction was also applied to each spectrum using a penalized smoothing method.²² Each spectrum was segmented into 1000 chemical shift regions with a binning size 0.01 ppm in the range of chemical shift δ 0-10 for all data analysis. The peak region of water signal between δ 4.5 to 5.2 ppm was removed prior to statistical analysis.

Statistical analysis of plasma ^1H NMR spectral data was performed with R 2.10 (R Development Core Team; <http://www.R-project.org>) using several data analysis strategies, as described below. All binned NMR spectral data were logged and normalized with a quantile normalization procedure. Identification of differential chemical signals (ie, binned variables) was performed using two-sample or paired *t*-tests as appropriate. Detailed statistical methods for metabolomics data analysis regarding chemical signal selection for sample classification are described in the

Table. Baseline clinical characteristics of peripheral arterial disease (PAD) participants, according to whether they were decedents vs survivors

Variables	Decedent (n = 20)	Survivor (n = 21)	P value ^a
Age, years	75.9 ± 7.8	73.9 ± 8.0	.42
Male, %	80 (16/20)	81 (17/21)	.99
Ankle-brachial index	0.74 ± 0.24	0.73 ± 0.20	.94
Body mass index, kg/m ²	26.5 ± 5.3	25.3 ± 3.5	.38
Total cholesterol, mg/dL	168.8 ± 38.1	173.0 ± 22.9	.70
High-density lipoproteins, mg/dL	41.1 ± 11.0	43.2 ± 17.6	.67
CRP, ^b mg/L (range)	6.11 ± 6.75 (0.2-23.9)	5.32 ± 7.98 (0.2-33.4)	.75
D-dimer, ^b µg/L (range)	0.80 ± 0.47 (0.23-1.71)	0.77 ± 0.45 (0.32-1.92)	.81
Current smoker, %	30 (6/20)	14 (3/21)	.64
Hypertension, %	95 (19/20)	75 (15/20)	.18
Statin use, %	40 (8/20)	45 (9/20)	.99
Diabetes, %	55 (11/20)	41 (7/17)	.61
Stroke, %	17 (3/18)	0 (0/19)	.21
Angina, %	37 (7/19)	19 (4/21)	.37
Cancer, %	26 (5/19)	29 (6/21)	.99
Congestive heart failure, %	16 (3/19)	0 (0/21)	.20
Myocardial infarction, %	42 (8/19)	25 (5/20)	.43
Pulmonary disease, %	21 (4/19)	10 (2/21)	.56
Atherosclerotic vascular disease, ^c %	74 (14/19)	33 (7/21)	.03
Classic claudication symptoms, %	15 (3/20)	24 (5/21)	.69
Antiplatelet use, %	50 (10/20)	71 (15/21)	.27
Physical activity			
Number of blocks walked last week	28.3 ± 38.8	33.6 ± 46.7	.70
Number of stair flights climbed last week	6.2 ± 8.1	17.7 ± 19.1	.03

CRP, C-reactive protein.

^aP values were obtained from two-sample t-tests or χ^2 tests.^bCRP and D-dimer data were only available for 16 survivors and 16 decedents from the WALCS and WALCS II cohorts. WALCS, Walking and Leg Circulation Study.^cComprising stroke, angina, or myocardial infarction.

Appendix (online only). Baseline clinical characteristics between decedents and survivors with PAD were compared using t-tests for continuous variables and χ^2 tests for categorical variables.

Self-Organizing Map. We used the Self-Organizing Map (SOM) to create an ordered representation of multi-dimensional metabolomic data to reveal complex correlation structures among samples and metabolite patterns. An algorithm similar to Mäkinen et al²³ was used to perform SOM analyses. Briefly, we produced a map with 5 by 5 hexagonal units using a Gaussian neighborhood function (2.4 samples per unit on average). The SOM algorithm resulted in spectral models of neighboring units that are more similar to each other than those from the remote sites of the SOM. After the map was constructed, each of 61 spectra was allocated to its best-matching spectral model (eg, a hexagonal unit) on the SOM. Samples mapped to the same unit represent a cluster of closely related metabolomic profiles. The interpretation of the SOM is provided in the **Appendix** (online only). To observe the differences between spectral models, the metabolic spectrum of each hexagonal unit was represented according to the average intensity of the key substructures of lipid or small molecules in the spectra mapped to the certain unit within one unit radius.

