AD_______________
(Leave blank)

Award Number:
W81XWH-08-FË0638

TITLE:
Using Genetics and Genomics for the Detection and Treatment of Breast Cancer

PRINCIPAL INVESTIGATOR:
Stephen J. Elledge, Ph.D.

CONTRACTING ORGANIZATION:
Brigham & Women’s Hospital
Boston, MA 02115

REPORT DATE:ÁÁUæ*\æ↑âæãÁG€€Ï
Á

TYPE OF REPORT:
Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

X Approved for public release; distribution unlimited

Distribution limited to U.S. Government agencies only; report contains proprietary information

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
30 September 2009

2. REPORT TYPE
Annual

3. DATES COVERED *(From - To)*
01 September 2008 – 31 August 2009

4. TITLE AND SUBTITLE
Using Genetics and Genomics for the Detection and Treatment
Of Breast Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-08-1-0638

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Stephen J. Elledge, Ph.D.

Email: selledge@genetics.med.harvard.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brigham & Women's Hospital
Boston, MA 02115

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research
And Material Command
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

None provided.

15. SUBJECT TERMS

None provided.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	unclassified	20	19b. TELEPHONE NUMBER *(include area code)*
unclassified	unclassified	unclassified			

INTRODUCTION

The evolution of human cells into malignant derivatives is driven by the aberrant function of genes that positively and negatively regulate various aspects of the cancer phenotype, including altered responses to mitogenic and cytostatic signals, resistance to programmed cell death, immortalization, neoangiogenesis, and invasion and metastasis[1]. The integrity of these gene functions is compromised by substantial genetic and epigenetic alterations observed in most cancer cell genomes. To understand the tumorigenic process, it is imperative to identify and characterize the genes that provide tumor cells with the capabilities requisite for their initiation and progression. However, the identities of those genes that contribute to the tumor phenotype are often concealed by the frequent alterations in genes that play no role in tumorigenesis.

Identifying genes that restrain tumorigenesis (tumor suppressors) has proven especially challenging due to their recessive nature. Further complicating their discovery are the multifaceted mechanisms by which tumor suppressor genes are inactivated including changes in copy number and structure, point mutations, and epigenetic alterations[2]. Moreover, the mechanisms by which tumor suppressor genes are inhibited may vary between tumors. With this in mind, a variety of molecular and cytogenetic technologies have been used to establish extensive catalogs of genetic alterations within human cancers[3,4]**.** And while it is generally accepted that highly recurrent aberrations signify changes that are important for tumor development, the causal perturbations underlying tumor genesis are often confounded by the extensive size of alterations and the large number that are incidental to the tumor phenotypes. As such, new strategies to delineate genes with functional relevance to tumor initiation and development are essential to understanding these processes.

One approach to this problem involves the use of *in vitro* models of human cell transformation. In such models, primary cells are transformed into tumorigenic derivatives by the coexpression of cooperating oncogenes[5-7]. These experimental models have been useful in delineating the minimum genetic perturbations required for transformation of various human cell types as well as evaluating the functional cooperation between a gene of interest and a defined genetic context. To date, these models of human cell transformation have incorporated genes already implicated in human tumorigenesis. However, such models also provide a potentially useful platform for the identification of new pathways that contribute to the transformed phenotype.

In this award, we originally proposed two basic areas of investigation. The first area is the development of methods to investigate the repertoire of the immune system to determine whether auto-antibodies exist that might predict the onset of breast cancer. The second area was the construction and use of shRNA libraries to find genes relevant to breast cancer and hopefully targets that might kill tumor cells.

A key part of our research plan has been the development and use of retroviral vectors expressing RNA interference RNAs to identify human genes involved in causing or restraining cancer. In our first progress reports we described our efforts to develop shRNA libraries and showed they could be used to identify tumor suppressors. Ultimately our goal is to screen of complex pools of shRNA expressing retroviruses each marked with a bar code that allows the results of the screen to be read out by microarray hybridization. We demonstrated this could be accomplished in enrichment screens for shRNAs that caused cellular transformation and growth

in soft agar and identified the REST gene and several other tumor suppressors. However, a key goal has been to identify shRNAs that debilitate or kill cancer cells. In order for this to be possible in complex pools, it is imperative that each vector knock down its target with high penetrance. We have successfully achieved this level of knockdown and can now see particular shRNA expressing viruses drop out of complex pools. We have used this methodology to search for genes whose knockdown enhance the proliferative capacity of normal breast epithelial cells, i.e. candidate tumor suppressors and genes whose knockdown are cancer specific lethals.

We have also searched for genes that when overproduced are capable of transforming human mammary epithelial cells. In this way we have found several proteins with oncogenic potential, several of which are amplified in breast cancers. We have performed a biological analysis of one of these in detail, PVRL4.

With respect to our goal of deconvoluting the auto-antibody response in patients with breast cancer. We have developed a representation of all human linear peptides and methodology to screen them with breast cancer patient sera to look for autoantibodies that might provide a biomarker for early detection of breast cancer.

BODY

Identification of cancer-relevant genes.

Development and exploitation of shRNA libraries to identify cancer relevant genes using human genetic screening.

Retroviral shRNA-mediated genetic screens in mammalian cells are powerful tools for discovering loss-of-function phenotypes. We have been working on the generation of shRNA libraries for the express purpose of performing screens to kill cancer cell. During this grant we developed a highly parallel multiplex methodology for screening large pools of shRNAs using half-hairpin barcodes for microarray deconvolution.

In the first Innovator award we worked to produce the generation of barcoded, microRNA-based shRNA libraries targeting the entire human genome that can be expressed efficiently from retroviral or lentiviral vectors in a variety of cell types for stable gene knockdown[8-9]**.** These constructs include silencing triggers designed to mimic a natural microRNA primary transcript, and each target sequence was selected on the basis of thermodynamic criteria for optimal small RNA performance. Biochemical and phenotypic assays indicated that the new libraries are substantially improved over our first generation reagents. We generated a sequence verified library comprising more that 140,000 second generation short hairpin RNA expression plasmids covering a substantial fraction of all predicted genes in the human genome. This work is described in **Ref 9** and the details are in the published paper.

We also developed a method of screening complex pools of shRNAs using barcodes coupled with microarray deconvolution to take advantage of the highly parallel format, low cost, and flexibility in assay design of this approach[9-10]. Although barcodes are not essential for enrichment screens (positive selection)[10-12], they are critical for dropout screens (negative selection) such as those designed to identify cell lethal or drug sensitive shRNAs which we have

proposed to do for this Innovator award[13]. Hairpins that are depleted over time can be identified through the competitive hybridization of barcodes derived from the shRNA population before and after selection to a microarray. We initially described the use of 60-mer barcodes for pool deconvolution[9-10], and then developed a methodology called half-hairpin (HH) barcoding for deconvoluting pooled shRNAs[14].

