AWARD NUMBER: W81XWH-06-1-0253

TITLE Novel Array-Based Target Identification for Synergistic Sensitization of Breast Cancer to Herceptin

PRINCIPAL INVESTIGATOR: Daniel Mercola M.D., Ph.D.

CONTRACTING ORGANIZATION: University of California, Irvine
Irvine, CA 92697

REPORT DATE: March 2008

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
March 2008	**Annual**	15Feb2007 - 14Feb2008

4. TITLE AND SUBTITLE

"Novel Array-Based Target Identification for Synergistic Sensitization of Breast Cancer to Herceptin"

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0253

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Daniel Mercola M.D., Ph.D.
E-Mail: dmercola@uci.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Irvine
Medical Sciences D440
Irvine, Ca 92697-4800

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

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

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During the current six month funding period (9/07-2/08) we have built on the results of the prior progress report to use the methods of chromatins immunoprecipitation developed with Jun Kinase and applied these to examine HER2-dependent binding of RNA Polymerase II (RNPII) to the DNA of human breast cancer cell lines MCF7HER2, MCF7PCDNA, BT474, and MDA453. The goal is identify genes that the oncogene HER2 specifically stimulates for transcription or 'primes' for transcription. The genes of the chromatin immunoprecipitatioin (ChIP) prepared DNA are identified by hybridizing the ChIP DNA to our in-house prepared arrays as for Jun Kinase and also Agilent promoter arrays with ~17,000 genes represented.

15. SUBJECT TERMS
Prostate cancer/formalin-fixed paraffin-embedded/diagnosis/microenvironment/stroma/validation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
U					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified	UU	9	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Page

ANNUAL PROGRESS REPORT 9/2007 – 2/2008.
Novel Array-Based Target Identification for Synergistic Sensitization of Breast Cancer to Herceptin

1. INTRODUCTION: Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.

During the current funding period (9/07-2/08) we have built on the results of the prior progress report to use the methods of chromatins immunoprecipitation developed with Jun Kinase and applied these to examine HER2-dependent binding of RNA Polymerase II (RNPII) to the DNA of human breast cancer cell lines MCF7HER2, MCF7PCDNA, BT474, and MDA453. The goal is identify genes that the oncogene HER2 specifically stimulates for transcription or 'primes' for transcription. The genes of the chromatin immunoprecipitatioin (ChIP) prepared DNA are identified by hybridizing the ChIP DNA to our in-house prepared arrays as for Jun Kinase and also Agilent promoter arrays with ~17,000 genes represented.

2, KEYWORDS. Breast cancer/oncogene/her2/expression analysis/chromatin immunoprecipitation/herceptin

3. OVERALL PROJECT SUMMARY.

Anti-RNAPII captured cellular DNA was hybridized to the arrays in competition with an equal amount of total DNA [content] from the same cross-linked cells without addition of antibody (Whole Cell Extract, WCA). In these studies we used the same hybridization protocol as for both in-house and Agilent arrays. Our arrays contain over 15,000 elements consisting of 5' regulatory sequence from 500 to 1500 kb in length and printed in triplicate (Arora ref). For these arrays hybridization is repeated using a reversed order of dye application for the labeling of ChIP and WGE DNA (dye-swap) thereby providing six hybridization intensity measurements in total for a given DNA sample. We compared biological duplicates, six measurements *vs.* six measurements. The criteria for significant binding by immuno-precepitated DNA compared to WGE is, $p < 0.05$, $B > 0$, as previously described in our prior report. The binding of RNPII to over 600 genes promoters satisfied these criteria. Biological duplicates have been examined for BT474 cells. Gene identification is 88.2% identical for the biological duplicates vs. random expectation of 12.8%. This is in good agreement with reproducibility of 79% observed using the Agilent arrays.. The sum of observations indicates that the methods used for cell culture and Chip DNA preparation, amplification and hybridization lead to reproducible RNPII DNA binding values. In addition we validated RNPII binding sites exclusively in the genome of MCF7-HER2 by utilizing traditional and quantitative ChIP-PCR. A negative control without antibody and non reactive antibody IgG were used for comparisons. We observed a good agreement between the RNPII binding sites detected by Agilent algorithm and the binding event that are observed by quantitative ChIP-PCR.,

HER2 Promotes a Large Scale Redistribution of RNPII binding among and within genes.

