Kinetochore genes are coordinately up-regulated in human tumors as part of a FoxM1-related cell division program

Prathapan Thiru[a], David M. Kern[a,b], Kara L. McKinley[a,b], Julie K. Monda[a,b], Florencia Rago[a,b], Kuan-Chung Su[a], Tonia Tsinman[a], Defne Yarar[c], George W. Bell[a], and Iain M. Cheeseman[a,b]
[a]Whitehead Institute for Biomedical Research, Cambridge, MA 02142; [b]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; [c]Merrimack Pharmaceuticals, Cambridge, MA 02139

ABSTRACT **The key player in directing proper chromosome segregation is the macromolecular kinetochore complex, which mediates DNA–microtubule interactions. Previous studies testing individual kinetochore genes documented examples of their overexpression in tumors relative to normal tissue, leading to proposals that up-regulation of specific kinetochore genes may promote tumor progression. However, kinetochore components do not function in isolation, and previous studies did not comprehensively compare the expression behavior of kinetochore components. Here we analyze the expression behavior of the full range of human kinetochore components in diverse published expression compendia, including normal tissues and tumor samples. Our results demonstrate that kinetochore genes are rarely overexpressed individually. Instead, we find that core kinetochore genes are coordinately regulated with other cell division genes under virtually all conditions. This expression pattern is strongly correlated with the expression of the forkhead transcription factor FoxM1, which binds to the majority of cell division promoters. These observations suggest that kinetochore gene up-regulation in cancer reflects a general activation of the cell division program and that altered expression of individual kinetochore genes is unlikely to play a causal role in tumorigenesis.**

Monitoring Editor
Kerry S. Bloom
University of North Carolina

Received: Mar 28, 2014
Revised: May 5, 2014
Accepted: May 6, 2014

INTRODUCTION

Proper chromosome segregation during mitosis is central to maintaining genome integrity. Even subtle chromosome segregation defects result in inappropriate chromosome numbers (termed aneuploidy), which are observed in >70% of tumors and have been suggested to promote tumorigenesis (Holland and Cleveland, 2009). Thus it is critical to understand the perturbations that contribute to chromosome missegregation during tumorigenesis. Proper chromosome segregation requires the multiprotein kinetochore complex, which connects chromosomal DNA to spindle microtubule polymers to provide the structure and forces required to align and segregate the replicated sister chromatids (Cheeseman and Desai, 2008). Because most kinetochore proteins exist in multisubunit complexes, increased expression of an individual kinetochore gene could cause subunit imbalances and create dominant-negative defects that disrupt high-fidelity chromosome segregation. Indeed, previous studies reported multiple examples of individual kinetochore genes that are up-regulated in tumor samples relative to normal tissue controls. This includes the kinetochore-localized microtubule-binding protein Ndc80/highly expressed in cancer 1 (Hec1; Chen *et al.*, 1997; Hayama *et al.*, 2006), the spindle assembly checkpoint proteins Mad1 (Ryan *et al.*, 2012), Mad2 (Sotillo *et al.*, 2007), and Bub1 (Shigeishi *et al.*, 2001; Grabsch *et al.*, 2003), regulatory components, including Cdc20 (Mondal *et al.*, 2007) and Aurora B (Bischoff *et al.*, 1998; Vischioni *et al.*, 2006), and key structural kinetochore proteins, including KNL1/CASC5

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E14-03-0837) on May 14, 2014.

Address correspondence to: Iain M. Cheeseman (icheese@wi.mit.edu).

Abbreviations used: BRCA, breast invasive carcinoma; CCLE, Cancer Cell Line Encyclopedia; ChIP-Seq, chromatin immunoprecipitation followed by high-throughput DNA sequencing; ENCODE, Encyclopedia of DNA Elements; expO, Expression Project for Oncology; GEO, Gene Expression Omnibus; HBI, Human Body Index; RMA, robust multiarray average; TCGA, The Cancer Genome Atlas; TFBS, Transcription Factor Binding Site.

Supplemental Material can be found at:
http://www.molbiolcell.org/content/suppl/2014/05/12/mbc.E14-03-0837.DC1.html

(Takimoto *et al.*, 2002), CENP-A (Tomonaga *et al.*, 2003), CENP-H (Tomonaga *et al.*, 2005; Shigeishi *et al.*, 2006), and many others.

The observations that selected kinetochore genes are up-regulated in tumors has led to proposals that this up-regulation may act as a causal event for tumor initiation or progression (Yuen *et al.*, 2005). In a subset of cases, this hypothesis has been tested using mouse models in which the expression of a given kinetochore gene is artificially up-regulated. Strikingly, up-regulation of Mad2 (Sotillo *et al.*, 2007), Ndc80/Hec1 (Diaz-Rodriguez *et al.*, 2008), or Aurora B (Ricke *et al.*, 2011) can increase the incidence of cancer formation in mice. Although these studies demonstrate that up-regulation of a kinetochore gene has the potential to promote tumorigenesis, the physiological relevance of these up-regulation events in human tumors remains unclear.

Here we seek to evaluate comprehensively kinetochore gene expression in diverse, publicly available data sets to determine the relative expression behavior of kinetochore genes. Our results demonstrate that core kinetochore genes are coordinately expressed under all analyzed conditions, including in normal human tissues, tumors, and cancer cell lines. In addition, kinetochore genes show similar expression behavior to other cell division genes. Most kinetochore genes do not display strong periodic mitosis-specific expression during the cell cycle, indicating that the observed up-regulation is not due solely to an increased proportion of mitotic cells in tumors. Instead, this expression behavior appears to reflect the induction of a broad but specific cell division program. To define key upstream factors involved in the induction of this cell division program, we analyze the relationship between kinetochore gene expression and established transcription factors. Our data suggest that this expression program is closely related to the behavior of the forkhead-family transcription factor, FoxM1. Thus kinetochore genes in tumors are coordinately expressed as part of a general activation of the cell division program that appears to function downstream of FoxM1.

RESULTS

Selection of kinetochore and cell division genes for expression analysis

To analyze the relative expression of kinetochore genes, we first generated a list of genes that could be used to assess expression profiles across diverse cell types. Work from our laboratory and many others has identified >100 different kinetochore-localized proteins in human cells (Cheeseman and Desai, 2008). These kinetochore proteins can be grouped into two broad functional categories: 1) core kinetochore proteins, which function exclusively at kinetochores; and 2) multifunctional kinetochore proteins, which function at kinetochores and additional subcellular locations (e.g., the microtubule-based motor cytoplasmic dynein or the nucleoporin Nup107-160 complex). To analyze the expression of other key genes required for dividing cells, we also selected genes involved in DNA replication and cell cycle regulation. Finally, to provide a comparison to both normal cellular processes and cancer states, we selected several representative housekeeping genes and genes whose up-regulation has been implicated in tumor progression (e.g., Myc and Ras). In total, we defined five functional categories that we refer to as follows: Core Kinetochore, Multifunctional Kinetochore, Cell Cycle/DNA Replication, Housekeeping, and Cancer Up-regulated. The selected genes and the corresponding categories are given in Supplemental Table S1.

Kinetochore genes display coordinated expression in normal tissues

We first assessed the expression of our selected genes in normal tissue samples using data from the Human Body Index (HBI; GEO Accession No. GSE7307; www.ncbi.nlm.nih.gov/geo). We normalized the expression of each gene relative to its median expression level across the entire set of samples and plotted these data as a heat map to indicate the level of up-regulation or down-regulation relative to this median expression (Figure 1). In this analysis, the mRNA levels of the Core Kinetochore genes were positively correlated (correlation of 0.44), representing coordinate expression across the entire selection of tissue samples. The highest levels of expression for these genes were observed in testes, bone marrow, and other tissues believed to show higher overall levels of cell division. Core Kinetochore components were expressed at lower levels in terminally differentiated tissues, including brain and neuronal samples. In this analysis, our selected Cell Cycle/DNA Replication genes displayed similar expression profiles to Core Kinetochore genes across these diverse samples. In contrast, as expected based on the role of the cytoplasmic dynein motor in axonal transport, we found that some Multifunctional Kinetochore genes, including dynein and its multiple associated proteins, were up-regulated in brain and neuronal samples. On the basis of the multiple cellular functions for these Multifunctional Kinetochore proteins, as well as their distinct and diverse behaviors in the data sets analyzed later (unpublished data), for simplicity we excluded these genes from subsequent figures. This analysis indicates that the selected Core Kinetochore and cell division genes are coordinately expressed across noncancerous human tissue samples.