We have made improvements over both the 60-mer barcodes and the half-hairpin barcodes. Both of those two barcodes suffer from cross hybridization and the occasional poor hybridization capacity for unknown reasons. We developed a framework for designing large sets of orthogonal barcode probes. We demonstrate the utility of this framework by designing 240,000 barcode probes and testing their performance by hybridization. From the test hybridizations, we also discovered new probe design rules that significantly reduce cross-hybridization after their introduction into the framework of the algorithm. These rules should improve the performance of DNA microarray probe designs for many applications. The details of this work are in the published paper **Xu et al, 2009 (Ref 17).** We have recently used some of these barcodes and they do seem to have much better properties than the half-hairpin barcodes.

Using these libraries and screening methods, we carried out dropout screens for shRNAs that affect cell proliferation and viability in cancer and normal cells. We identified many shRNAs to be anti-proliferative that target core cellular processes such as the cell cycle and protein translation in all cells examined. More importantly, we identified genes that are selectively required for proliferation and survival in different cell lines. Our platform enables rapid and cost-effective genome-wide screens to identify cancer proliferation and survival genes for target discovery. Such efforts are complementary to the Cancer Genome Atlas and provide an alternative functional view of cancer cells. This initial screening technology resulted in two papers in Science, **Schlabach et al, 2008 (Ref 18) and Silva et al, 2008 (Ref 19)**, and I will not go into great detail here with respect to the data because both manuscripts are published. The bottom line is that these methods work, they are robust and we can use them to find cancer relevant genes. I also describe below the identification of addition genes relevant to breast cancer cell survival using this technology

Control of REST Degradation

The transcription factor REST/NSRF (RE1-Silencing Transcription Factor) is a master repressor of neuronal gene expression and neuronal programs in non-neuronal lineages[23-25]**.** In the course of the first Innovator grant, we identified REST as a human tumor suppressor in breast epithelial tissues[20], suggesting that REST regulation may have important physiologic and pathologic consequences. We showed that REST knockdown using shRNA libraries could cause human mammary epithelial cells to acquire tumorigenic properties such as growth in soft agar. Many pathways controlling REST have yet to be elucidated. We went forward to further study this problem and found that REST is actually regulated by ubiquitin-mediated proteolysis[21]. We found through an RNAi screen that $SCF^{\beta TRCP}$ is an E3 ubiquitin ligase responsible for REST degradation. βTRCP binds and ubiquitinates REST and controls its stability through a conserved phosphodegron. During neural differentiation and in breast cells REST is degraded in a βTRCP-dependent manner. βTRCP is required for proper neural differentiation only in the presence of REST, indicating that βTRCP facilitates this process through degradation of REST. Conversely, failure to degrade REST attenuates differentiation.

Furthermore, we find that βTRCP overexpression, which is common in human epithelial cancers including breast, causes oncogenic transformation of human mammary epithelial cells and this pathogenic function requires REST degradation. Thus, REST is a key target in βTRCP-driven transformation and the βTRCP-REST axis is a new regulatory pathway controlling neurogenesis.

The data we generated demonstrate that REST is a labile protein targeted for ubiquitin-dependent proteasomal degradation by $SCF^{\beta TRCP}$ through a phospho-degron on REST. We showed that $SCF^{\beta TRCP}$ is a critical regulator of both physiologic and pathologic REST activities, constituting a new pathway controlling neural differentiation and cellular transformation. We provided the first genetic evidence that REST and $SCF^{\beta TRCP}$ regulate an early stage in neural specification as an inhibitor of neural differentiation. This is likely to be important in epithelial cancers as forced $SCF^{\beta TRCP}$ activation is known to cause breast cancer when expressed in mouse mammary glands. Our data are consistent with a model in which developmental cues induce degradation of REST, resulting in the derepression of REST targets. The ability of stable REST to inhibit terminal differentiation of neurons also predicts that REST may promote proliferative properties in the neuronal lineage when overproduced or inappropriately stabilized. Consistent with this notion, REST is overexpressed in human medulloblastoma and ectopic REST expression in *v-myc*-immortalized neural stem cells promotes medulloblastoma formation in mice[(26-27)]**.** Thus, the contrasting roles of REST as an oncogene and tumor suppressor are highly dependent on the developmental lineages.

βTRCP is overexpressed and oncogenic in epithelial cancers[(28-30)] and we identified REST as a key target in this context. This suggests that pharmacologic inhibition of βTRCP may provide a means to restore REST tumor suppressor function in human cancer. The presence of a phosphodegron motif within REST suggests a role for upstream kinase(s) and/or phosphatase(s) that control REST degradation. We propose a model in which differentiation into the neural state is induced by this yet to be discovered signal transduction cascade that targets REST for degradation by $SCF^{\beta TRCP}$, acting cooperatively with induction of βTRCP expression during neural differentiation.

Conversely, hyperactivation of such pathway(s) priming REST degradation may be oncogenic in epithelial tissues and thus serve as new therapeutic targets in cancers with compromised REST function. Thus, exploration of these pathways will likely provide new opportunities for modulating neural stem cell and cancer cell behavior. We need to find the kinase that regulates REST as it will be a potential oncogene. We have also found a new phospho-tyrosine phosphatase that is positively regulated by REST that when knocked down can cause cell transformation so we should be working that up in the new year.

We also participated in a collaboration with Yang Shi's lab who was working on methylatransferases. They found a protein CDYL that also binds REST and Methyltransferases to negatively regulate target gene expression through REST binding to promoters. We showed with them that CDYL is also a tumor suppressor candidate and knocking down CDYL expression transforms HMECs[(22)].

Identification of novel tumor suppressor genes

As we have described previously, we have developed viral shRNA libraries targeting the entire human genome to explore loss-of-function phenotypes in mammalian systems and have applied these libraries towards identification of novel tumor suppressor genes. With recent improvements in our library, we have revisited our search for novel tumor suppressor genes in breast cancer by using a genome-wide enrichment screen to identify genes whose knockdown increases proliferation and/or survival of normal human mammary epithelial cells (HMECs).