RESULTS

Characteristics of PAD participants. The baseline clinical characteristics of the participants are shown in the

Table. There were no significant differences in the majority of the clinical variables between survivors and decedents. However, the prevalence of several comorbidities (eg, angina, stroke, myocardial infarction, and pulmonary disease) was higher among the decedents, and there is a significant difference between the decedents and survivors (74% vs 33%; $P = .03$) in active atherosclerotic vascular diseases (ie, a combination of stroke, angina, and myocardial infarction). Additionally, the number of stair flights walked was significantly higher in the survivor group (6.2 vs 17.7; $P = .03$). For the 20 decedents, seven died of coronary heart disease or CVD, six died of cancer, five died of non-coronary heart disease/CVD or noncancer causes, and two died of an unknown cause.

Metabolic patterns of decedents vs survivors. We observed a large number of spectrum signals with distinct intensity levels between the decedent index, decedent preindex, and survivor blood samples from the diffusion-edited spectra. Specifically, 26% of signals for index decedent samples vs survivor samples, 23% of signals for preindex decedent samples vs survivor samples, and 12% of signals for index decedent samples vs preindex decedent samples were significantly different at P values $<.05$. In contrast, no distinct CPMG ¹H NMR metabolomic patterns were observed between these three groups (data not shown). Thus, we focused on the diffusion-edited spectra data for the remaining analyses.

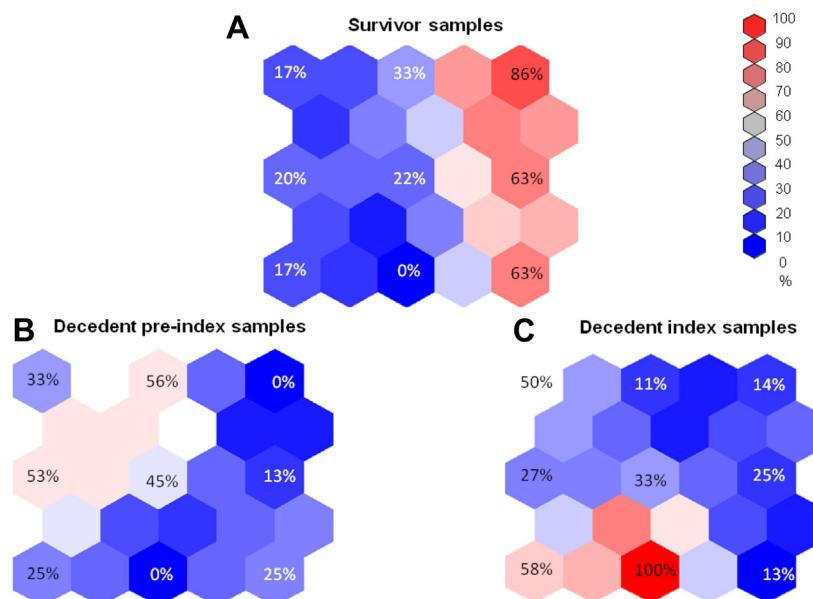


Fig 1. Self-Organizing Maps (SOMs) of 61 proton nuclear magnetic resonance (^1H NMR) spectra of plasma from peripheral arterial disease (PAD) patients with and without experiencing near-term death. Each hexagonal unit of the SOM represents a specific metabolic model spectrum and the color coding represents the percentage estimate of samples from a specific category (decedent index, decedent preindex, and survivor) mapped to the corresponding model spectrum. A-C represent the SOMs of 21 survivor samples, 20 decedent preindex samples, and 20 decedent samples, respectively.

The SOM analysis used 222 signals (false discovery rate <10%) of the diffusion-edited spectral data from the comparison between the decedent preindex samples and the survivor samples. As shown in Fig 1, most survivor samples shared metabolomic patterns on the upper-right hexagonal units of the SOM while most decedent preindex and decedent index samples had similar metabolomic patterns on the left and lower hexagonal units. For example, most survivor samples (86%; Fig 2) had metabolomic patterns characterizing metabolites from signals of the following chemical groups: $-\text{CH}_3$ (0.7 ppm), $(-\text{CH}_2-)_n$ (1.3 ppm), βCH_2 (1.45 ppm), $=\text{CH}-\text{CH}_2\text{CH}_2-$ (1.95 ppm), αCH_2 (2.2 ppm), and $\text{CH}-\text{CH}_2-\text{CH}=$ (2.82 ppm) of fatty acyl chain protons of various lipid types as part of lipoproteins as well as $-\text{N}(\text{CH}_3)_3$ (3.26 ppm) and NCH_2 (3.6 ppm) from the choline head group protons of phospholipids. The intensity levels of these chemical signals are significantly different between survivors and decedents (Fig 3). Relative to survivor samples, most of these chemical signals that have higher (or lower) mean intensity levels in the decedent preindex samples have even higher (or lower) mean intensity levels in the decedent index samples, although none of them reaches statistical significance ($P < .05$).