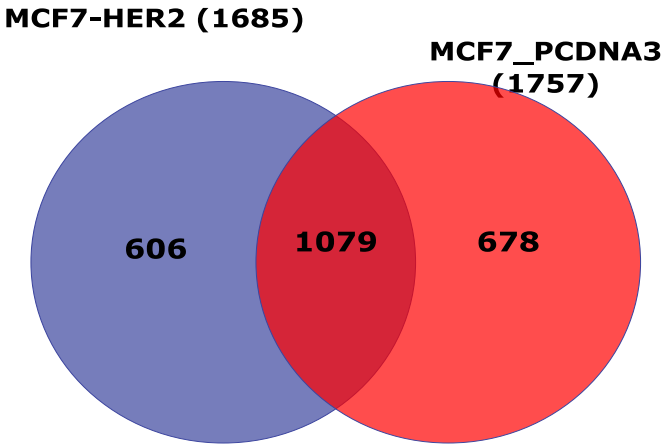

Figure 1. Venn Diagram of the overlap of genes with differential RNPII binding in MCF7 breast cancer cell line which exogenously expresses full length HER2 and control MCF7 cells with empty PCDNA3 plasmid. ChIP-chip analysis of these cell lines identified 606 genes that gained RNPII binding site and 678 genes that lost RNPII binding upon expression of HER2 in MCF7 breast cancer cell line. The two cell types share 1079 (45%)bound RNPII binding genes that are unchanged by HER2 expression and 1284 genes are shown altered POLII bindings (60%) indicating a large effect of HER2 in POLII binding . The coincidence (or p-value) of the observed overlap is less than $p<0.0001$.

We have identified the genes that bind RNPII by high HER2 expressing MCF7HER2 cells and HER2-null MCF7PCDNA cells. We observed a large number of genes in both MCF7HER2 and MCF7PCDNA cells that met our criteria for significant binding of RNPII. 2363 genes for the Agilent arrays of ~17,000 genes (**Figure 1**). Comparison of the distribution of binding of RNPII in HER2-expressing cells to control cells revealed a striking pattern of rearrangement upon the expression of HER2. The Venn diagram of **Figure 1** 1indicates that HER2-expressing cells exhibit 606 RNPII bound genes *not* bound by the control cells. 678 genes that bind RNPII in control cells are lost upon expression of HER2. These observations indicate that RNPII is commonly chromatin bound in MCF7 cells and that express of HER2 promotes a shift in RNPII s binding to new gene sites not present in MCFR7 cells.

1079 other genes bind RNPII in both cell types. Moreover, numerous genes exhibit multiple RNPII bound sites, up to 10 sites/gene. The number of bound RNPII sites per gene is commonly altered in HER2-expressing cells both by gain and loss of RNPII binding. Moreover this picture is underestimate the numbers of genes involved because Agilent only samples DNA in the range -5kb to 2 kb about the transcription start site and only provide information on about 17000 of 23000 genes..

In order to further understand the basis of the binding of RNPII by genes especially the binding of RNPII at multiple sites, we examined the affinity of RNPII for various genes. The bind probability for array data reflects the amount of hybridization or the amount of RNPII-DNA complex formation. High complex formation in turn reflects affinity of RNPII for a site. High complex formation leads to low (highly significant) probabilities. In contrast low complex formation may reflect loose binding and mobile or transcribing RNPII.

To compare many genes with different numbers of RNPII sites, we calculated a joint probability for the gene and expressed the probability as per site for genes with multiple sites**.** **Figure 2** summarizes the results. We find for genes with multiple sites (from n=1 to n=13) the probability per site decreases with n. The genes that bind RNPII at many sites have "tight" or high affinity sites. This provides a consistent picture of the origin of multiple binding as occurring in genes which "maintain" high affinity sites. Such genes are predicted to have poor RNPII mobility and poor transcription. Since the probability-number relationship of **Figure 2** is similar for MCF7HER2 and MCF7PCDNA. In all, 33 % of 2363 significant RNPII binding probes are associated with multiple binding (i.e. %30 with $n \geq 2$ 396/1284 , gain and loss).