A detailed description of our HMEC enrichment screen is illustrated in Figure 1. Early passage HMECs used for the screen were obtained from a reduction mammoplasty and immortalized by expression of hTERT and spontaneous silencing of p16. To screen the entire genome-wide library of ~80,000 shRNAs, screens were performed in 6 separate pools of ~13,000 shRNAs. HMECs were infected in triplicate with a representation of 1000 cells per shRNA at an MOI of 2 viruses per cell. Initial reference samples were collected 72 hours post-infection. The remaining cells were puromycin-selected for 4 days and propagated with a representation of ≥1000 cells per shRNA maintained at each passage. Cells were collected as the end samples after ~7 population doublings (PDs). Probes were prepared from both samples, and Cy3- or Cy5-labeled probes were competitively hybridized to half-hairpin barcode microarrays to measure the change in representation of each shRNA over time.

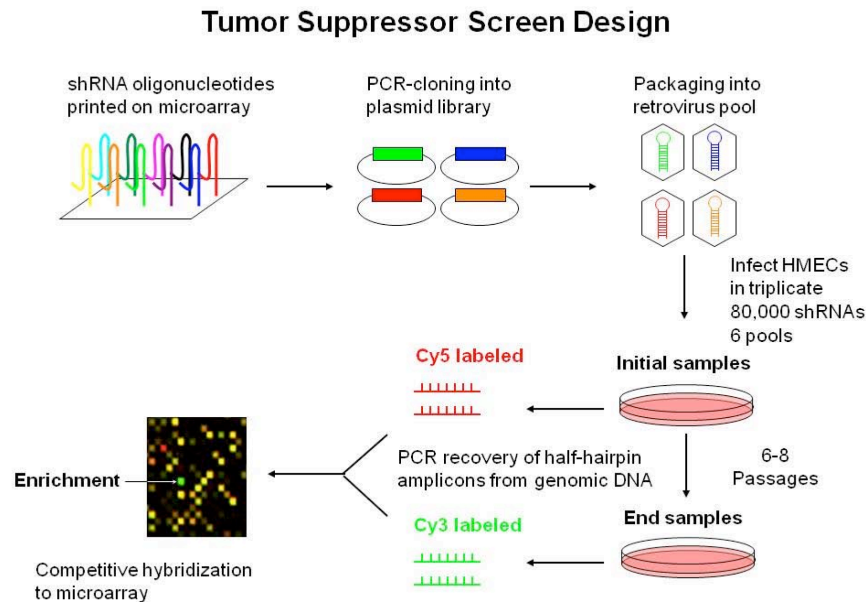

Figure 1. Tumor suppressor screen design. A genome-wide enrichment screen was performed in human mammary epithelial cells for shRNAs against genes whose knockdown increases proliferation or cell survival. Such genes are novel tumor suppressor candidates

Statistical analysis indicate that most probes consistently yielded signals >2-fold above the mean background of negative control probes across all triplicates. The correlations among the initial samples across triplicates and between the initial and end samples within each replica were high, indicating high reproducibility and maintenance of representation. As expected, most shRNAs showed little change over the time course of the screen. However, using statistical analysis for microarray (SAM) with a false discovery rate (FDR) of 5%, we observed

that the abundance of 4257 shRNAs against 3653 genes were increased > 2-fold over the course of the screen, indicating these shRNAs increase proliferation and/or survival of HMECs. Many of these genes show statistically significant overlap with deletion regions in breast cancers.

Many of the genes whose knockdown lead to increased proliferation and/or survival are known tumor suppressor genes or key mediators of proliferation and apoptotic pathways. For example, shRNAs against genes encoding the tumor suppressors Rb, PTEN, and p53 were all increased over the course of the screen. Furthermore, shRNAs against proliferation genes, such as those encoding the Rb-like proteins p107 and p130 and cyclin dependent kinase inhibitors p21, p27, and p57, as well as shRNAs against prop-apoptotic genes, such as those encoding caspase 3, caspase 6, and Apaf-1 scored as well.

We are currently using a multi-color competition assay (MCA) to validate shRNAs which have scored in our screen and determine whether knockdown of candidate genes increases proliferation and/or survival. Thus far, we have examined 142 candidate shRNAs against 123 known genes that scored in our screen for their effect on cell proliferation and/or survival. Using the MCA assay, 58 of 142 candidate shRNAs (40%) validated to increase proliferation and/or survival compared to FF2 control shRNAs over the 6 day assay. We are currently investigating these genes to determine the mechanisms by which their knockdown increases proliferation and/or survival and whether they are novel tumor suppressors. Furthermore, we are using genomic profiling of tumor samples to determine whether these genes are located in focal deletion regions, which would suggest that they are involved in tumor suppression.

We have also synthesized a new library of shRNAs that has a much deeper representation of shRNAs for our candidate genes, 12 hairpins per gene and have screened that library to look for sequences enriched after growth in HMEC cells. This will give us the ability to say whether or not the gene of interest is indeed the target of the shRNAs we have identified. If we find multiple shRNAs to a single gene, that will mean it is a real target as opposed to an off target.

Identification of breast cancer-specific lethal genes

Genetic loss-of-function screening to identify genes that are essential to cancer cell proliferation and survival is a powerful and complementary approach to large sequencing efforts and is expected to provide many potential cancer drug targets. Towards this end, we have performed lethality screens to identify genes that are selectively required for proliferation and survival of breast cancer cells but not normal mammary epithelial cells, which we call Breast Cancer Lethal (BCAL) genes. For lethality screens, the abundance of shRNAs targeting genes that are essential for cell viability will be reduced following cell passaging and will thus "drop-out" of the shRNA population. By comparing each shRNA's abundance in an initial cell population taken shortly after retroviral shRNA library infection to its abundance in samples taken after several cell population doublings, lethal shRNAs can be identified. Additionally, comparisons between the shRNA lethality profiles of breast cancer cells and normal human mammary epithelial cells can identify BCAL genes (Figure 2).

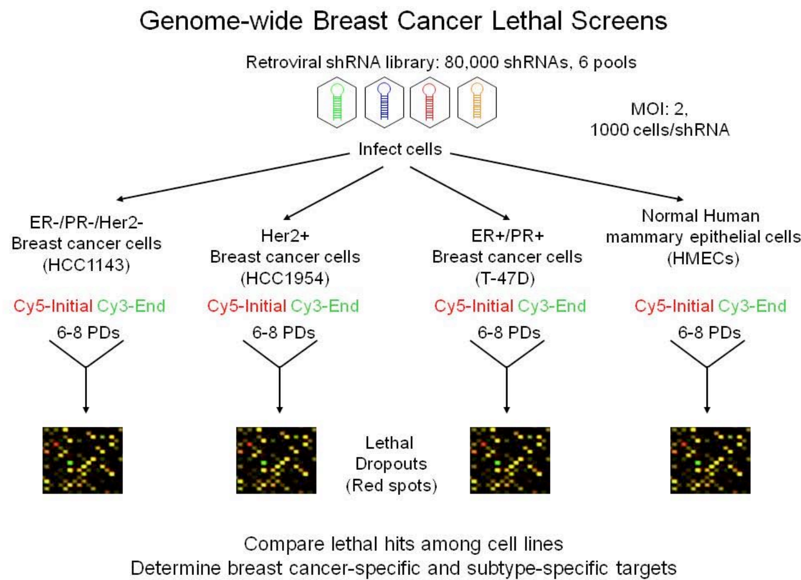

Figure 2.
Breast cancer lethal screen design. A genome-wide lethality screen was performed in three breast cancer cell lines and one normal human mammary epithelial cell line for shRNAs against genes whose knockdown decreases viability of breast cancer but not normal cells. Such genes are novel cancer drug targets.