Sensitivity and specificity of metabolic profile for predicting near-term mortality. ^1H NMR metabolomic profiles provide a high predictive accuracy for the near-term all-cause deaths in these PAD participants. For the classification between decedent preindex samples and survivor samples, the overall classification performance was

optimal at 78% (32 out of 41 participants), with a specificity of 71% (15 out of 21 survivors) and a sensitivity of 85% (17 out of 20 decedents) using the top 40 to 150 chemical signals of the diffusion-edited data. The predictive accuracy was similar when decedent index and survivor samples were compared. Specifically, the overall classification success was optimal at 83% (34 out of 41 participants) with a specificity of 86% (18 out of 21 survivors) and a sensitivity of 80% (16 out of 20 decedents) using the top 10 to 20 chemical signals. The four decedents who were misclassified included two who died of cancer, one who died of CVD, and one who died of an unknown cause. Logistic regression shows that the metabolomic signature remained significantly associated with near-term all-cause death ($P < 10^{-4}$) after adjustment for active atherosclerotic diseases status.

DISCUSSION

Our results demonstrate that the global metabolomic profiling of plasma differed between 20 PAD participants who died within 8 months after their blood samples were obtained and 21 PAD participants who survived for at least 30 months after their blood samples were obtained. Importantly, the metabolomics profiling of decedent preindex samples shows that these differences are already detectable more than 1 year before death. Since death involves a global breakdown of the homeostatic regulation of multiple metabolic pathways that fails to counter the overwhelming destructive metabolic reactions,²⁴ the dramatically altered metabolic composition in decedents may reflect underlying

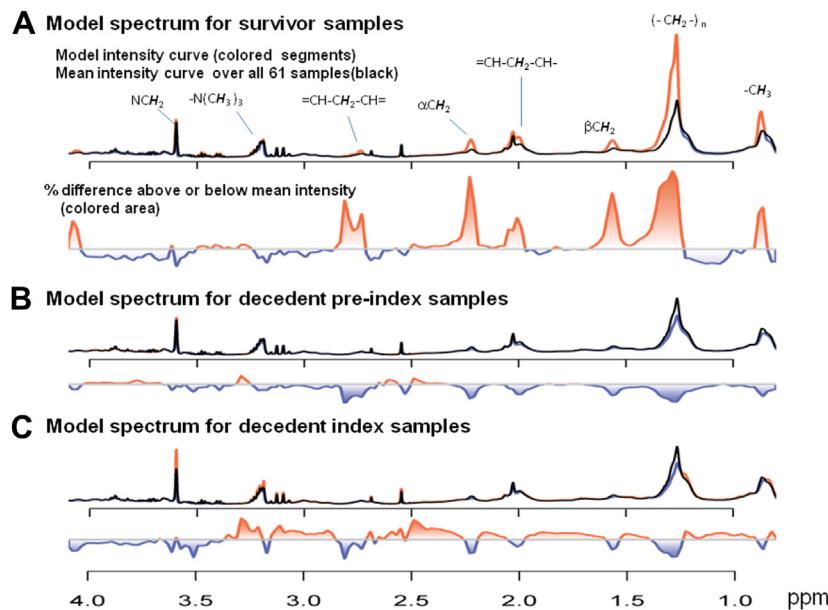


Fig 2. Three key metabolic model spectra of the Self-Organizing Map (SOM). **A**, The survivor model spectrum mapped to the grid (1,5) in Fig 3, A. **B**, The decedent preindex model spectrum mapped to the grid (3,1) in Fig 3, B. **C**, The decedent index model spectrum mapped to the grid (5,3) in Fig 3, C. The red curve indicates the current metabolic model; the black curve indicates the mean spectrum over all 61 spectra, thus serving as a constant reference. The proportional differences of each metabolic model spectrum and the mean spectrum are shown by the colored area below each model spectrum.

physiological changes leading to death. These findings suggest that metabolomic profiling of plasma may be a useful tool to delineate the pathophysiological and biologic mechanisms immediately preceding death.