Correlation of gene expression with HER2-dependent RNPII binding.

We have begun to study the relation of HER2-expression to RNA expression using data from Affymetrix U133 plus 2.0 expression arrays with approximately 54,000 probe sets prepared in our lab.. The two microarray platforms (Affy and Agilent) share 14231 unique genes making it possible to assess RNA expression by most gene with bound RNPII. Total steady state RNA was prepared from cells grown exactly as for the ChIP-chip analysis. The RNA expression results pertain largely to full-length transcription owing to the preponderance of Affymetrix probe sets being located in the 3' end of the coding sequence. Expression analysis was initially carried using a single RNA preparation for each of the MCF7 cell lines and repeated later as biological duplicates for both MCF7HER2 and MCF7PCDNA cells, 6 arrays in all. An agreement analysis revealed a Pearson Correlation Coefficient of 0.92 and $\boldsymbol{p} < 2x10^{-6}$

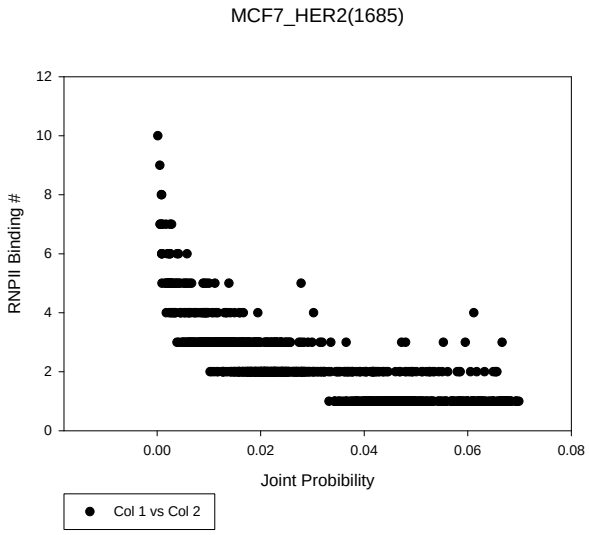

Figure 2. :To define the interaction of RNPII and a DNA binding site number, n, we used a modified probability for a single, p_x_bar, as described by Agilent Inc. When more than one site with $p<0.05$ were observed within +5000/-2500 bp of a given gene's transcriptional start site (i.e. the zone probed by the Agilent array), we derived joint p values, p_{jt}, and the geometric average probability, p_{gav}.

Of the 606 genes with RNPII binding sites (exclusively) in MCF7-HER2, we observed 139 (23%) genes with significantly altered RNA expression levels compared to MCF7PCDNA cells. 678 genes of MCF7PCDNA (control cell) that are lost upon expression of HER2 protein also exhibit significant differential transcription. However transcription by the HER2-expressing cells is significantly higher levels than those genes that lost their RNPII binding sites. Wilcoxon signed rank test was performed and showed a significant difference ($\boldsymbol{p} < 0.001$) in global gene expression levels of those with enriched RNPII compared to genes with depleted RNPII upon expression of HER2. Thus in general HER2 promotes increased full-length gene expression.

In order to further test whether these transcription changes likely represent HER2-dependent alterations, we compared the percent of RNPII sites exhibiting these changes with the frequency of significant changes found in rest of the array by Chi-square analysis which yielded probabilities of ($\boldsymbol{p}<0.001$) and ($\boldsymbol{p}<0.001$) for MCF7HER2 and MCF7PCDNA cells respectively indicating that both loss and gain of RNPII is indeed associated with functional transcriptional changes.