To identify BCAL genes, we have performed highly-parallel, genome-wide, pooled, shRNA lethality screens in three breast cancer (HCC1954, HCC1143, T47D) and one normal mammary epithelial (HMEC) cell lines. Cells were infected in triplicate with the entire genome-wide library of ~80,000 shRNAs in 6 separate pools of ~13,000 shRNAs/pool. Cells were infected in triplicate with a representation of 1000 cells per shRNA at an MOI of 2 viruses per cell. Initial reference samples were collected 72 hours post-infection. The remaining cells were puromycin-selected for 4 days and propagated with a representation of ≥1000 cells per shRNA maintained at each passage. Cells were collected as the end samples after ~7 population doublings (PDs). Probes were prepared from both initial and end samples, and Cy3- or Cy5-labeled probes were competitively hybridized to half-hairpin barcode microarrays to measure the change in representation of each shRNA over time.

Statistical analysis indicate that most probes consistently yielded signals >2-fold above the mean background of negative control probes across all triplicates. The correlations among the initial samples across triplicates and between the initial and end samples within each replica were high, indicating high reproducibility and maintenance of representation. To identify cancer-specific lethal shRNAs, we utilized statistical analysis for microarray (SAM) with a false discovery rate (FDR) of 10% as well as several fold change cutoffs. For a given shRNA to score as a BCAL shRNA, its abundance decreased > 2-fold in one of the three breast cancer cell lines, but was not decreased > 2-fold in HMECs. Furthermore, a BCAL shRNA displayed > 1.8 fold difference in abundance between the normal and breast cancer cell line. We identified 3787 shRNAs against 3410 genes that met these BCAL criteria. These shRNAs and genes were further classified into 4 groups: "pan" BCAL shRNAs that were selectively lethal to 2 or more breast cancer cell lines, and HCC1954, HCC1143, and T47D cell type-specific BCAL shRNAs that were selectively lethal to a single breast cancer cell line. We are currently validating whether knockdown of candidate BCAL genes leads to reduced viability of breast cancer cell lines but not normal HMECs using a Cell Titer Glo viability assay. Furthermore, we are

investigating these candidates using mechanistic studies to determine how their reduction leads to cancer cell lethality. We expect these BCAL genes will reveal novel oncogene or non-oncogene addictions of breast cancer cells that can lead to new therapeutic targets for breast cancer.

Identification of synthetic lethal genes with K-ras

A major challenge in cancer therapeutics is the identification of cellular drug targets whose inhibition leads to the selective killing of cancer cells while sparing normal cells. Recent advances in mammalian RNA interference (RNAi) technologies have made it possible to systematically interrogate the human genome for genes whose loss of function constitute synthetic lethality either with the oncogenic state or with particular oncogenic mutations[13, 18, 19]. We have developed barcoded, retroviral/lentiviral-based short hairpin RNA (shRNA) libraries targeting the entire human genome to enable genome-wide loss-of-function analysis through stable gene knockdown[9]. Our design also allowed us to develop a multiplex screening platform that enables the highly parallel screening of >10,000 shRNAs in a pool-based format using microarray deconvolution[18,19]. These technological breakthroughs have made it possible to rapidly interrogate the genome for functional vulnerability of cancer cells and here we apply these to the Ras oncogene.

The Ras family of small GTPases are frequently mutated in human cancers [Reviewed in **Ref 34]**. Ras is a membrane-bound signaling molecule that cycles between the inactive, GDP-bound state and the active, GTP-bound state. Growth factor receptor signaling promotes GTP loading and activation of Ras, which in turn activates an array of downstream pathways to promote cell proliferation and survival. Among the major Ras effector pathways is the MAP kinase pathway, the PI3-kinase (PI3K) pathway, RalGDS proteins, phospholipase-Cε and Rac. Each of these has been implicated in mediation of Ras oncogenesis. Ras GAPs (GTPase activating proteins) inactivate Ras by stimulating GTP hydrolysis. Oncogenic mutations in Ras are invariably point mutations that either interfere with Ras GAP binding or directly disrupt Ras GTPase activity, locking Ras in a constitutively active, GTP-bound state. Oncogenic mutations have been found in all three members of the Ras gene family with *KRAS* being the most frequently mutated. *KRAS* mutations are found at high frequencies in pancreatic, thyroid, colon, lung and liver cancers and in myelodyspastic syndrome and are correlated with poor prognosis[34]**.**

Despite its prominent status as a cancer drug target, therapeutics aimed at disrupting the Ras pathway have proven challenging thus far. Inhibitors of farnesyl transferase, the enzyme that prenylates Ras for its membrane localization, have met with only limited success[34]**.** Chemical screens in isogenic Ras mutant and wild type cell lines have identified compounds that exhibit preferential toxicity towards Ras mutant cells[35-36]**.** However, the translation of these chemical screens into clinical practice has been impeded by the challenge in identifying the protein targets of these chemical entities and subsequent drug development. Inhibitors targeting various Ras effecter pathways could also prove efficacious in treating tumors with Ras mutations, as it was recently shown that a combined application of MEK and PI3K/mTOR inhibitors can reduce tumor burden in a mouse model of Ras-driven lung cancer[37]**.** However, the prevalence of *de novo* and acquired drug resistance to other targeted therapies suggests that combinations of multiple therapeutic agents will be required to effectively inhibit malignant progression.