Cell cycle expression of kinetochore genes

In the human tissue samples analyzed, Core Kinetochore gene expression appeared to correlate with the presumptive cell division frequency (mitotic index) of a given tissue. If these genes are up-regulated during the mitotic phase of the cell cycle, the increased expression observed in these samples could be due to an increased proportion of mitotic cells. Similarly, if Cell Cycle genes are expressed in a specific cell cycle stage, their increased expression in frequently dividing tissues could be due to increased occupancy of that cell cycle stage. Therefore we sought to assess the relationship between the expression of these genes and cell cycle stage. To test the relationship between Core Kinetochore gene expression and cell cycle progression, we analyzed periodic gene expression in a data set that analyzed time points in synchronously dividing human cells (Grant *et al.*, 2013). We found that a subset of Core Kinetochore genes was periodically expressed during the cell cycle (Figure 2). These periodically expressed genes are enriched for outer kinetochore proteins and regulatory components that localize to kinetochores only during mitosis. However, the majority of Core Kinetochore genes and other cell division components did not show dramatic differences in their expression behavior during the cell cycle (Figure 2). This suggests that the up-regulation of Core Kinetochore and Cell Cycle genes observed in frequently dividing normal tissues represents a general activation of the cell division program rather than expression due to increased occupancy of a specific cell cycle stage.

Kinetochore genes are up-regulated in tumors

Multiple studies demonstrated that individual kinetochore genes are up-regulated in tumors relative to normal tissue controls (see *Introduction*). To test whether this up-regulation is broadly true for other Core Kinetochore and Cell Cycle/DNA Replication

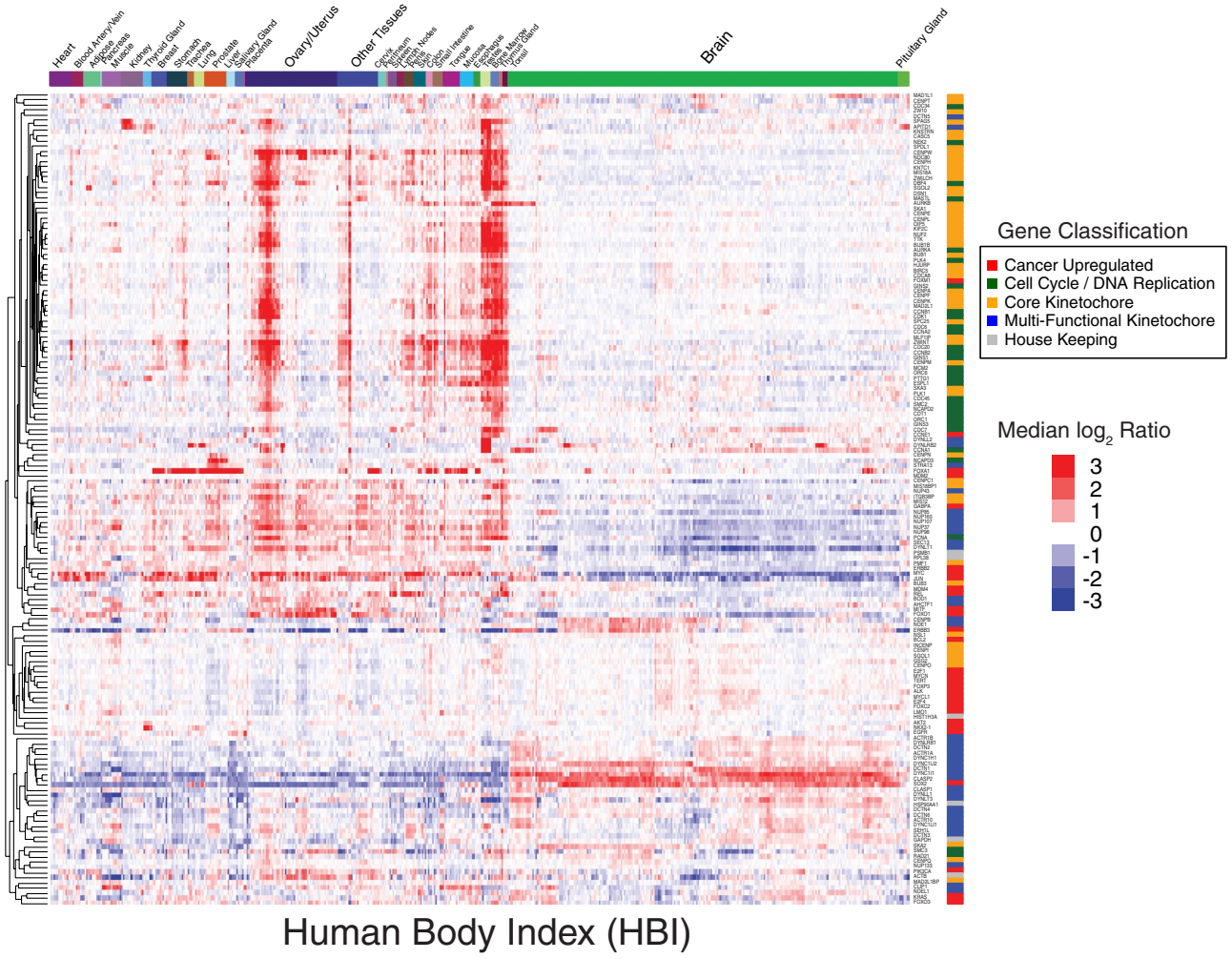

FIGURE 1: Kinetochore gene expression across normal human tissues. Gene expression data were obtained from the HBI (GEO Accession No. GSE7307). Selected genes were median centered, with changes in expression level color coded as indicated. Multiple samples within a given tissue category (e.g., Brain) were grouped and clustered. To the left of the expression data, a hierarchical clustering tree indicates the relationship between genes based on expression patterns. To the right of the expression data, genes are classified based on their functional category.

components, we analyzed their gene expression in The Cancer Genome Atlas (TCGA) study of breast cancer (Cancer Genome Atlas Network, 2012), in which matched breast tumor samples and normal tissue controls were analyzed for 28 different patients (Figure 3). For this analysis, gene expression in each individual tumor sample was normalized relative to the level of gene expression in the corresponding matched normal tissue sample. In this analysis, half of the Core Kinetochore and Cell Cycle/DNA Replication genes (36 of 71 genes) were significantly up-regulated (false discovery rate $p < 0.05$ and at least twofold change, *t* test) in the tumor samples relative to the normal tissue controls. These data suggest that tumors broadly activate genes required for cell division rather than up-regulate individual cell division genes.

Kinetochore genes are coordinately regulated in diverse tumor and cancer samples

The matched tumor/normal tissue samples provide an important comparison for evaluating the up-regulation of kinetochore genes. However, these paired experiments are of limited sample size, preventing an extensive analysis of the coordinated behavior of kinetochore genes across different conditions. Therefore we analyzed the relative expression of our selected genes in the Broad Cancer Cell Line Encyclopedia (Figure 4), which contains expression data for 991 cancer cell lines (www.broadinstitute.org/ccle; Barretina *et al.*, 2012), and the Expression Project for Oncology (expO) data set (Figure 5), which contains 2158 tumor samples (GEO Accession No. GSE2109). For these analyses, we first defined the median expression level for a given gene across all of the samples in a data set. We then compared its expression level in each sample to that median expression level to define the relative level of up- or down-regulation. This normalized expression level for a given gene across the diverse samples allowed us to evaluate the variation in gene expression among tumor populations. Based on this methodology, half of the samples should display reduced expression relative to this median level. However, based on the analysis of the matched TCGA breast cancer samples (Figure 3), each kinetochore gene is likely to be up-regulated relative to normal tissue.