We tested the conjecture that tight binding of RNPII impedes transcription we examined the relationship of binding probability and transcription. We observed that most RNPII binding events downstream of transcription start site (TSS) were associated with smaller number of RNPII binding sites (n=1) with increased transcription compared to stalled near the TSS. We asked whether tightly bound sites were preferentially located with respect to coding DNA. In order to quantify this relationship, the stalling index, SI, as defined by Young and coworkers for RNPII ([1,2] was calculated. Briefly, the stalling index is calculated as the ratio of the number of RNPII sites within 600 bp of the transcriptional start site over the total number of sites within 2500 bp. Stalling index values >4 are considered as "stalled" or paused RNPII complexes not involved in transcription whereas stalling index <4 are considered as "active RNPII". Genes with no significant RNPII

binding sites are categorized as "inactive". The Agilent promoter arrays have an average probe spacing of 196-200 bp however, probe sites do not extend beyond 2500pb downstream of the transcription start site and therefore SI is insensitive to binding of RNPII in genes with coding sequences beyond 2500 bp 3' to the transcription start site. Within this limited range of genes with single RNPII binding sites, n =1, exhibit stronger binding (lower probability) for sites near the transcription start site than for downstream positions (**Figure 3**). Moreover, as n increases, the increased binding of proximally bound sites becomes more apparent. For genes with multiple binding the difference in binding strength of proximal and downstream sites as judged by analysis of probability is significant indicating significant decreased binding strength for downstream bound RNPII. In terms of the definition of SI the observation suggests active or mobile RNPII sites are indeed associated with weaker binding. This relationship is general for all significantly bound RNPII sites of the MCF7HER2 model system and all sites lost (678) or gained (606) upon expression of HER2.

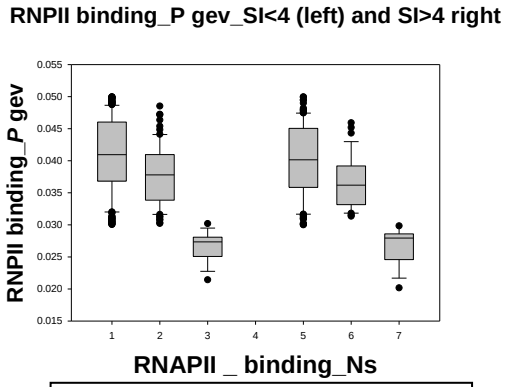

Figure 3.

The common (1078) sites reflect a similar phenomenon (**Figure1**). These sites retain at least one significant site but exhibit an increase or decrease in number upon expression of HER2. Again proximally located sites exhibit significantly higher binding strength examined by probability. Examination of SI also reveals a marked asymmetric redistribution of RNPII binding upon expression of HER2. 502 out of 678 (74%) significantly bound RNPII sites in MCF7PCDNA that are lost upon expression of HER2 are derived from downstream "active" sites while 404/606 genes (67%) of newly recognized sites in MCF7HER2 cells are in downstream "active" positions. Thus 564/1284 (42%) genes with only one RNPII binding sites, gained or lost, are predominately derived from downstream sites.

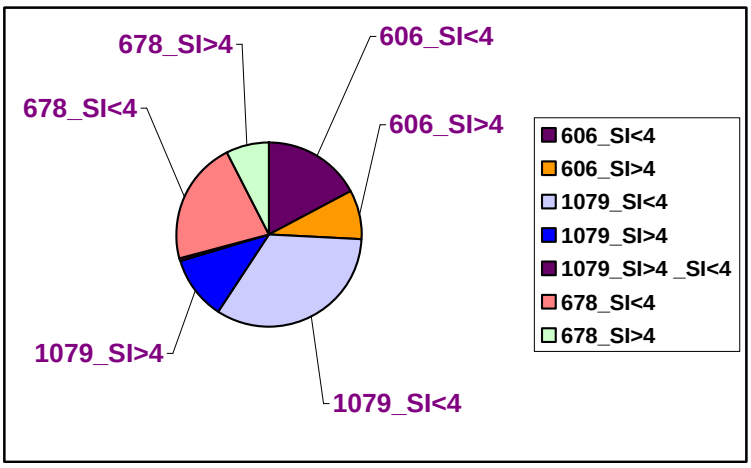

Figure 4. Pie chart of distribution of RNPII binding near the TSS (SI>4) or downstream (SI<4) sites in both cell lines.