Our results show that the major chemical signals differentiating survivors from decedents were derived from the diffusion-edited profiling but were not from the CPMG profiling, suggesting lipid metabolism may have a stronger association with near-term death than other small molecules such as carbohydrates or amino acids. Although specific molecular mechanisms may exist in different diseases, the lipid signature identified in the current study may represent a common pathway associated with mortality in different diseases. Evidence is emerging to support this assertion. For example, large epidemiological studies have found an increased risk of developing coronary heart disease and a variety of cancers in individuals with diabetes,^{25,26} suggesting the existence of common molecular mechanisms that are associated with the development of various chronic diseases. Additionally, our metabolomics finding is in line with a recent transcriptomic finding by Hirsch et al,²⁷ who identified an unexpected molecular link between lipid metabolism and diverse human diseases including cancer, atherosclerosis, diabetes, and autoimmune diseases. Structural analysis of our metabolomic data showed that the most significant chemical signals arose from lipoproteins and phospholipids. These two lipid groups can be converted through arachidonic acid metabolism to prostaglandins, which play a central role in platelet aggregation, vascular smooth muscle tone,

and inflammation.²⁸ This suggests that these two lipid groups have a fundamental role in plaque homeostasis or pathophysiological changes prior to death.²⁹ Experimental studies also demonstrated that inhibition of cyclooxygenase, the rate-limiting enzyme for conversion of arachidonic acid to important signal molecules (ie, prostaglandins), has a profound effect on cellular functions in many tissues, including tumor growth suppression and reduction of atherosclerosis.³⁰⁻³³ Since the cyclooxygenase pathway plays a critical role in inflammation, future investigation is needed to determine whether these two groups of lipid species (ie, those with fatty acyl chain protons of lipoproteins and the choline head group protons of phospholipids) are involved in this pathway and induce further inflammation or changes in other biological processes related to the pathophysiology of all-cause near-term death. These two groups of lipid species may also be associated with physical activity. The survivor group of our study's participants had better walking performance than the decedent group. A recent metabolomics analysis³⁴ of the Framingham Heart Study participants showed that baseline plasma concentration of metabolites in the lipolysis pathway that were altered in response to exercise was associated with cardiovascular fitness. Furthermore, several cohort and twin studies show that long-term leisure-time physical activity affects serum concentrations of a wide spectrum of lipoprotein subclasses, independent of body mass index and age.³⁵ Therefore, future studies will be needed to examine whether our observation in the alteration of lipid profiling is linked to physical activity – thereby