Our analysis of gene expression behavior in tumor and cancer samples demonstrated that there is a strong similarity between changes in the mRNA levels of the Core Kinetochore components and Cell Cycle/DNA Replication genes (Figures 4 and 5). This expression profile was distinct from the other classes of genes

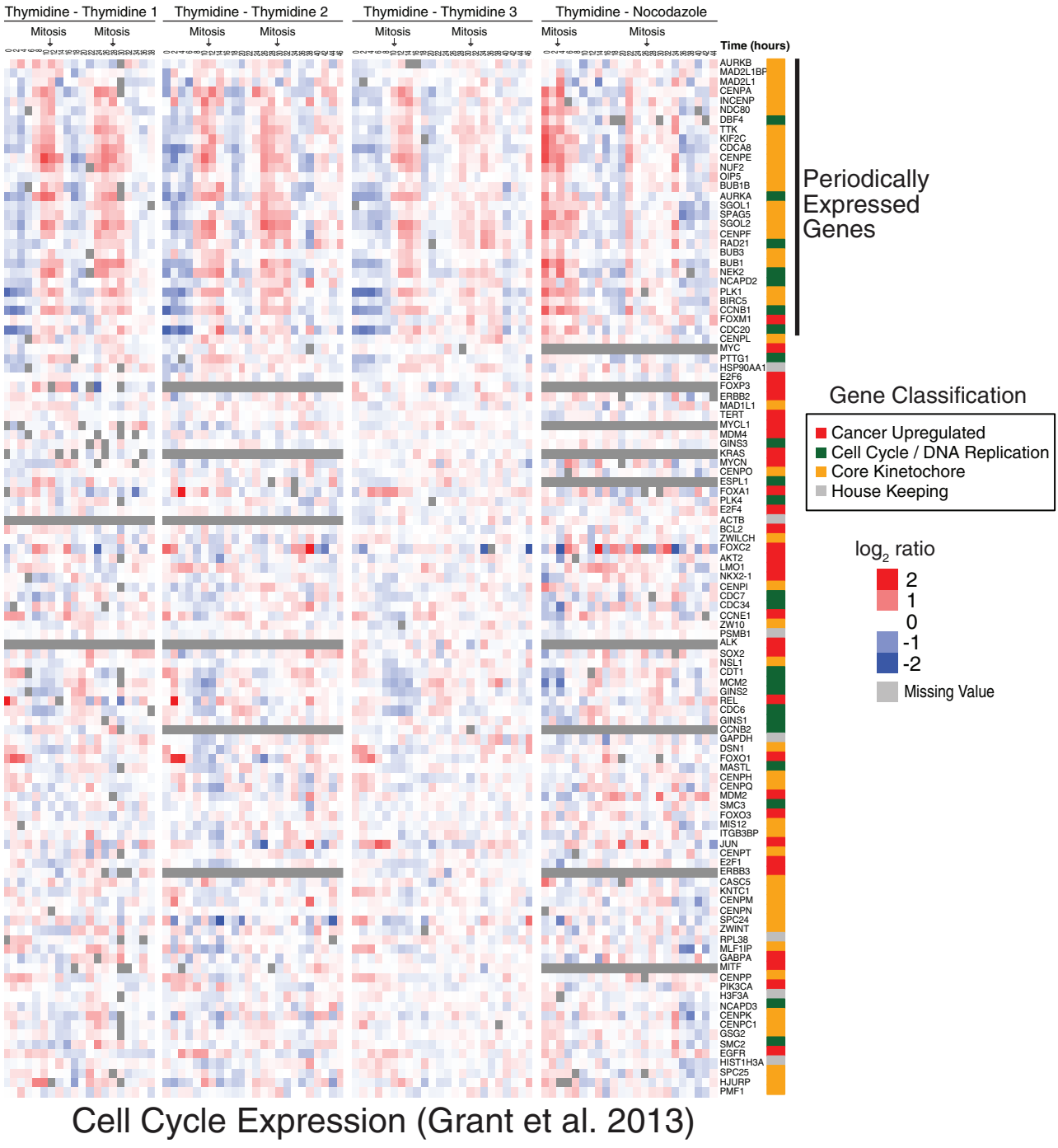

FIGURE 2: Kinetochore gene expression throughout the cell cycle. Gene expression data were obtained from Grant *et al.* (2013), who analyzed synchronous cell cycles in U20S cells blocked and released with the time points as indicated (top). Selected genes were displayed relative to the reference expression level, with changes in expression level color coded as indicated. To the right of the expression data, genes are classified based on their functional category. Periodically expressed genes are clustered at the top as indicated. Arrows indicate approximate mitotic times based on cell synchronization (Grant *et al.*, 2013).

included in our analysis, including Housekeeping genes and classically defined Cancer Up-regulated genes. In contrast to previous proposals, Core Kinetochore genes are not differentially expressed but are instead largely coordinately expressed across diverse tumors and cancers. In fact, we observed only a few, rare cases in which an analyzed Core Kinetochore gene showed clearly uncoordinated expression behavior. This includes increased expression of CENP-N in prostate tumors in the expO data set (Figure 5) and increased expression of Mad1 (Mad1L1) and p31/Comet (Mad2L1BP) in skin cancer cell lines from the CCLE data set (Figure 4). Therefore most individual kinetochore genes are not up-regulated in isolation, but instead their expression represents the broad induction of Core Kinetochore and Cell Cycle genes.

Although we found that Core Kinetochore genes generally behaved similarly to each other, as a group they exhibited variation in mRNA levels across the different tumors and cancer cell lines.

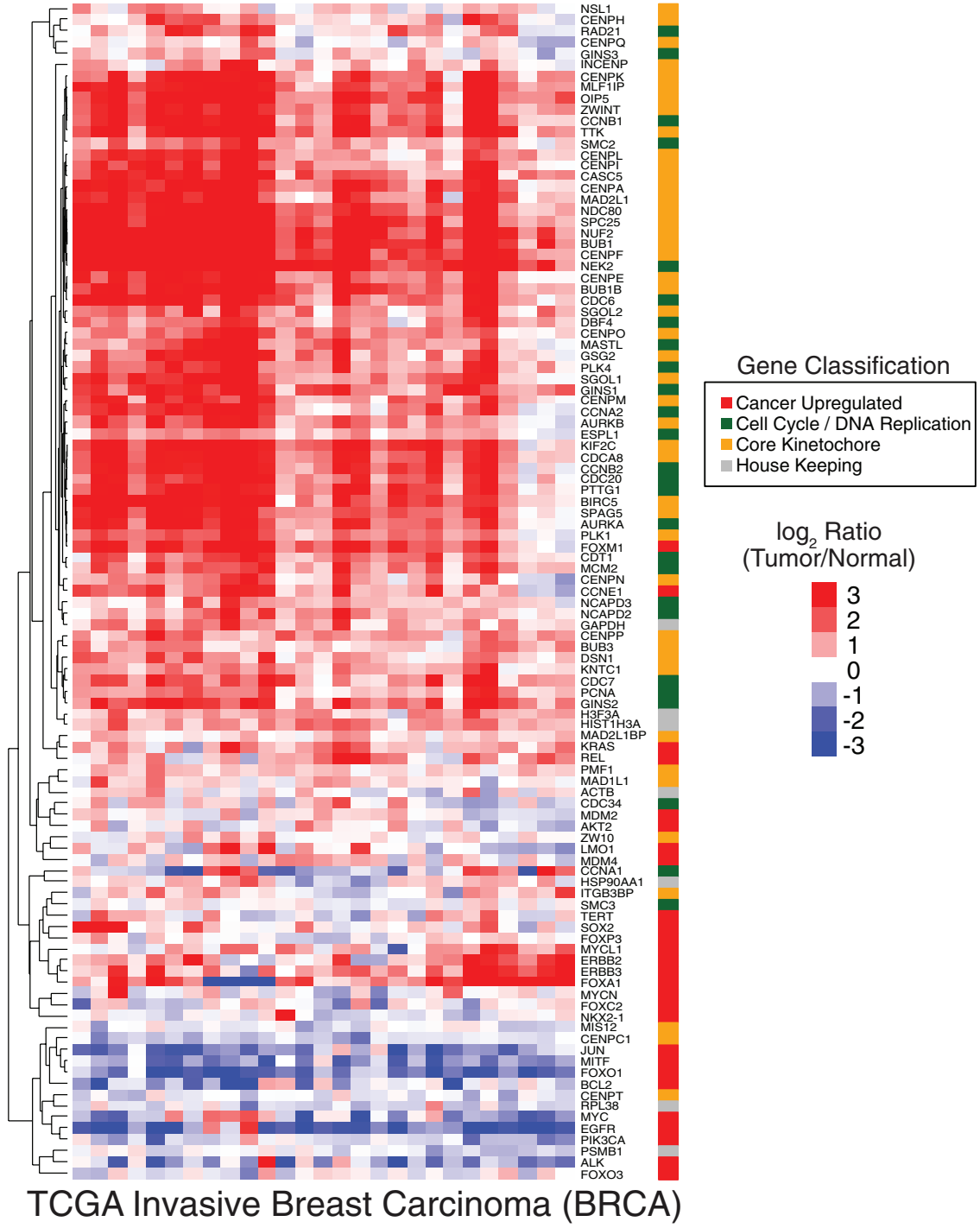

FIGURE 3: Kinetochore gene expression in tumors vs. normal tissue. Gene expression data were obtained from the TCGA study of invasive breast cancer (Cancer Genome Atlas Network, 2012). Expression levels of the selected genes in tumor samples were compared with the expression levels in the corresponding matched normal tissue control, with changes in expression level color coded as indicated. To the left of the expression data, a hierarchical tree indicates the relationship between expression patterns. To the right of the expression data, genes are classified based on their functional category.