The SI calculations show that of the 2301 (**Figure 3** Venn diagram) RNPII-bound genes with binding of RNPII to MCF7HER2 and MCF7PCDNA promoters and the coding region, 16.6 % (399/2301) had gained RNPII binding sites down stream of TSS whereas 28% (688/2301) genes had lost RNPII binding sites down stream of TSS upon Expression of HER2 oncogene (**Figure 3**). Moreover, 8.5 % (208/2301) of RNPII binding occurs near the TSS in MCF7-HER2 cells, whereas, 7.5% genes lost RNPII binding site around the TSS in HER2 over expressing cells (**Figure 4**). Compared to the previously identified 606 genes (newly bound RNPII Genes) in MCF7-HER2 cells, 396 (65%) had RNPII binding sites down stream of TSS and 205 (33%) had RNPII stalled at TSS. Only 1% (28) of RNPII bound genes had binding sites down stream of TSS in MCF7HER2 cells and stalled at the TSS in control MCF7cells These results illustrate a large effect of HER2 overexpression in shifting the RNPII binding toward the coding region,

downstream of TSS (coding region). This result is consistent with the observed significant increased transcription of MCF7HER2 cells compared to control cells.

4. KEY RESEARCH ACCOMPLISHMENTS: Bulleted list of key research accomplishments emanating from this research.

- Applied chromatin immunoprecipitation to our model system of MCF7HE2 and MCF7PCDNA cells. Used Agilent promoter arrays to hybridize the ChIP DNA to the arrays in parallel with home-made promoter arrays. Identified over 2300 genes on the Agilent arrays. Showed that where there was overlap with the home-made arrays there was excellent agreement.
- Obtained expression data in biological duplicates for the model cell system and identified the genes with RNPII binding that exhibited significant transcription Identified the HER2-dependent transcribing genes that do not transcribe in control cells: the HER2-dependent transcribing genes.
- We investigated the basis for transcribing vs. non-transcribing genes. We showed that very tight bing of RNPII was associated with genes with multiple tight binding sites which tended to be located near the transcription start sites. Identified HER2-dependent poised and expressed genes in human breast cancer cell lines.
- We showed the HER2 expressing cells had statistically significant great transcription with sites located in 3' positions consistent with RNPII that transverses the gene leading to full-length transcription.

5. PUBLICATIONS, ABSTRACTS, PRESENTATIONS.

(Nothing to report)

6. INVENTIONS, PATENTS, LICENCES. (Nothing to Report)

7. CONCLUSION. Summarize the results to include the importance and/or implications of the completed research and when necessary, recommend changes on future work to better address the problem. A "so what section" which evaluates the knowledge as a scientific or medical product shall also be included in the conclusion of the report.

Overall, the results during this funding period have shown that we can identify HER2-dependent RNPII binding, determine which of the identified genes are involved in transcription, and have learned something about the basic relationship between 'tightness' of RNPII binding and transcription. In general genes with high affinity for RNPII bind multiple RNPII molecules mostly near the TSS and transcribe less than genes with loose binding undoubtedly many factors control tightness of binding including chromatin modification and the availability of numerous initiation factors. HER2 alone is not the sole director but expression of HER2 leads to loser RNPII binding, downstream binding and increased full-length transcription.

A future goal will be to examine the FNPII bound genes in actual breast cancer tissues where HER2 and all the essential factors of expression or suppression are present to .identify where HER2 genes are are altered in their regulation compared to cell line. Although not in the CDMPR SOW, this will provide essential information of how HER2 works.

For the next funding period we will extend the above studies to BCa cell lines BTA454, MDA453 and others in order to generalize the above results to real HER2-ajmplief tumor cells. We will determine which of

the identified genes both with and without expression are inhibited or increased in expression by the addition of Herceptin. We will examine the function nature of the genes and pathways involved.

8. REPORTABLE OUTCOMES.

Dr. R. Rahmatpanah was awarded a training grant fellowship position of the UCI Cancer Research Institute, May, 2009 for 2 years.

9. OTHER ACHIEVEMENTS. (Nothing to Report)

10. **REFERENCES**: List all references pertinent to the report using a standard journal format (i.e. format used in *Science, Military Medicine*, etc.).

1 Zeitlinger, J. *et al.* RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat Genet* **39**, 1512-1516, doi:ng.2007.26 [pii]10.1038/ng.2007.26 (2007).

2 Seila, A. C. *et al.* Divergent transcription from active promoters. *Science* **322**, 1849-1851, doi:1162253 [pii] 10.1126/science.1162253 (2008).