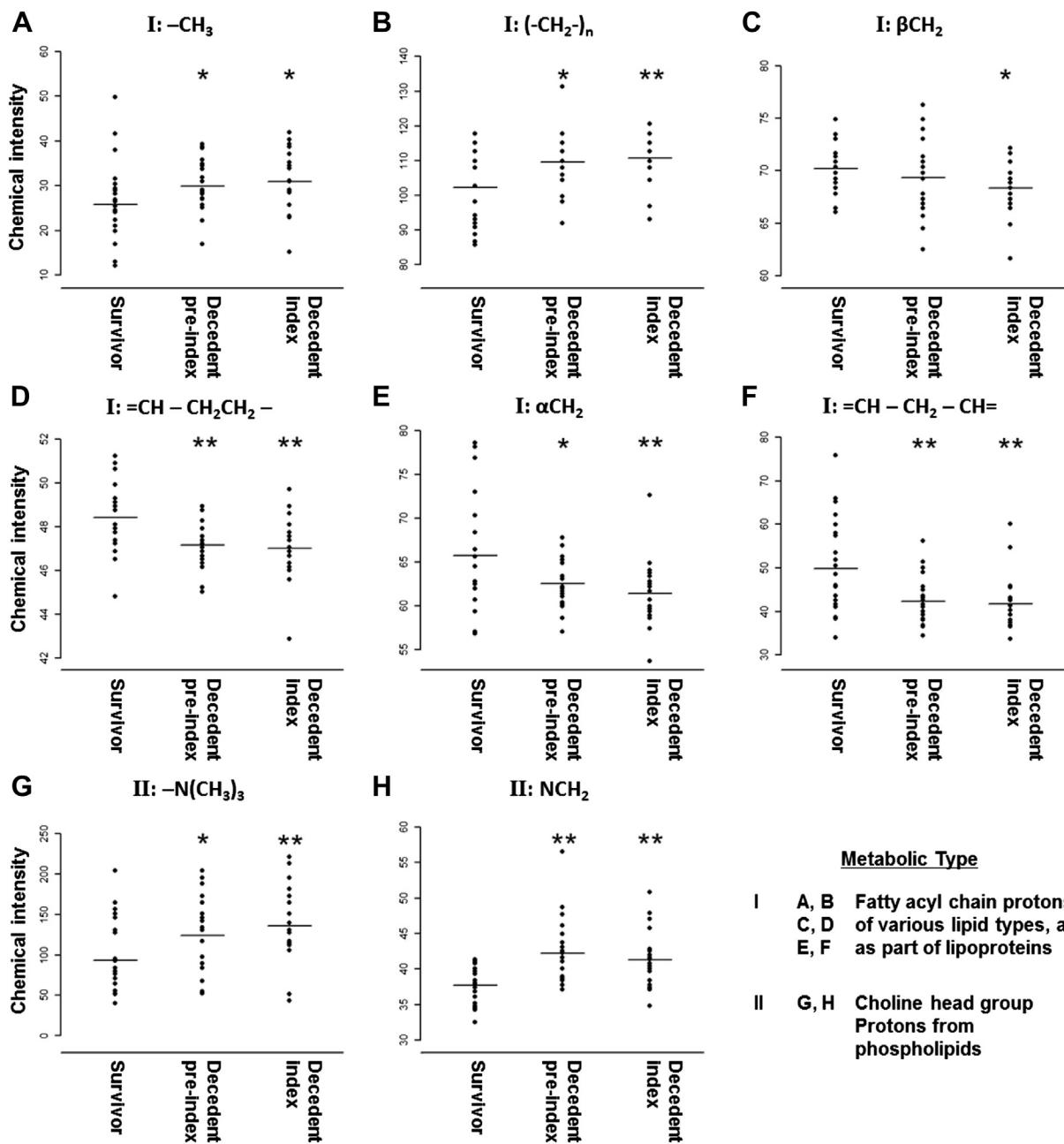


Fig 3. Concentration levels of substructures of lipid molecules from plasma samples of survivors and decedents. Quantified were: **A**, $-\text{CH}_3$ signal at 0.7 parts per million (ppm). **B**, $(-\text{CH}_2)_n$ signal at 1.3 ppm. **C**, βCH_2 signal at 1.45 ppm. **D**, $=\text{CH}-\text{CH}_2\text{CH}_2-$ signal at 1.95 ppm. **E**, αCH_2 signal at 2.2 ppm. **F**, $=\text{CH}-\text{CH}_2-\text{CH}=$ signal at 2.82 ppm. **G**, $-(\text{NH}_3)_3$ signal at 3.26 ppm. **H**, NCH_2 signal at 3.6 ppm. Statistically significant levels for comparison between survivor and preindex decedent and comparison between survivor and index decedent samples are indicated by red star: * $P < .05$; ** $P < .01$. Intensity levels are not significantly different between preindex and index decedent samples for all eight chemical signals.

identifying an enhanced secondary prevention strategy for the high-risk patients with PAD. The intensity levels of the majority of these chemical signals show linear increases or decreases from survivor to decedent preindex to decedent index samples (Fig 3). Intensity levels are significantly

different between survivors and decedent index samples and between survivor and decedent preindex samples (except for βCH_2), but are not significantly different between decedent preindex and index samples. Since both decedent metabolomics profiles are near-term

mortality measurements (18.4 vs 3.6 months), the relatively stable concentration of these metabolites over the 1 year or so before death among the decedents (but significantly different from survivors) suggests that these lipid signals are a robust metabolite signature to discriminate those at risk for imminent acute events including death. However, further prospective study is needed, including confirmation of these findings in a larger cohort of individuals with PAD. Identifying these specific metabolites using high-sensitive technologies such as mass spectrometry may help elucidate metabolic mechanisms of mortality and may lead to new therapies to improve longevity.