In principle, tumors can be attacked by either reversing the effects of oncoproteins through inhibition (i.e. exploiting oncogene addiction), or by attacking tumor-specific vulnerabilities caused by the oncogenic state, often by inhibiting proteins that are not oncoproteins themselves (i.e. exploiting non-oncogene addiction)[32, 33]. The inappropriate rewiring of cellular signaling through oncogene activation should result in vulnerabilities that could be exploited for cancer therapies in theory. Since these vulnerabilities are not obvious and cannot be predicted, the most direct approach to their discovery is through genetic exploration. The systematic identification of genes and pathways necessary for the Ras-driven oncogenic state would provide additional drug targets for therapeutic exploration, shed new light on Ras' mechanisms of action and potentially provide new biomarkers for patient stratification. To this end, we undertook a genome-wide RNAi screen to identify synthetic lethal interactions with the KRAS oncogene. We discovered a diverse set of proteins whose depletion selectively impaired the viability of Ras mutant cells. Among these we observed a strong enrichment for genes with mitotic functions. We found a pathway involving the mitotic kinase PLK1, the anaphase promoting complex/cyclosome and the proteasome that, when inhibited, results in prometaphase accumulation and the subsequent death of Ras mutant cells. Gene expression analysis indicates that reduced expression of genes in this pathway correlates with increased survival of patients bearing tumors with a Ras transcriptional signature. Our results suggest a previously underappreciated role for Ras in mitotic progression and demonstrate a pharmacologically tractable pathway for the potential treatment of cancers harboring Ras mutations.

We have found that the synthetic lethal approach to be very informative with respect to identifying potential targets for anti-cancer therapeutics that can be used in combination with other drugs to attack a cancer of a particular genotype with particular oncogenic or tumor suppressor mutations. We hope to continue these sorts of screens in the future to identify genes whose depletions are toxic in different cancers. The experimental details of this work are included in the Luo et al Cell paper [38].

Anti-tumor Antibody profiling

Anti-tumor auto-antibodies have been proposed to be highly sensitive and specific biomarkers for early cancer detection. To identify antibody binding profiles specific for cancer patients, a number of groups have utilized phage display, a high- throughput affinity assay. In this approach, libraries of peptides or protein fragments are displayed on the surface of a bacteriophage, thereby allowing a proteomic screen for binding properties of a patient's antibody repertoire. Previous efforts towards auto-antibody biomarker discovery using phage display have used either random peptides or tumor-derived cDNA libraries to "pan" displayed peptides against patient sera. Such studies have been difficult to interpret, however, since resulting "hits" are frequently out of frame or derived from noncoding sequences.

Our novel strategy has been to encode the human "peptidome" as a set of individual DNA microarray-derived oligonucleotides. We have generated a library of approximately 467,000 oligos that tile the entire set of human open reading frames. For phage display, the oligos were cloned into the T7Select (Novagen) system, allowing for low copy number display fused to the

T7 coat protein, 10B. The library was then extensively characterized by sequencing several hundred individual phage clones at random. As we reported last year, 74% of our phage population encodes in-frame peptides, and 55% of the population expresses completely correct sequences. This phage library is unique in several respects. First, this is the first example of a synthetic phage display library encoding protein fragments from microarray-derived oligonucleotides. Second, this library is the only example of a normalized representation of the human peptidome. The alternative random peptide and cDNA derived libraries are much less powerful from a cancer biomarker screening standpoint.

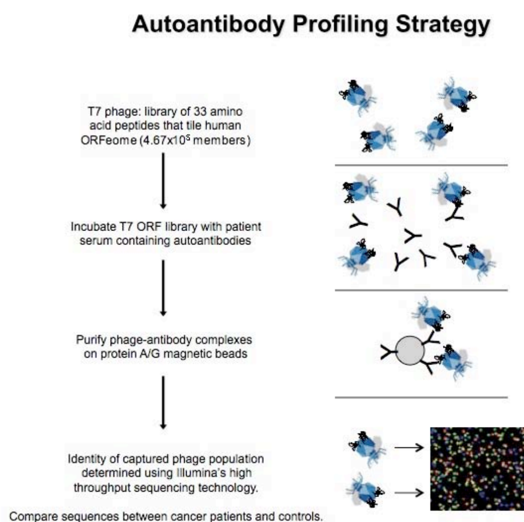

Figure 3.
Novel approach to patient autoantibody profiling. T7-Peptidome library is mixed with patient antibodies and specific complexes are allowed to form. Complexes are purified on magnetic protein A/G beads. High throughput sequencing is utilized to identify captured phage, and profiles are compared between patients.

The sensitivity of our screen for auto-antibodies associated with breast cancer can be determined by the degree of specific enrichment that we are able to obtain by immunoprecipitating antibody-bound phage. To this end, we have optimized the conditions for enrichment since our last report. By diluting a FLAG-tagged T7 phage (1:1000) into a native phage population and diluting an anti-FLAG antibody (1:1000) into a non-specific isotype control antibody, we were able to mimic the high complexity of our assay with just four variables. We optimized for enrichment (measured quantitatively with the plaque lift assay) by systematically varying the following parameters: T7 phage concentration, antibody concentration, time of immunoprecipitation, and number of washes before phage elution. The concentrations of phage and antibody during complex formation were assumed to be dependent on each other, and were thus optimized simultaneously (Figure 4). We successfully optimized all 4 parameters, allowing for a FLAG phage enrichment factor of over 4000-fold showing the basic premise of the screen is sound and we should be able to screen breast serum samples.

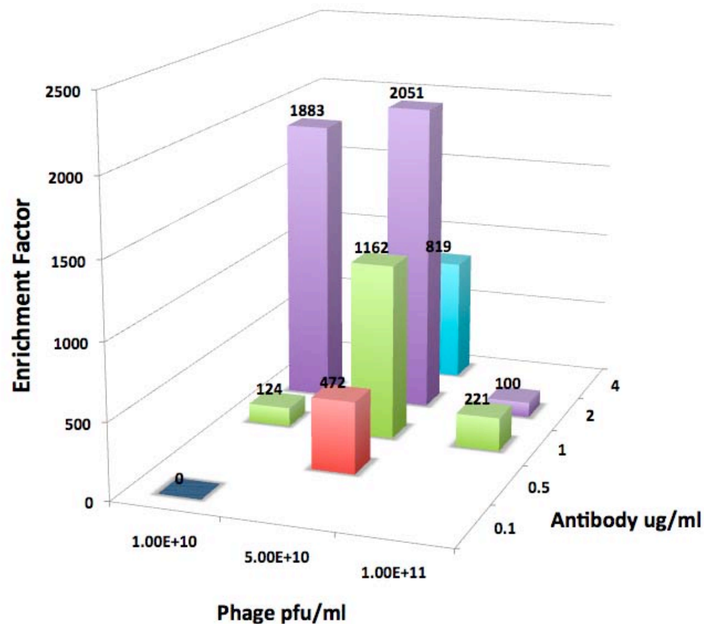

Figure 4.
Optimization of enrichment with respect to phage and serum antibody concentration. Total phage and serum antibody were mixed at concentrations indicated and let to form complexes overnight at 4°C. Enrichment was determined by plaque lift assay for FLAG expressing phage.