We therefore sought to determine whether this relative expression behavior correlated with different tumor properties or behaviors. We did not detect an obvious correlation between the level of Core Kinetochore gene expression and the pathological stage of the tumor in the expO tumor data (unpublished data). Similarly, we did not observe any noticeable difference in the expression behavior in those tumors in the expO data set from patients who smoked relative to nonsmokers (unpublished data). In addition, we did not detect differences in the sensitivity of the cell lines from the CCLE data set to antimitotic drugs, including Taxol, based on the varying Core Kinetochore gene expression levels (unpublished data).

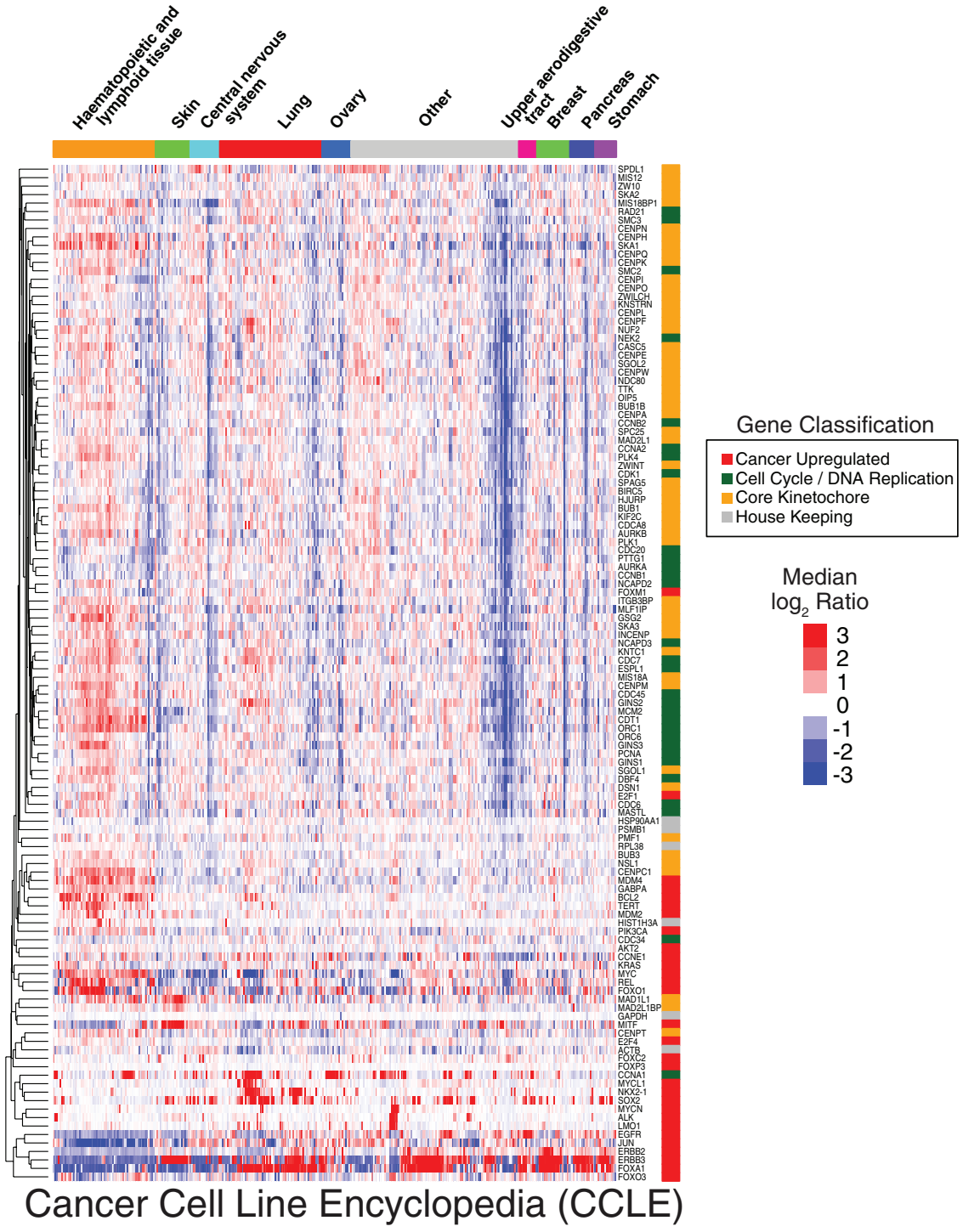

FIGURE 4: Kinetochore gene expression in cancer cell lines. Gene expression data were obtained from the CCLE (Barretina *et al.*, 2012). Selected genes were median centered, partitioned into groups by tissue of origin (as shown across the top), and clustered within that group, with changes in expression level color coded as indicated. To the left of the expression data, a hierarchical tree indicates the relationship between expression patterns. To the right of the expression data, genes are classified based on their functional category.

In contrast, we observed a clear relationship between the mRNA level and the identity of the tissue from which the tumor or cell line was derived. For example, much higher levels of Core Kinetochore and Cell Cycle/DNA Replication genes were observed in colon, cervix, and uterus samples, with lower relative levels present in kidney, prostate, and thyroid samples (Figures 4 and 5). Because matched normal/tumor samples are not available in the expO data set, we compared the expO tumor expression data to the tissue-specific expression levels from the HBI. Based on this comparison, the expression of Core Kinetochore genes is up-regulated to a roughly similar extent in these diverse tumor samples relative to the corresponding tissue of origin (Figure 6). Thus these differing expression levels primarily reflect the basal state of the starting tissue rather than a specific feature of tumorigenesis in a given tissue.