Some limitations of the current study should be acknowledged. First, the small sample size of our pilot study does not allow adjustment for other risk factors and clinical variables in our classification model. Similarly, our study has a low statistical power to evaluate the correlation between metabolites and known risk factors such as cholesterol levels, thereby limiting the clinical utility of the metabolomic profile in modifying the clinical course. Future studies with a larger sample size are needed to confirm the findings reported here. Second, our pilot study did not have all the clinical variables necessary to construct the 10-year Charlson comorbidity score.³⁶ Therefore, we are not able to assess the incremental predictive value of our metabolomics profiling for near-term mortality in addition to the Charlson comorbidity score. Third, ¹H NMR profiling is advantageous because it is capable of screening a large number of metabolites and revealing a global metabolomic pattern. However, the molecular species giving rise to the peaks are generally unknown, and this method tends to have a low sensitivity that can detect only highly abundant metabolites.³⁷ Further studies using technology with high precision and accuracy such as liquid chromatography mass spectrometry may identify unambiguously specific lipid metabolites and other nonlipid metabolites that may complement current existing biomarkers for near-term mortality. Fourth, study participants were outpatients with PAD. Our findings may not be generalizable to individuals at risk for PAD or those with critical limb ischemia. Fifth, we did not have data on diabetes mellitus severity. However, we are unaware of evidence demonstrating that more severe diabetes mellitus is associated with higher mortality in PAD patients without critical limb ischemia.

In summary, our result, as a proof-of-principle, shows that a specific metabolite signature from the lipid species correlated well with near-term death. To our knowledge, this is the first study to examine the utility of metabolomics profiling for near-term mortality prediction in any patient population. Our results suggest that metabolomic profiling may be a useful prognostic tool to identify individuals at high risk for near-term mortality. However, confirmation of these findings in a larger cohort is needed. In addition, further in-depth investigation of these metabolites is required to understand how the alteration of these metabolites' concentration influences the metabolic pathways that are activated during the days and weeks leading up to death.

AUTHOR CONTRIBUTIONS

Conception and design: MM, KL, YT
Analysis and interpretation: CH, CK, YT
Data collection: MM, KL, CK
Writing the article: CH, MM, YT
Critical revision of the article: CH, MM, YT
Final approval of the article: CH, YT
Statistical analysis: CH, SW, HT
Obtained funding: MM, YT
Overall responsibility: YT
CH and MM contributed equally to this work.

REFERENCES

- Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, et al. Heart disease and stroke statistics—2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2008;117:e25-146.
- Fowkes FG, Murray GD, Butcher I, Heald CL, Lee RJ, Chambliss LE, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA* 2008;300:197-208.
- Heald CL, Fowkes FG, Murray GD, Price JF. Risk of mortality and cardiovascular disease associated with the ankle-brachial index: systematic review. *Atherosclerosis* 2006;189:61-9.
- Desai MY, Nasir K, Braunstein JB, Rumberger JA, Post WS, Budoff MJ, et al. Underlying risk factors incrementally add to the standard risk estimate in detecting subclinical atherosclerosis in low- and intermediate-risk middle-aged asymptomatic individuals. *Am Heart J* 2004;148:871-7.
- Greenland P, Knoll MD, Stamler J, Neaton JD, Dyer AR, Garside DB, et al. Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. *JAMA* 2003;290:891-7.
- Khot UN, Khot MB, Bajzer CT, Sapp SK, Ohman EM, Brener SJ, et al. Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA* 2003;290:898-904.
- Magnus P, Beaglehole R. The real contribution of the major risk factors to the coronary epidemics: time to end the “only-50%” myth. *Arch Intern Med* 2001;161:2657-60.
- Rioufol G, Finet G, Ginon I, Andre-Fouet X, Rossi R, Vialle E, et al. Multiple atherosclerotic plaque rupture in acute coronary syndrome: a three-vessel intravascular ultrasound study. *Circulation* 2002;106:804-8.
- Moss AJ, Goldstein RE, Marder VJ, Sparks CE, Oakes D, Greenberg H, et al. Thrombogenic factors and recurrent coronary events. *Circulation* 1999;99:2517-22.
- Libby P. Current concepts of the pathogenesis of the acute coronary syndromes. *Circulation* 2001;104:365-72.
- Libby P, Ridker PM, Maseri A. Inflammation and atherosclerosis. *Circulation* 2002;105:1135-43.
- Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. Metabolomics applied to diabetes research: moving from information to knowledge. *Diabetes* 2009;58:2429-43.
- Bather TF, Sitter B, Sjøbakk TE, Tessem MB, Gribbestad IS. Magnetic resonance metabolomics of intact tissue: a biotechnological tool in cancer diagnostics and treatment evaluation. *Cancer Res* 2010;70:6692-6.
- Barba I, de Leon G, Martin E, Cuevas A, Aguade S, Candell-Riera J, et al. Nuclear magnetic resonance-based metabolomics predicts exercise-induced ischemia in patients with suspected coronary artery disease. *Magn Reson Med* 2008;60:27-32.
- McDermott MM, Ades P, Guralnik JM, Dyer A, Ferrucci L, Liu K, et al. Treadmill exercise and resistance training in patients with peripheral arterial disease with and without intermittent claudication: a randomized controlled trial. *JAMA* 2009;301:165-74.
- McDermott MM, Greenland P, Liu K, Guralnik JM, Celic L, Criqui MH, et al. The ankle brachial index is associated with leg function and physical activity: the Walking and Leg Circulation Study. *Ann Intern Med* 2002;136:873-83.