After optimizing enrichment conditions, we performed our optimized immuno-precipitation assay on our synthetic phage library to validate our enrichment protocol. To this end, we immunoprecipitated a pool of our library with a combination of 9 antibodies, all of which are directed against the C-termini of different human proteins. This pool consisted of 27,000 unique phage displaying the C-terminal peptidome. Finally, a plaque lift of the immunoprecipitated phage were immunoblotted with the same mix of C-terminal antibodies (Figure 5). Whereas there were apparently no "hits" on a plaque lift derived from the input library, we observed a large fraction of positively staining plaques on the immunoprecipitated sample. Of the thirty hits that were sequenced, all of these corresponded to peptides specifically targeted by the C-terminally-directed input antibodies, suggesting that our optimized enrichment protocol is highly selective for specific interactions.

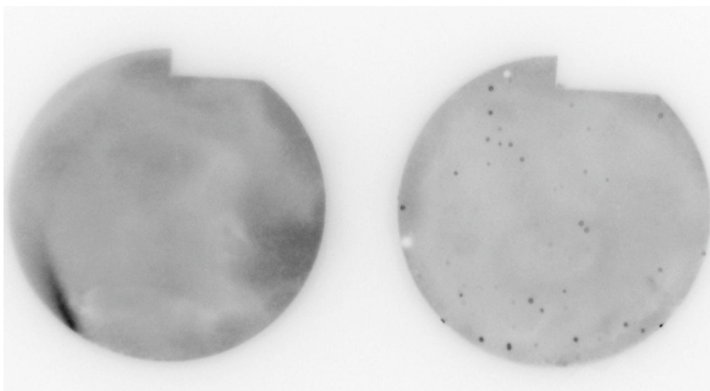

Figure 5.
Plaque lifts of the C-terminal peptidome libraries, stained with a pool of 9 anti-C-terminal protein antibodies. Blot on the left is from the input library, and blot on the right is from the phage immunoprecipitated with the same pool of antibodies.

As reported last year, we had experienced difficulty using microarrays to uniquely identify the phage species present in a given sample. This was largely due to the fact that our probe design was highly constrained by the sequences of the oligonucleotide inserts. A high degree of cross-hybridization was observed, which compromised the quality of our data. The alternative, high throughput sequencing of the oligo inserts, is expensive and therefore prohibitive for a screen of >100 samples. We have therefore developed an indexing scheme to allow multiplexing of samples during high throughput sequencing, thereby reducing the cost per sample (Figure 6).

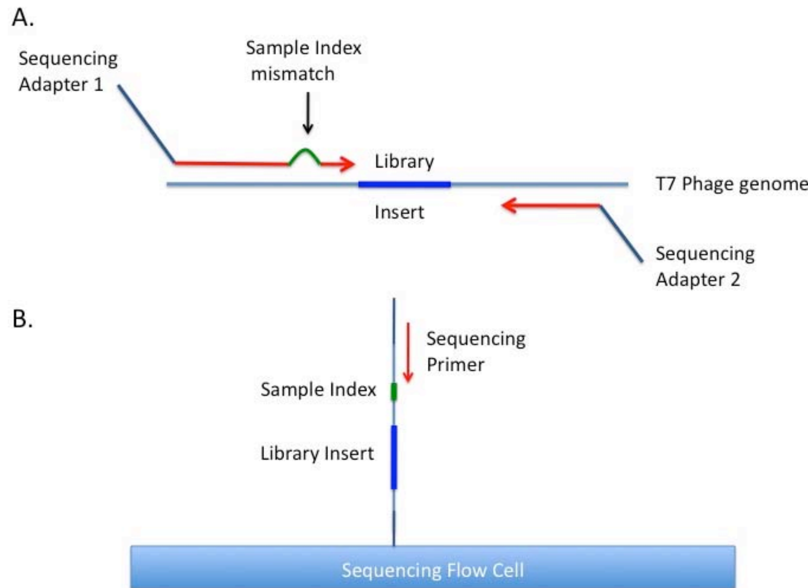

Figure 6.
Indexing and sequencing strategy for the library inserts. *A.* Two primers are used to amplify the library insert from the phage genome. Each primer includes the appropriate adapter for bridge PCR. Additionally, one primer contains a mismatch at one of three bases, which will be used to uniquely identify the sample within a multiplexed pool.
***B.* Sequencing of the index and library insert in the Solexa flow cell.**

We performed Solexa sequencing on the T7 C-terminal peptidome library. After alignment of the 8.4 million reads which passed QC, we noted that about 94% of all 22,454 expected sequences were indeed observed. Next, after performing immunoprecipitation using healthy control serum, 90% of the input phage population remained, suggesting that the IP did not significantly bottleneck the population. As an IP control, we spiked in three of the C-terminal directed antibodies (2 ng/ml) that had been validated previously. The three target peptides corresponding to these antibodies were among the most strongly enriched sequences in the dataset.

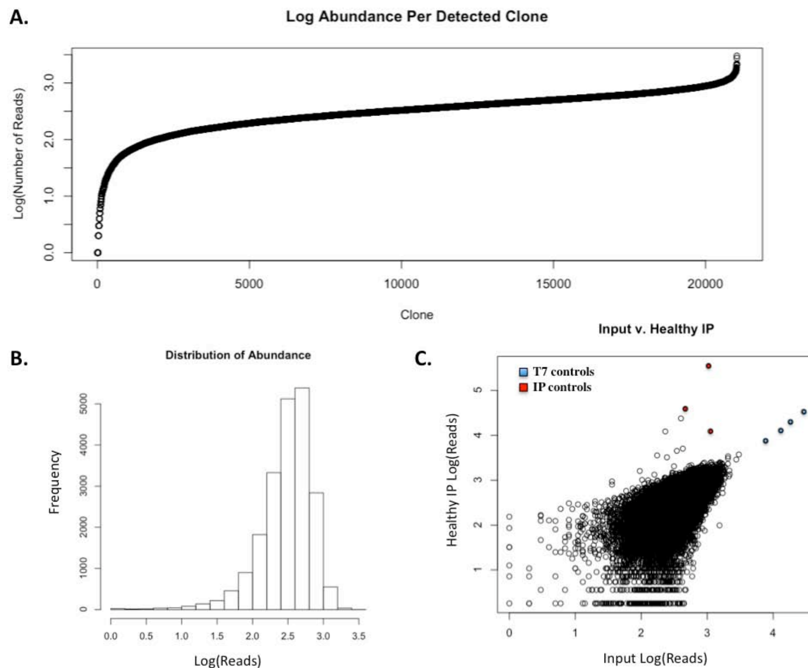

Figure 7. Sequencing of the C-terminal peptidome phage population. *A-B.* Distribution of the sequence abundance for each clone in the C-terminal library. *C.* Scatter plot of sequence abundance for each clone, comparing input population with population after IP from healthy serum. High abundance T7 controls are shown in blue. Targets of the three spiked in antibodies (ATR, nibrin, SAPK4) are shown in red.