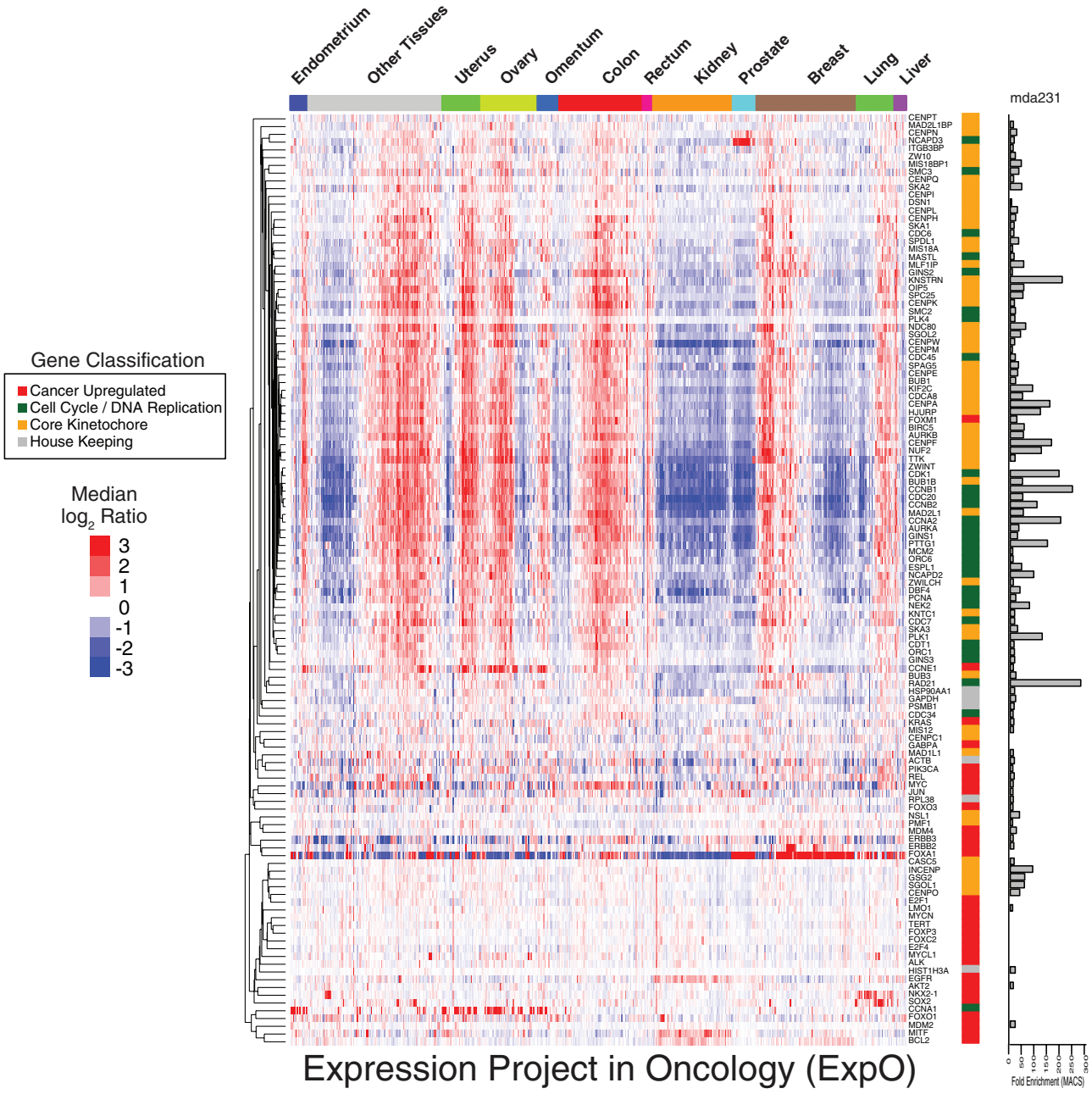

FIGURE 5: Kinetochore gene expression in tumors. Gene expression data were obtained from the expO (GEO Accession No. GSE2109). Selected genes were median centered, partitioned into groups by tissue of origin (as shown across the top), and clustered within that group, with changes in expression level color coded as indicated. To the left of the expression data, a hierarchical tree indicates the relationship between expression patterns. To the right of the expression data, genes are classified based on their functional category. To the far right, the bar graph indicates the relative fold enrichment of FoxM1 binding to the promoter of each gene based on ChIP-seq data (from ENCODE Project Consortium *et al.*, 2012).

Kinetochore protein interaction partners display closely coordinated expression

The similar expression behavior observed in tumors and cancer cell lines for the Core Kinetochore genes has important implications for the consequences of this altered expression. Because most kinetochore proteins exist in multisubunit protein complexes, the relative protein levels of the different subunits, rather than the precise amounts of a given subunit, will affect the behavior of that complex. Therefore we sought to analyze the statistical correlation for the gene expression behaviors of Core Kinetochore genes. Overall we observed highly correlated expression profiles for kinetochore genes in the CCLE cancer cell line data and expO tumor data set (Figure 7).

We also sought to analyze the expression of specific kinetochore protein complexes for which up-regulation in cancers has been reported previously. Ndc80/Hec1 assembles as part of a four-subunit Ndc80 complex that includes Nuf2, Spc24, and Spc25. Ndc80/Hec1 has been shown to be up-regulated in tumors (Chen *et al.*, 1997; Hayama *et al.*, 2006), and mice overexpressing Ndc80 display increased cancer formation (Diaz-Rodriguez *et al.*, 2008). If the Ndc80

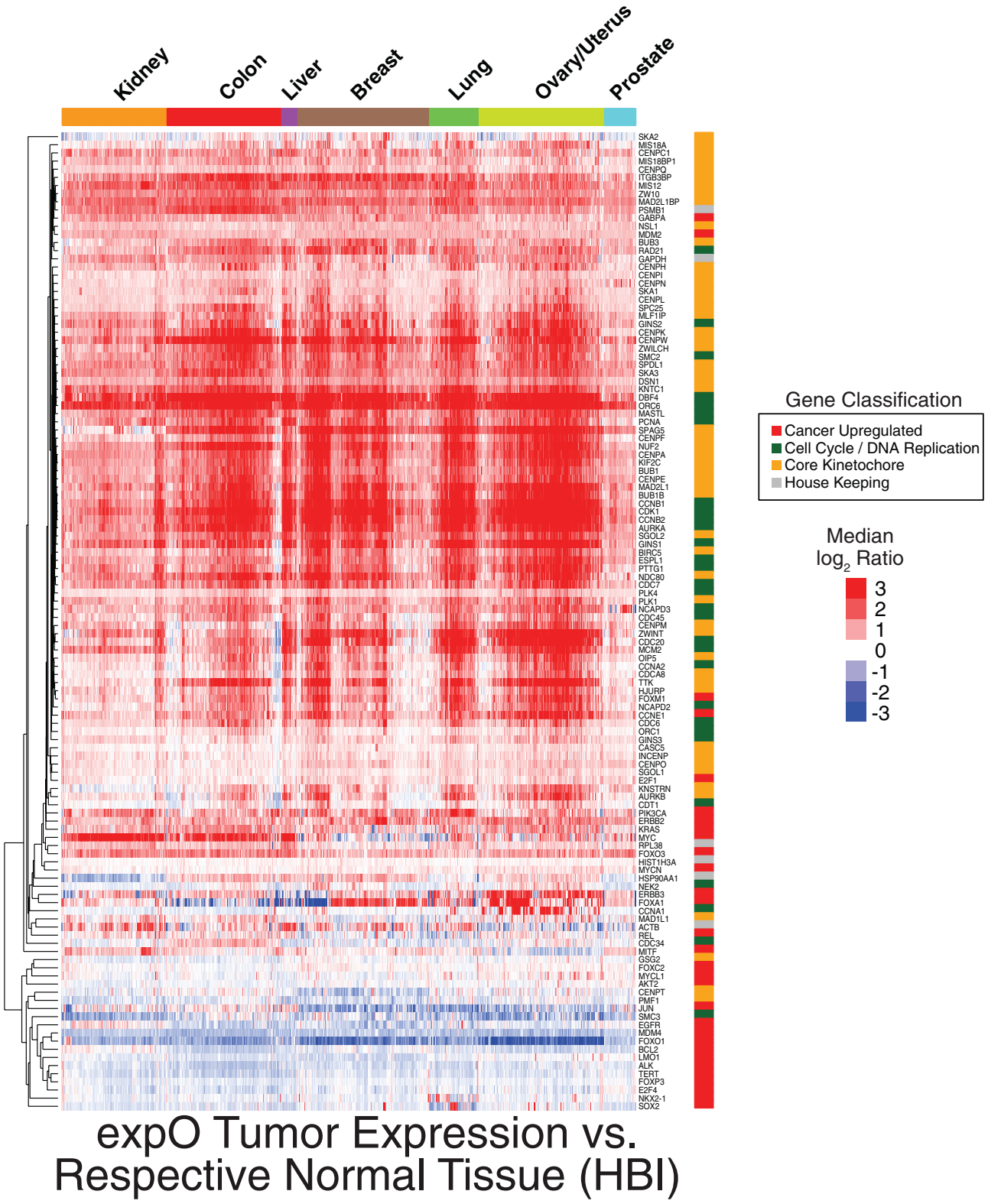

FIGURE 6: Tumor gene expression normalized relative to corresponding normal tissue. Gene expression data from the expO tumor data set were compared with the median expression in the corresponding normal tissue from the HBI data set. For example, kidney tumors were compared with the median expression in normal kidney tissues. These data suggest that tumors display similar degrees of up-regulated expression of kinetochore genes. To the left of the expression data, a hierarchical tree indicates the relationship between expression patterns. To the right of the expression data, genes are classified based on their functional category.

