Author's response to reviews

Title: Expression profiling of blood samples from an SU5416 Phase III metastatic colorectal cancer clinical trial: a novel strategy for biomarker identification.

Authors:

- Dr Samuel E DePrimo (samuel-deprimo@sugen.com)
- Lily M Wong (lily-wong@sugen.com)
- Dr Deepak B Khatry (deepak-khatry@sugen.com)
- Susan L Nicholas (susan-nicholas@sugen.com)
- William C Manning (bill-manning@sugen.com)
- Beverly D Smolich (beverly-smolich@sugen.com)
- Anne-Marie O'Farrell (marie-ofarrell@sugen.com)
- Julie M Cherrington (julie-cherrington@sugen.com)

Version: 2 Date: 14 Jan 2003

PDF covering letter
Dear Dr. Veitch:

We are pleased to learn that our manuscript, “Expression profiling of blood samples from an SU5416 Phase III metastatic colorectal cancer clinical trial: a novel strategy for biomarker identification”, is potentially suitable for publication in *BMC Cancer*. We thank the reviewers for their insightful and constructive comments regarding the manuscript. We also appreciate the fact that both reviewers acknowledged the validity of the approach and recognized that the main import of this study is establishment of proof of concept for the approach. Our responses to the reviewers’ comments are listed below.

**Compulsory revisions (Dr. Chinnaiyan):**

1. The typographical error has been corrected, and the blank page (pg 32) has been deleted.

2. Dr. Chinnaiyan rightly makes the point that there are other analytical approaches besides the ones we used, that are applicable to mining of microarray datasets. A methodology such as SAM analysis is certainly an accepted and rigorous option. However, we do believe that the approach we have used, wherein we identified candidates according to p-values and Affymetrix Difference Call scores, was appropriate to the goal of investigating the feasibility of blood as a surrogate tissue. We successfully used independent verification via qRT-PCR to validate the significance of the candidate biomarker transcripts. Nevertheless, as suggested, we have done some analysis with the SAM program, wherein we submitted the 500 most significant genes (SU5416 vs treated) as defined by the Spotfire t-test to SAM analysis. Using 300 permutations of these data, we found that all 4 of the markers we independently validated by qRT-PCR are among the 16 most significant genes as measured by SAM. In fact, most of the genes listed in Table 1 are ranked near the top of the significance list. We have now briefly summarized these results in the second paragraph of the Discussion section. Thus, it appears that we would have identified many of the same genes by this approach as we did in our described initial approach. At any rate, independent confirmation of all potential biomarkers identified by either method is not within the scope of this study.

3. Regarding the histogram representations of data in Figure 2, we have designated a threshold of ‘2-fold’ to indicate up-regulation for the RT-PCR data; this was not based on statistical rigor but instead chosen to illustrate that there is a difference in the number of cases of up-regulation between the 2 patient groups for at least some of the transcripts that were tested. We relied on results of the Mann-Whitney U test to indicate significant
differences between the 2 groups, and pursued this further with samples from the second trial and in the discriminant analysis as described in Table 5. For the Affymetrix data, samples were scored as increased or not based on Difference Calls (which are displayed in Figure 1) rather than on fold-change. The legend to Figure 2 has been slightly modified to more clearly describe this. The suggestion to include actual measurements from the analysis is a good one. We have added a table, in Additional File 1, wherein we include all of the relative expression values for patients from both trials A and B. This file is referenced in the legends to Figures 2, 3, and 4, as all of these figures represent some or all of the values listed in the table. We have not sought to measure absolute expression levels in our RT-PCR analysis, but rather used a relative expression approach, thus we are not able to provide absolute levels measured in the individual RNA samples (raw data is in the form of Ct scores).

Discretionary Revisions:

Dr. Miller recommended acceptance without revision. His point regarding additional in vitro studies to address the possible mechanisms responsible for induction of these transcripts in the SU5416 arm is well taken, and we acknowledge that these questions remain unanswered. In general, these issues might be difficult or impossible to resolve using in vitro approaches, as the changes seen in patient samples are presumably happening during several days or weeks of systemic therapy, and thus, for example, a negative result in vitro may not translate to a lack of effect in vivo. Also, it might be necessary to test primary PBMC cultures from several donors, in order to capture the inter-individual variability typically seen in patients. Further, from a practical point of view in terms of future studies, we anticipate application of our approach to clinical investigations of newer therapeutic agents that are orally available and thus will not involve vehicle effects or a need for concomitant medications.

Dr. Miller’s comments regarding utilizing protein as a potentially more reliable biomarker is valid, and we completely agree that may well be the case for many surrogate biomarkers. The greatest utility of the gene expression profiling approach, we feel, lies in the opportunity to identify potential biomarkers, which can then be further applied using whatever assay is most robust and practical.

Some preliminary analysis to address expression differences in responders vs non-responders in the GeneChip dataset has been done, using a similar approach to the one we describe, but results were much less convincing than for SU5416 vs control arm. Indeed, we agree that markers of response are the most valuable surrogate markers, and this is a primary objective in new studies.

Regarding Dr. Chinnaiyan’s comment regarding inclusion of gene expression data as supplementary material, we would prefer not to include this data at present as it is related to a recent clinical trial.

Sincerely,

Sam DePrimo