In order to assess the accuracy of the sequence-based enrichment measurement, we performed sequencing and a plaque lift on the same immuno-precipitated material. By plaque lift assay we observed that about 1 of every 35 phage in the population was SAPK4. By sequencing the same sample, 1 of every 25 sequences aligned with the SAPK4 clone (Figure 7).

The C-terminal peptidome IP was then performed using 3 pathological sera. Sera from one breast cancer (BC) patient, one confirmed paraneoplastic disease (PND) patient, and one multiple sclerosis (MS) patient were used in the screen. SAPK4 antibody had been spiked in at 1:1000 (relative to patient antibody) and this clone was strongly enriched in all of the samples. We did not see any obvious correlation between enriched clones and the particular pathology, but this is not surprising since the screen utilized only C-terminal peptides, and not the full proteome, which is the next step. This dataset was used to model the effect of screening against the more complex (~19 fold) full peptidome library, as well as the effect of multiplexing samples (by 8 fold) during a large-scale screen.

Key research accomplishments

1. Generation of genome wide phage display libraries that cover the entire coding capacity of the human genome for the auto-antibody profiling. The determination that IPs can be performed on these phage display libraries and enrich target phage.
2. Development of a deconvolution strategy that allows multiplex sequencing of enriched phage for deconvolution through next generation sequencing methods.
3. Identification of BetaTRCP as the ubiquitin ligase that controls REST stability.
4. Identification of CDYL as a bridge between REST and methyl transferases to repress gene expression and determination that CDYL suppresses cellular transformation in breast cells.
5. Identification of several genes that when overexpressed cause transformation of mammary cells.
6. Identification of several genes that appear to be synthetically lethal with ras.
7. Pharmacological validation of Plk1 as a synthetic lethal with Kras.
8. Demonstration that we can perform genome-wide shRNA screens on breast lines.
9. Development of a library of 250,000 new orthogonal barcodes of 25mers that do not cross hybridize to faithfully report on relative abundance of shRNA vectors.
10. Identification of genes that are selectively toxic to multiple breast cancer cell lines.

Statement of Work

Year 1.

Task 1 (Months 1-12)
In the first year we anticipate completing the first whole genome lethality screens for 2 breast cancer lines and one normal mammary epithelial line. We will also characterize the genes found from those screens.

We have completed the screening of two cancer cell lines and one normal human mammary epithelial cell line. We have been focusing on retesting the genes we have found to show that they are reproducibly killing cancer cells or enhancing the growth of normal mammary epithelial cells.

Task 2 (Months 1-24 and possibly longer, an ongoing effort)
We will further develop the auto-antibody profiling method readout and screen the 200 normal and 200 breast cancer patients sera we have already obtained.

We have made the libraries and obtained the sera samples. In this last year we have established that the methods should work with reconstruction experiments and are working out ways to screen the data generated. We cannot use microarray readouts because of cross hybridization. We have solved this by sequencing using Next Generation sequencing using the Solexa platform. We have developed the system to the point where we have confidence that an antibody that recognizes a peptide will immuoprecipitate it out of a complex library covering all human peptides and enrich it by over 1000-fold. We are in the process of performing the initial IPs to characterize the immunome's interaction with the human peptidome.

Task 3 (Months 6-24)
We will be completing the ras synthetic lethality screen we have begun. This genome-wide analysis should be finished and verified hits worked up.

We have accomplished this aim. We carried out a genomic screen for shRNAs that would specifically kill cells with activated ras. In this screen we identified a large number of genes involved in mitosis revealing a previously unappreciated mitotic defect in ras mutant tumors. We also found that the expression levels of genes required to support ras tumorigenesis correlated with patient survival.

Task 5 (Months 1-12) We will continue to characterize PVRL4/Nectin-4 for its role in transformation of breast cancer cells and try to determine the identity of its downstream effectors and if and how they contribute to transformation.

We have performed a structure function analysis on PVRL4 and found that only its extra-cellular domain is required for transformation. We also found that breast cancer cells that over-express PVRL4 completely change their morphology when PVRL4 is knocked down by shRNA expression.

REPORTABLE OUTCOMES

THE FOLLOWING PAPERS WERE PUBLISHED BASED ON FINDINGS THAT EMANATED FROM THIS GRANT IN 2008 AND 2009. MOST OF THESE WERE INITIATED IN THE ORIGINAL INNOVATOR GRANT FUNDING PERIOD.

Schlabach, M.R., Luo, J., Solimini, N.L., Hu, G., Xu, Q., Li, M.Z., Zhao, Z., Smogorzewska, A., Sowa, M.E., Ang, X.L., Westbrook, T.F., Liang, A., Chang, K., Hackett, J.A., Harper, W.J., Hannon, G.J., and Elledge, S.J. (2008) Cancer Proliferation Gene Discovery Through Functional Genomics. ***Science*** 319:620-4.

Silva, J.M., Marran, K., Parker, J.S., Silva, J., Golding. M., Schlabach, M.R., Elledge, S.J., Hannon, G.J., Chang, K. (2008) Profiling essential genes in human mammary cells by multiplex RNAi screening. ***Science*** 319:617-20.

Westbrook, T.F., Hu, G., Ang, X.L., Mulligan, P., Pavlova, N.N., Liang, A., Leng, Y., Maehr, R., Shi, Y., Harper, W.J., and Elledge, S.J. (2008) SCF-bTRCP Controls Oncogenic Transformation and Neural Differentiation Through REST Degradation. ***Nature*** 452:370-4.

Mulligan, P., Westbrook, T.F., Ottinger, M., Pavlova, N., Chang, B., Macia, E., Shi, Y.J., Barretina, J., Liu, J., Howley, P.M., Elledge, S.J., Shi, Y. (2008) CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. ***Mol Cell*** 32:718-26.

Luo J, Solimini NL, Elledge SJ. (2009) Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction. ***Cell***. 136:823-37.