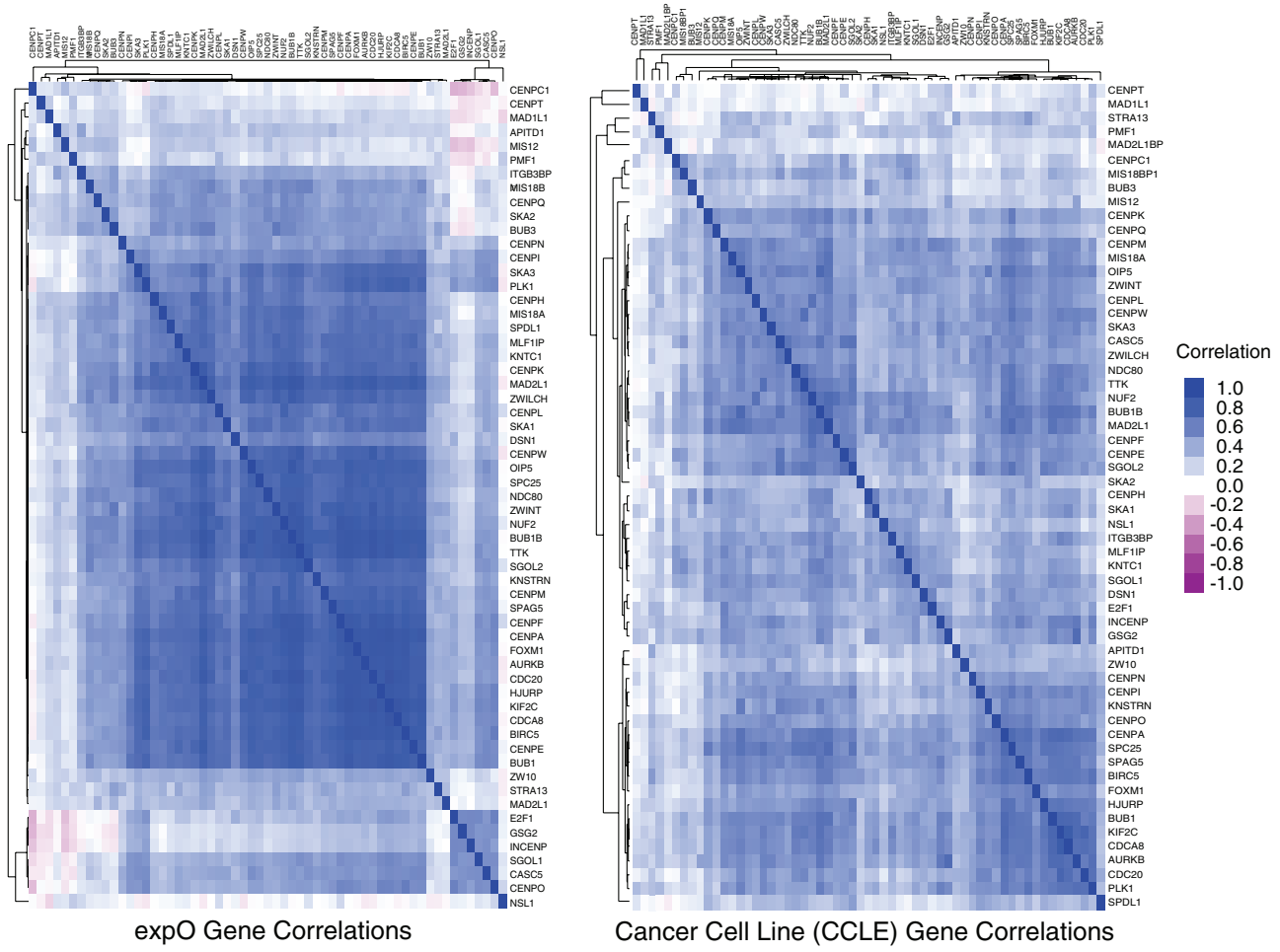

FIGURE 7: Correlations for kinetochore gene expression data from the CCLE and expO data sets. To assess the correlations in the gene expression profiles of the kinetochore genes (as well as the FoxM1 and E2F1 transcription factors), pairwise comparisons between the expression profiles for each pair of genes were tested in the CCLE and expO data sets. To the left of the expression data, a hierarchical tree indicates the relationship between gene correlations. Overall kinetochore genes display highly correlated gene expression profiles in tumors and cancer cell lines.

protein were specifically up-regulated in the absence of its binding partners, this would cause severe defects with the potential for dominant-negative consequences to the fidelity of chromosome segregation. In contrast, coordinate mRNA up-regulation of the Ndc80 complex subunits would avoid defects due to unbalanced subunit components. On the basis of our analysis, we found similar expression behavior for Ndc80, Nuf2, and Spc25 in cancer cell lines and tumors (correlation of 0.64 in CCLE data and 0.78 in expO data; the Spc24 probeset was unavailable for this analysis).

Similarly, previous work implicated overexpression of the spindle assembly checkpoint protein Mad2 (MAD2L1) as a potential causal factor in promoting tumorigenesis (Sotillo *et al.*, 2007). Although artificial overexpression of Mad2 can induce tumor formation in mice (Sotillo *et al.*, 2007), it is unclear whether such up-regulation of Mad2 alone occurs in human tumors. Cells are exquisitely sensitive to the level of Mad2, with subtle changes in concentration altering spindle assembly checkpoint signaling and mitotic progression (Heinrich *et al.*, 2013). However, the absolute level of Mad2 in a cell does not appear to be the primary factor controlling checkpoint function. Instead, the level of Mad2 relative to its binding partner and downstream target Cdc20 is the key, determining factor (Heinrich *et al.*, 2013). Our analysis of the tumor and cell line expression data revealed that Mad2 and Cdc20 displayed highly coordinated gene expression across the diverse samples (correlation of 0.43 in the CCLE data and 0.85 in the expO data).

Finally, we tested the correlation between the subunits of the Mis12 complex. Mis12 complex subunits (Mis12, Dsn1, Nnf1/Pmf1, and Nsl1) form an integrally associated, stable subcomplex (Kline *et al.*, 2006) but have not been a focus of previous studies on kinetochore gene up-regulation in cancer. Although Mis12 complex subunits show similar expression to other kinetochore genes in our analyses, the four subunits of this complex are not as strongly correlated as the Ndc80 complex (0.26 in the CCLE data set and 0.16 in the expO data sets). However, previous work on the Mis12 complex found that the protein levels of the Mis12 complex subunits are very sensitive to each other. For example, Dsn1 appears to be very unstable in the absence of the other subunits, both in cells (Kline *et al.*, 2006) and in biochemical reconstitution experiments (Cheeseman *et al.*, 2006). Although we focused here on the transcriptional control of kinetochore gene expression, translational and posttranslational control of protein levels will also affect the final subunit balances present in the cell.

Overall our analyses indicate that despite previous reports, Mad2 and Ndc80/Hec1 do not appear to be specifically up-regulated in cancer and tumor samples and instead are expressed as part of a larger cluster containing the Core Kinetochore and Cell Cycle/DNA Replication genes.

Expression of the FoxM1 transcription factor is highly correlated with expression of cell division genes

The coordinated expression of the Core Kinetochore and Cell Cycle/DNA Replication genes across the diverse data sets that we analyzed suggests that there is a transcriptional basis for this similar expression. To evaluate this, we analyzed the expression of genes that have been shown to control key gene expression programs, including the forkhead-type transcription factors and the proliferation promoting transcription factors E2F1 and E2F4.

E2F1 displayed moderately correlated gene expression to Core Kinetochore components in the CCLE and expO tumor data sets (correlations of 0.38 and 0.26, respectively). However, the E2F4 transcription factor did not show similar expression to kinetochore components (correlation –0.09). This is consistent with our observations that the expression behavior of the cell division genes is not strictly related to increased expression of periodically expressed genes, many of which are regulated downstream of E2F (Ren *et al.*, 2002).

The majority of forkhead-family transcription factors (including FoxA1, FoxC2, FoxO1, FoxO3, and FoxP3) did not display detectable similarity to the expression of the Core Kinetochore genes and Cell Cycle/DNA Replication components. In contrast, we found that FoxM1 showed highly correlated expression to kinetochore components in the expO tumor data (Figure 7; correlation 0.65). FoxM1 binds to promoter regions of the majority of kinetochore genes and cell division components based on recent chromatin immunoprecipitation followed by high-throughput DNA sequencing (ChIP-seq) analyses (Figure 5; ENCODE Project Consortium *et al.*, 2012 ; Chen *et al.*, 2013; Grant *et al.*, 2013; Sanders *et al.*, 2013). In particular, FoxM1 displays at least 10-fold enrichment of its ChIP-seq signal compared with a control IP in a region encompassing 1 kb upstream and 100 base pairs downstream of the transcriptional start site for 48 of 56 Core Kinetochore genes. This suggests that FoxM1 contributes to the coordinated expression of Core Kinetochore and Cell Cycle genes observed in normal tissues and cancer samples. Indeed, previous work found that FoxM1 is required for cell cycle progression and expression of multiple cell division genes (Laoukili *et al.*, 2005; Grant *et al.*, 2013). Coordinated expression of FoxM1-dependent genes suggests that regulation of FoxM1 itself could be a fundamental step in activating the cell division expression program in normal tissues and cancers.