17. McDermott MM, Greenland P, Liu K, Guralnik JM, Criqui MH, Dolan NC, et al. Leg symptoms in peripheral arterial disease: associated clinical characteristics and functional impairment. *JAMA* 2001;286:1599-606.
18. McDermott MM, Mazor KM, Reed G, Pagoto S, Graff R, Merriam P, et al. Attitudes and behavior of peripheral arterial disease patients toward influencing their physician's prescription of cholesterol-lowering medication. *Vasc Med* 2010;15:83-90.
19. McDermott MM, Ferrucci L, Guralnik J, Tian L, Liu K, Hoff F, et al. Pathophysiological changes in calf muscle predict mobility loss at 2-year follow-up in men and women with peripheral arterial disease. *Circulation* 2009;120:1048-55.
20. McDermott MM, Liu K, Carr J, Criqui MH, Tian L, Li D, et al. Superficial femoral artery plaque, the ankle brachial index, and leg symptoms in peripheral arterial disease: the Walking and Leg Circulation Study (WALCS) III. *Circ Cardiovasc Imaging* 2011;4:246-52.
21. Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, et al. Metabolic profiling, metabolomic and metabolomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat Protoc* 2007;2:2692-703.
22. Xi Y, Rocke DM. Baseline correction for NMR spectroscopic metabolomics data analysis. *BMC Bioinformatics* 2008;9:324.
23. Makinen VP, Soininen P, Forsblom C, Parkkonen M, Ingman P, Kaski K, et al. ¹H NMR metabonomics approach to the disease continuum of diabetic complications and premature death. *Mol Syst Biol* 2008;4:167.
24. Marieb EN, Hoehn K. Human Anatomy & Physiology. 7th ed. San Francisco, CA: Pearson Benjamin Cummings; 2007.
25. Adami HO, McLaughlin J, Ekbom A, Berne C, Silverman D, Hacker D, et al. Cancer risk in patients with diabetes mellitus. *Cancer Causes Control* 1991;2:307-14.
26. Coughlin SS, Calle EE, Teras LR, Petrelli J, Thun MJ. Diabetes mellitus as a predictor of cancer mortality in a large cohort of US adults. *Am J Epidemiol* 2004;159:1160-7.
27. Hirsch HA, Iliopoulos D, Joshi A, Zhang Y, Jaeger SA, Bulyk M, et al. A transcriptional signature and common gene networks link cancer with lipid metabolism and diverse human diseases. *Cancer Cell* 2010;17:348-61.
28. Cipollone F, Cicolini G, Bucci M. Cyclooxygenase and prostaglandin synthases in atherosclerosis: recent insights and future perspectives. *Pharmacol Ther* 2008;118:161-80.
29. Santovito D, Mezzetti A, Cipollone F. Cyclooxygenase and prostaglandin synthases: roles in plaque stability and instability in humans. *Curr Opin Lipidol* 2009;20:402-8.
30. Earnest DL, Hixson LJ, Alberts DS. Piroxicam and other cyclooxygenase inhibitors: potential for cancer chemoprevention. *J Cell Biochem Suppl* 1992;16I:156-66.
31. Scioscia KA, Snyderman CH, Rueger R, Reddy J, D'Amico F, Comsa S, et al. Role of arachidonic acid metabolites in tumor growth inhibition by nonsteroidal antiinflammatory drugs. *Am J Otolaryngol* 1997;18:1-8.
32. Lin DW, Nelson PS. The role of cyclooxygenase-2 inhibition for the prevention and treatment of prostate carcinoma. *Clin Prostate Cancer* 2003;2:119-26.
33. Burleigh ME, Babaev VR, Oates JA, Harris RC, Gautam S, Riendeau D, et al. Cyclooxygenase-2 promotes early atherosclerotic lesion formation in LDL receptor-deficient mice. *Circulation* 2002;105:1816-23.
34. Lewis GD, Farrell L, Wood MJ, Martinovic M, Arany Z, Rowe GC, et al. Metabolic signatures of exercise in human plasma. *Sci Transl Med* 2010;2:33ra37.
35. Kujala UM, Makinen VP, Heinonen I, Soininen P, Kangas AJ, Leskinen TH, et al. Long-term leisure-time physical activity and serum metabolome. *Circulation* 2013;127:340-8.
36. Froehner M, Koch R, Litz RJ, Oehlschlaeger S, Twelker L, Hakenberg OW, et al. Detailed analysis of Charlson comorbidity score as predictor of mortality after radical prostatectomy. *Urology* 2008;72:1252-7.
37. Mayr M. Metabolomics: ready for the prime time? *Circ Cardiovasc Genet* 2008;1:58-65.