Luo, J., Emanuele, M.J., Li, D., Creighton, C.J., Schlabach, M.R., Westbrook, T.F., Wong, K.-K., Elledge, S.J. (2009) A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. ***Cell*** 137:835-48.

Xu, Q., Schlabach, M.R., Hannon, G.J., Elledge, S.J. (2009) Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci. U. S. A.* 106:2289-94.

CONCLUSIONS

Progress on understanding the REST Tumor suppressor pathway.

It is clear that REST acts as a tumor suppressor in mammary cells. That means that the pathways controlling REST are also likely to be important in tumor suppression as well. We have discovered a new protein kinase driven pathway that targets REST for proteolytic degradation via the SCF using the F-box protein BTRCP. Since BTRCP overproduction is oncogeneic and causes cellular transformation through REST degradation, we have now established a cellular transformation pathway that links a known oncogene to the negative control of a tumor suppressor gene. We have also found a new phospho-tyrosine phosphatase that is positively regulated by REST that when knocked down can cause cell transformation so we should be working that up in the new year.

Progress on barcode screening for essential genes

It is clear from our current studies that we have overcome the main problem with performing bar code screens which is getting sufficiently good knockdown from single copy vectors and being able to reproducibly measure their abundance in complex pools by microarray hybridization. We have now gone most way through a genome screen for cancer-specific lethals in three genetically distinct breast cancer lines. This has been our major goal from the very start and we are now verifying the results and determining which are truly cancer specific. In addition, we have completed a screen to find genes that are synthetically lethal with Kras mutations. We are beginning cMyc and PI3K synthetic lethals in the next year. This should tell us which mutations in the breast lines are causing the synthetic lethality we are seeing.

Screens for Tumor suppressors using the RNAi library and for Oncogenes using the ORFeome library

We originally thought last year that this ongoing effort should be completed this year. However, we ran into trouble with our cell transformation assay. Apparently the supplier of our specialized media for HMECs switched some of their components and nothing worked. We worked hard for 6 months and thought we had finally overcome that problem which is important for both the shRNA screens as well as the overproduction screens for oncogenes but we were wrong and still have a problem. We think we may be back on track but it will be a while. But we have now constructed new shRNA libraries that have 12 hairpins per gene which we have begun to screen with using TLM-HMECs and we have been reading these screens out using Solexa sequencing. The interpretation of this data is complex and we are still working through it.

In addition we are performing the same screens with retroviral ORFeome libraries which are the equivalent to normalized full length cDNA libraries. One gene we are following up is PVRL4/Nectin-4. It potently transforms HMECs and is overproduced in 62% of ductal carcinomas. We have found several breast lines that express PVRL4 and when it is knocked down the cells show reduced proliferation and no longer grow in clumps like stem cells. We hope to have a paper on this next year.

Investigation of Auto-Antibodies as Breast Cancer Biomarkers

We now have the libraries in hand and the patient samples. We ran into read-out problems on the microarrays because we are forced to use the sequences that are coded as opposed to optimized barcodes. We have solved this problem by converting to DNA sequencing which, unfortunately is much more expensive but we have shown that it can work in reconstruction experiments. We have begun screening the first patient samples and we hope that this will be done in the next year.

REFERENCES

1. D. Hanahan and R.A. Weinberg, Cell 100, 57-70 (2000)
2. A. Balmain, J. Gray, J., and B. Ponder. Nat Genet *33 Suppl*, 238-244 (2003).
3. D. G. Albertson, et al. Nat Genet *34*, 369-376 (2003).
4. P.A. Futreal et al. Nat Rev Cancer *4*, 177-183 (2004).
5. B. Elenbaas et al. Genes Dev *15*, 50-65 (2001).
6. W.C. Hahn et al, Nature *400*, 464-468 (1999).
7. J.J. Zhao et al, Trends Mol Med *10*, 344-350 (2004).
8. P. J. Paddison *et al.*, *Nature* 428, 427 (2004).
9. J. M. Silva *et al.*, *Nat. Genet.* 37, 1281 (2005).
10. T. F. Westbrook *et al.*, *Cell* 121, 837 (2005).
11. N. Popov *et al.*, *Nat. Cell Biol.* 9, 765 (2007).
12. I. G. Kolfschoten *et al.*, *Cell* 121, 849 (2005).
13. V. N. Ngo *et al.*, *Nature* 441, 106 (2006).
14. L. D. Wood *et al.*, *Science* 318, 1108 (2007).
15. K. Berns *et al.*, *Nature* 428, 431 (2004).
16. T. R. Brummelkamp *et al.*, *Nat. Chem. Biol.* 2, 202 (2006).
17. Q. Xu, et al. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2289-94 (2009).
18. M.R. Schlabach et al. *Science* 319, 620-4 (2008).
19. J. R. Silva et al. 319, 617-20 (2008).
20. T.F. Westbrook, et al. *Cell* 121, 837-48 (2005).
21. T.F. Westbrook, et al *Nature* 452:370-4 (2008).
22. P. Mulligan, et al. *Mol Cell.* 32:718-26 (2008).
23. J.A. Chong, J. A. et al. *Cell* 80, 949-57 (1995).
24. C. J. Schoenherr and D. J. Anderson. *Science* 267, 1360-3 (1995).
25. N. Ballas, and G. Mandel, *Curr Opin Neurobiol* 15, 500-6 (2005).
26. G.N. Fuller, et al. *Mol Cancer Ther* 4, 343-9 (2005).
27. X. Su, et al. *Mol Cell Biol* 26, 1666-78 (2006).
28. Y. Kudo, et al. *Mol Cell Biol* 24, 8184-94 (2004).
29. T. Saitoh, and M. Katoh, *Int J Oncol* 18, 959-64 (2001).
30. F.K. Noubissi, et al. *Nature* 441, 898-901 (2006).
31. F. Stegmeier et al. *Proc. Natl. Acad. Sci. USA*. 102:13212-7.
32. N.L. Solimini, J., Luo, and S.J., Elledge *Cell*. 130:986-8 (2007).
33. J. Luo, N.L. Solimini and S.J. Elledge SJ. *Cell*. 136:823-37 (2009).
34. A. E. Karnoub and R. A. Weinberg, Nat Rev Mol Cell Biol *9*, 517-531 (2008).
35. C. J. Torrance,et al. Nat Biotechnol *19*, 940-945 2001).
36. S. Dolma, et al. Cancer Cell *3*, 285-296 (2003).
37. J.A. Engelman, et al. Nat Med *14*, 1351-1356 (2008).
38. J. Luo et al. *Cell* 137:835-48 (2009).