DISCUSSION

Previous work documented numerous cases in which a given kinetochore gene is overexpressed in a selection of tumors. These observations led to proposals that such overexpression could be a causal event in promoting tumorigenesis (Yuen *et al.*, 2005). In fact, recent work shows that artificially up-regulating individual kinetochore genes can promote tumor formation in mice (Sotillo *et al.*, 2007; Diaz-Rodriguez *et al.*, 2008; Ricke *et al.*, 2011), but a broad-based consideration of whether such up-regulation of individual genes occurs in human tumors has been lacking.

To evaluate the significance of observed gene up-regulation in cancer, defining the basis for this up-regulation is of critical importance. There are several mechanisms by which a gene could be up-regulated. First, the expression of its corresponding upstream transcription factors could be changed. In this case, the altered expression of a transcription factor would be likely to affect the behavior of multiple downstream genes. Second, a promoter or enhancer region of an individual gene could be altered to increase the level of transcription. Third, the copy number of the gene could be increased through amplification or other genomic changes. However, because large genomic regions are typically amplified in cancers, it can be hard to assess which individual gene(s) within this amplified region are responsible for tumor behavior.

In cases in which a kinetochore gene is amplified or its promoter region is altered, this would be predicted to affect the mRNA level of that gene but not other kinetochore components. Such cases could represent causal events in tumor progression, as they could disrupt subunit balances within the macromolecular kinetochore structure and result in dominant-negative defects in kinetochore function, compromising the fidelity of chromosome segregation. However, it was unclear whether up-regulation of individual kinetochore genes occurs in human tumors. Here we analyzed kinetochore gene expression in diverse, publicly available expression compendia, which revealed that kinetochore genes are largely coordinately expressed across diverse conditions. On the basis of this strongly correlated behavior, we conclude that the up-regulation of an individual kinetochore gene is unlikely to be a causal factor in the vast majority of human cancers. However, we cannot exclude the possibility that minor changes in gene expression (less than twofold) do occur for some kinetochore components. Based on recent analysis in fission yeast (Heinrich *et al.*, 2013), such subtle imbalances for some kinetochore components could result in changes in mitotic progression or fidelity. We note that detecting such subtle changes in expression would be beyond the sensitivity of most experimental approaches, and experiments testing the consequences of artificially up-regulating kinetochore genes have induced >10-fold overexpression of a selected component.

Instead of specifically up-regulating specific kinetochore components, cancers appear to activate a broad-based cell division program that involves both the core kinetochore components and other genes required for cell cycle progression and DNA replication. We propose that this is not strictly related to mitotic index or doubling time of the cancer cells. Indeed, most kinetochore genes are not strongly periodically expressed during the cell cycle. Instead, this appears to be a general activation of the cell division program. Previous work found similar gene clusters for genes that function in related processes, such as ribosomal or proteasomal components (Niehrs and Pollet, 1999). On the basis of the observed coexpression behavior of the FoxM1 transcription factor and the binding of FoxM1 to the promoters of cell division genes, we propose that the up-regulation of the cell division program occurs downstream of FoxM1. Activation of this cell division program may allow a cell to be poised for cell division such that these proteins are ready to facilitate mitosis once additional cues signal cell cycle entry.

Our analyses provide an important context for considering the roles of kinetochore components in driving tumor progression and highlight the general activation of the cell division program in tumors. Because many cancer cells have been shown to display error-prone chromosome segregation (chromosomal instability; Thompson *et al.*, 2010), it is important to define the molecular basis for this behavior. Our analysis suggests that dramatic changes in the expression of an individual kinetochore gene are unlikely to be responsible for chromosome instability. In contrast, the recent sequencing of cancer cell (Barretina *et al.*, 2012) genomes revealed widespread mutations in components of the kinetochore and cell division apparatus that represent interesting candidates for future analyses.

TCGA tumor samples	TCGA normal samples
TCGA-A7-A13E-01A-11R-A12P-07	TCGA-A7-A13E-11A-61R-A12P-07
TCGA-A7-A13F-01A-11R-A12P-07	TCGA-A7-A13F-11A-42R-A12P-07
TCGA-BH-A0AU-01A-11R-A12P-07	TCGA-BH-A0AU-11A-11R-A12P-07
TCGA-BH-A0B5-01A-11R-A12P-07	TCGA-BH-A0B5-11A-23R-A12P-07
TCGA-BH-A0BS-01A-11R-A12P-07	TCGA-BH-A0BS-11A-11R-A12P-07
TCGA-BH-A0BZ-01A-31R-A12P-07	TCGA-BH-A0BZ-11A-61R-A12P-07
TCGA-BH-A0C3-01A-21R-A12P-07	TCGA-BH-A0C3-11A-23R-A12P-07
TCGA-BH-A0DD-01A-31R-A12P-07	TCGA-BH-A0DD-11A-23R-A12P-07
TCGA-BH-A0DT-01A-21R-A12D-07	TCGA-BH-A0DT-11A-12R-A12D-07
TCGA-BH-A0HA-01A-11R-A12P-07	TCGA-BH-A0HA-11A-31R-A12P-07
TCGA-BH-A18J-01A-11R-A12D-07	TCGA-BH-A18J-11A-31R-A12D-07
TCGA-BH-A18K-01A-11R-A12D-07	TCGA-BH-A18K-11A-13R-A12D-07
TCGA-BH-A18L-01A-32R-A12D-07	TCGA-BH-A18L-11A-42R-A12D-07
TCGA-BH-A18M-01A-11R-A12D-07	TCGA-BH-A18M-11A-33R-A12D-07
TCGA-BH-A18N-01A-11R-A12D-07	TCGA-BH-A18N-11A-43R-A12D-07
TCGA-BH-A18P-01A-11R-A12D-07	TCGA-BH-A18P-11A-43R-A12D-07
TCGA-BH-A18Q-01A-12R-A12D-07	TCGA-BH-A18Q-11A-34R-A12D-07
TCGA-BH-A18R-01A-11R-A12D-07	TCGA-BH-A18R-11A-42R-A12D-07
TCGA-BH-A18S-01A-11R-A12D-07	TCGA-BH-A18S-11A-43R-A12D-07
TCGA-BH-A18U-01A-21R-A12D-07	TCGA-BH-A18U-11A-23R-A12D-07
TCGA-BH-A18V-01A-11R-A12D-07	TCGA-BH-A18V-11A-52R-A12D-07
TCGA-BH-A1EO-01A-11R-A137-07	TCGA-BH-A1EO-11A-31R-A137-07
TCGA-BH-A1EU-01A-11R-A137-07	TCGA-BH-A1EU-11A-23R-A137-07
TCGA-E2-A153-01A-12R-A12D-07	TCGA-E2-A153-11A-31R-A12D-07
TCGA-E2-A158-01A-11R-A12D-07	TCGA-E2-A158-11A-22R-A12D-07
TCGA-E2-A15I-01A-21R-A137-07	TCGA-E2-A15I-11A-32R-A137-07
TCGA-E2-A15M-01A-11R-A12D-07	TCGA-E2-A15M-11A-22R-A12D-07
TCGA-E2-A1BC-01A-11R-A12P-07	TCGA-E2-A1BC-11A-32R-A12P-07

TABLE 1: TCGA samples.

MATERIALS AND METHODS

Data processing and analysis

For each expression compendium (HBI [GEO Accession No. GSE7307], expO [GEO Accession No. GSE2109], and CCLE [www.broadinstitute.org/ccle]), Affymetrix Human Genome U133 Plus 2.0 arrays were normalized with robust multiarray average (RMA) using the affy package from Bioconductor (www.bioconductor.org/packages/release/bioc/html/affy.html) and Entrez Gene (www.ncbi.nlm.nih.gov/gene) ID custom probeset definitions as defined previously (Dai *et al.*, 2005). The RMA values of selected genes were subsequently median centered and hierarchically clustered (based on uncentered correlation and average linkage) using Cluster 3.0 (de Hoon *et al.*, 2004) and visualized with Java TreeView (Saldanha, 2004). Principal component analysis plots gave similar results to the hierarchical clustering (unpublished data).