Submitted Oct 3, 2012; accepted Apr 15, 2013.

Additional material for this article may be found online at www.jvascsurg.org.

APPENDIX (online only)**MEASUREMENT OF CLINICAL VARIABLES**

Ankle-brachial index measurement. A hand-held Doppler probe (Pocket-Dop II, Nicolet Vascular, Golden, Colo) was used to measure systolic pressures in the right brachial artery, right dorsalis pedis and posterior tibial arteries, left dorsalis pedis and posterior tibial arteries, and left brachial artery.

Total and high-density lipoprotein cholesterol levels. Total cholesterol levels were measured by using enzymatic reaction with peroxidase-phenol-4-aminophenazone indicator reaction. High-density lipoprotein cholesterol levels were measured by using a direct enzymatic colorimetric assay.

Cigarette smoking, body mass index, and medical history. We determined smoking status by patient report using a structured interview. Body mass index was calculated as weight in kilograms dividing by height in meters squared. Diabetes status was determined based on patient self-report in the SILC and the RRF clinical trials. In the WALCS II, III cohorts, diabetes was considered present if the participant was taking insulin or an oral hypoglycemic agent or if a patient report of diabetes mellitus was confirmed by medical record review or a primary care physician questionnaire. Hypertension was defined as either a patient report of physician-diagnosed hypertension or a physician report of a history of hypertension on the primary care physician questionnaire. The patient report of a physician diagnosis of each comorbid condition was used in the SILC and RRF clinical trials. In WALCS cohorts, we obtained patient reports of physician-diagnosed comorbidities, but the data were also adjudicated from medical record review and a primary care physician questionnaire.

D-dimer and CRP levels. D-dimer levels were measured using an Asserachrom D-Di it (Diagnostica Stago, Asnière res-sur-Seine, France) with an enzyme-linked immunosorbent assay procedure. CRP levels were measured using an immunotchnique on the Behring BN II analyzer (Dade Behring, Wilmington, Del). CRP and

D-dimer data were only available for 16 survivors and 16 decedents from the WALCS and WALCS II cohorts.

Physical activity. Patient-reported physical activity was measured with a questionnaire derived from the Harvard Alumni Activity Survey.¹ Participants were asked, “During the last week, how many city blocks or their equivalent did you walk? Let 12 city blocks equal 1 mile.” and “In the last week, about how many flights of stairs did you climb up? A flight is 10 steps.”

Statistical methods for chemical signal selection and sample classification. To assess the probability of false positive findings, we used the positive False Discovery Rate method proposed by Storey.² For predicting all-cause deaths, the leave-one-out cross validation procedure with support vector machine classifiers was used to evaluate the classification accuracy. For each binned variable, a c statistic to evaluate its discriminative power for near-term death was calculated. Binned variables with the largest c statistics were sequentially added to support vector machine classifiers.

Interpretation of SOM. After the locations of the metabolomic spectra on the SOM were obtained, color coding was used for each hexagonal unit to represent the percentage of a specific category of spectra within one unit radius of a certain hexagonal unit. For example, if two spectra of survivor samples are mapped to a certain hexagonal unit, and eight spectra of survivor samples are mapped to the nearby units, then there were 10 (2 + 8) spectra of survivor samples mapped to the certain unit within one unit radius. If a total of 16 experimental spectra were mapped to the certain unit within one unit radius, the color value of the certain hexagon unit for survivor category is 63% (10/16).

REFERENCES

- Lee IM, Paffenbarger RS Jr, Hsieh CC. Time trends in physical activity among college alumni, 1962-1988. *Am J Epidemiol* 1992;135: 915-25.
- Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 2003;100:9440-5.