For the cell cycle study (Grant *et al.*, 2013), processed Agilent Whole Human Genome Oligonucleotide array ratios (sample/reference) were downloaded from Supplemental Table S1. The profiles of selected genes were sorted by arctan2 values, as provided in the table.

For TCGA breast invasive carcinoma data, processed Agilent expression data (level 3) for matching normal and tumor pairs (sample IDs are given in Table 1) were downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov). For each cancer sample, $\log_2$ ratios of selected genes were calculated relative to the matching normal sample, and the heatmap was created as described.

A subset of the expO tumor samples was also compared with normal counterparts in the HBI compendium. For each of the common HBI tissues (kidney, colon, liver, breast, lung, ovary/uterus, and prostate), the median expression level of each gene in that tissue was calculated, and each expO tumor sample was compared with the set of corresponding normal medians. The heatmap was generated as described.

Pearson correlations were calculated in R by comparing the expression ($\log_2$ ratios) of each gene against every other gene in the data set. The correlation values were then clustered in Cluster3 and visualized in Java TreeView, as described. Correlations were summarized by the mean via Fisher transformation.

FOXM1 ChIP-Seq peaks were downloaded from the ENCODE project (https://genome.ucsc.edu/ENCODE/downloads.html, under

Uniform TFBS). To associate peaks with genes, they were intersected with RefSeq's (www.ncbi.nlm.nih.gov/refseq/) gene promoters (defined as 1 kb upstream to 100 bases downstream of the transcription start site). When more than one peak overlapped a promoter, we chose the maximum fold enrichment. Model-based analysis of ChIP-Seq data fold-enrichments, as provided in the download, was used to make the bar graph.

ACKNOWLEDGMENTS

This work was supported by a Scholar Award to I.M.C. from the Leukemia and Lymphoma Society, grants to I.M.C. from the National Institutes of Health, National Institute of General Medical Sciences (GM088313), and the American Cancer Society (Research Scholar Grant 121776), and National Science Foundation Graduate Research Fellowships to J.M. and F.R.

REFERENCES

Barretina J *et al.* (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483, 603–607.

Bischoff JR *et al.* (1998). A homologue of *Drosophila* aurora kinase is oncogenic and amplified in human colorectal cancers. EMBO J 17, 3052–3065.

Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. Nature 490, 61–70.

Cheeseman IM, Chappie JS, Wilson-Kubalek EM, Desai A (2006). The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. Cell 127, 983–997.

Cheeseman IM, Desai A (2008). Molecular architecture of the kinetochore-microtubule interface. Nat Rev Mol Cell Biol 9, 33–46.

Chen X, Muller GA, Quaas M, Fischer M, Han N, Stutchbury B, Sharrocks AD, Engeland K (2013). The forkhead transcription factor FOXM1 controls cell cycle-dependent gene expression through an atypical chromatin binding mechanism. Mol Cell Biol 33, 227–236.

Chen Y, Riley DJ, Chen PL, Lee WH (1997). HEC, a novel nuclear protein rich in leucine heptad repeats specifically involved in mitosis. Mol Cell Biol 17, 6049–6056.

Dai M *et al.* (2005). Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res 33, e175.

de Hoon MJ, Imoto S, Nolan J, Miyano S (2004). Open source clustering software. Bioinformatics 20, 1453–1454.

Diaz-Rodriguez E, Sotillo R, Schvartzman JM, Benezra R (2008). Hec1 overexpression hyperactivates the mitotic checkpoint and induces tumor formation in vivo. Proc Natl Acad Sci USA 105, 16719–16724.

ENCODE Project Consortium, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.

Grabsch H, Takeno S, Parsons WJ, Pomjanski N, Boecking A, Gabbert HE, Mueller W (2003). Overexpression of the mitotic checkpoint genes BUB1, BUBR1, and BUB3 in gastric cancer—association with tumour cell proliferation. J Pathol 200, 16–22.

Grant GD, Brooks L3rd, Zhang X, Mahoney JM, Martyanov V, Wood TA, Sherlock G, Cheng C, Whitfield ML (2013). Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors. Mol Biol Cell 24, 3634–3650.

Hayama S, Daigo Y, Kato T, Ishikawa N, Yamabuki T, Miyamoto M, Ito T, Tsuchiya E, Kondo S, Nakamura Y (2006). Activation of CDCA1-KNTC2, members of centromere protein complex, involved in pulmonary carcinogenesis. Cancer Res 66, 10339–10348.

Heinrich S, Geissen EM, Kamenz J, Trautmann S, Widmer C, Drewe P, Knop M, Radde N, Hasenauer J, Hauf S (2013). Determinants of robustness in spindle assembly checkpoint signalling. Nat Cell Biol 15, 1328–1339.

Holland AJ, Cleveland DW (2009). Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nat Rev Mol Cell Biol 10, 478–487.

Kline SL, Cheeseman IM, Hori T, Fukagawa T, Desai A (2006). The human Mis12 complex is required for kinetochore assembly and proper chromosome segregation. J Cell Biol 173, 9–17.

Laoukili J, Kooistra MR, Bras A, Kauw J, Kerkhoven RM, Morrison A, Clevers H, Medema RH (2005). FoxM1 is required for execution of the mitotic programme and chromosome stability. Nat Cell Biol 7, 126–136.

Mondal G, Sengupta S, Panda CK, Gollin SM, Saunders WS, Roychoudhury S (2007). Overexpression of Cdc20 leads to impairment of the spindle assembly checkpoint and aneuploidization in oral cancer. Carcinogenesis 28, 81–92.

Niehrs C, Pollet N (1999). Synexpression groups in eukaryotes. Nature 402, 483–487.

Ren B, Cam H, Takahashi Y, Volkert T, Terragni J, Young RA, Dynlacht BD (2002). E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. Genes Dev 16, 245–256.

Ricke RM, Jeganathan KB, van Deursen JM (2011). Bub1 overexpression induces aneuploidy and tumor formation through Aurora B kinase hyperactivation. J Cell Biol 193, 1049–1064.

Ryan SD, Britigan EM, Zasadil LM, Witte K, Audhya A, Roopra A, Weaver BA (2012). Up-regulation of the mitotic checkpoint component Mad1 causes chromosomal instability and resistance to microtubule poisons. Proc Natl Acad Sci USA 109, E2205–E2214.

Saldanha AJ (2004). Java Treeview—extensible visualization of microarray data. Bioinformatics 20, 3246–3248.

Sanders DA, Ross-Innes CS, Beraldi D, Carroll JS, Balasubramanian S (2013). Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. Genome Biol 14, R6.

Shigeishi H, Higashikawa K, Ono S, Mizuta K, Ninomiya Y, Yoneda S, Taki M, Kamata N (2006). Increased expression of CENP-H gene in human oral squamous cell carcinomas harboring high-proliferative activity. Oncol Rep 16, 1071–1075.

Shigeishi H, Oue N, Kuniyasu H, Wakikawa A, Yokozaki H, Ishikawa T, Yasui W (2001). Expression of Bub1 gene correlates with tumor proliferating activity in human gastric carcinomas. Pathobiology 69, 24–29.

Sotillo R, Hernando E, Diaz-Rodriguez E, Teruya-Feldstein J, Cordon-Cardo C, Lowe SW, Benezra R (2007). Mad2 overexpression promotes aneuploidy and tumorigenesis in mice. Cancer Cell 11, 9–23.

Takimoto M *et al.* (2002). Frequent expression of new cancer/testis gene D40/AF15q14 in lung cancers of smokers. Br J Cancer 86, 1757–1762.

Thompson SL, Bakhoum SF, Compton DA (2010). Mechanisms of chromosomal instability. Curr Biol 20, R285–R295.

Tomonaga T, Matsushita K, Ishibashi M, Nezu M, Shimada H, Ochiai T, Yoda K, Nomura F (2005). Centromere protein H is up-regulated in primary human colorectal cancer and its overexpression induces aneuploidy. Cancer Res 65, 4683–4689.

Tomonaga T, Matsushita K, Yamaguchi S, Oohashi T, Shimada H, Ochiai T, Yoda K, Nomura F (2003). Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res 63, 3511–3516.

Vischioni B, Oudejans JJ, Vos W, Rodriguez JA, Giaccone G (2006). Frequent overexpression of aurora B kinase, a novel drug target, in non-small cell lung carcinoma patients. Mol Cancer Ther 5, 2905–2913.

Yuen KW, Montpetit B, Hieter P (2005). The kinetochore and cancer: what's the connection? Curr Opin Cell Biol 17, 576–582.