Applying high throughput to CLL

By Chris Cain, Senior Writer

Boston researchers have provided a detailed look at how cancer genome evolution alters clinical outcomes in chronic lymphocytic leukemia,¹ and a Michigan team has identified a new transcriptional fusion that could broadly contribute to the pathogenesis of CLL.² Both studies provide insights that could guide the development of new diagnostics and therapeutics for patients with CLL.

CLL is the second most common type of adult leukemia and has highly variable clinical progression. To better understand the origin of this variability, multiple groups have set out to conduct high throughput DNA sequencing of CLL tumor samples from patients. In the last two years, whole-exome analysis of hundreds of these tumors has shown that the disease is associated with a large number of mutations, with no one predominant mutation or pathway.³⁻⁵

“The mutational landscape of CLL that is emerging from sequencing studies is extremely challenging, in that most recurrent mutations are present at low frequencies,” said Victor Quesada, a postdoctoral fellow at the **University of Oviedo** who led one of those sequencing studies. “At this point, we need to gather information to characterize one of at least two scenarios: either there is an underlying mechanism that connects all of the driver mutations or there are many independent mechanisms that may lead to CLL. So far, the second scenario has seemed more likely.”

Now, two teams have applied new high throughput sequencing data analysis methods to gain further insight into the mechanisms underlying CLL disease progression.

A joint team from the **Dana-Farber Cancer Institute** and the **Broad Institute of MIT and Harvard** focused on tumor DNA and used whole-exome sequencing to measure mutational heterogeneity in patient samples taken before and after chemotherapy. The findings suggest that identifying driver mutations present in only small fractions of cells could help predict the clinical course of CLL.

The team expects that its computational tools could be broadly applied to measure tumor heterogeneity in a variety of cancers.

University of Michigan Medical School researchers instead focused on tumor RNA and used a recently developed approach for whole-transcriptome data analysis to identify new chimeric transcripts comprised of normally disconnected genes. The group found one such transcript in more than 95% of CLL samples and presented preliminary data showing that the truncated protein phosphatase it encodes could be oncogenic.

However, no DNA mutations or translocations were found that could explain how the chimera is generated, and validation in additional samples will be needed before its significance becomes clear.

Evolving thinking

Recent studies have uncovered substantial mutational heterogeneity within individual tumor samples, but the reliance of those studies on single-cell methods or whole-genome deep sequencing limited the numbers of samples that could be analyzed.

For example, two of the largest tumor evolution studies in 2012 measured the mutational landscape of a total of about 30 acute myelogenous leukemia (AML) samples.^{6,7}

Because of the genetic and clinical variation in CLL, researchers at Dana-Farber wanted to scale up the analysis to ensure it could comprehensively capture diversity. They teamed up with a Broad Institute team to apply recently developed computational algorithms for DNA sequence analysis to whole-exome data, which is far less costly to produce than whole-genome sequencing data.^{8,9}

Gad Getz, director of cancer genome computational analysis at the Broad Institute and co-corresponding author on the study, told *SciBX* that two algorithms developed in his lab, MuTect and ABSOLUTE, were essential to handle the complexity of quantifying populations of mutations in whole-exome sequencing data derived from bulk tumors.

In heterogeneous tumor samples, some copies of a given DNA sequence may be wild-type whereas others may be mutated. This wild-type/mutant ratio is known as an allelic fraction, and analysis of the allelic fraction can predict which mutations are clonal—found in most or all cancer cells in the sample—and which are subclonal, meaning those found only in some cancer cells in the sample.

“MuTect enabled the detection of mutations that have a low allelic fraction and therefore are present in smaller subclones,” Getz said. “ABSOLUTE enables the conversion of allelic fractions to the fraction of cancer cells harboring the mutation.”

The joint team used these two approaches and additional algorithms to measure the prevalence of mutations and copy number variations within whole-exome data generated from 149 pairs of matched CLL and germline DNA samples.

The group identified 1,543 clonal mutations and 1,266 subclonal mutations in total. On average, there were about 10 clonal mutations per sample and about 5 subclonal mutations.

Catherine Wu, co-corresponding author of the study, told *SciBX* that the scale of the analysis provides new insights into the heterogeneity of patients with CLL.

“By studying such a large cohort we were able to trace back the genetic history of the disease,” she said. “We observed, for example, a higher proportion of clonal mutations in older individuals, probably reflecting the accumulation of passenger

“This work shows that pretreatment clonal heterogeneity and specifically the presence of subclonal drivers can predict clinical outcomes.”

—Catherine Wu,
Dana-Farber Cancer Institute

mutations throughout life in pretransformed cells. In addition, we were able to define groups of recurrent putative driver mutations that appear earlier and later along the disease history of CLL. The earlier mutations tended to be more B cell specific, whereas the later ones seemed to be more generic cancer drivers found across many tumor types. This suggests interesting relationships between earlier and later events that we hope to study further in model systems.”

Wu is an associate physician of medical oncology at Dana-Farber and associate professor of medicine at **Harvard Medical School**.

The team then analyzed the effect of therapeutics on the evolution of tumor mutational heterogeneity.

In 12 matched patient samples taken before and after treatment with chemotherapy or anti-CD20 antibodies, the sequencing analysis identified driver mutations that increased in prevalence after treatment and predicted poorer survival compared with what was seen in patients lacking the mutations.

Treatment was associated with an expansion of cells carrying these driver mutations, as 10 of 12 treated patients experienced clonal evolution, whereas only 1 of 5 samples from untreated patients analyzed at 2 time points underwent clonal evolution.

Results were published in *Cell*.

Wu told *SciBX* that these results could have clinical implications. “This work shows that pretreatment clonal heterogeneity and specifically the presence of subclonal drivers can predict clinical outcomes,” she said.

Quesada agreed. “Probably the most useful consequence of this work is the ability to predict the evolution of the disease before the worst symptoms arise. This is one of the most pressing issues in CLL clinical management,” he said.

“Nowadays, clinicians need to wait for the symptoms to develop before providing a treatment. The risk for overtreatment is too high if no additional information can be gathered.”

He added, “The fact that mutations that are associated with poor prognosis are already present in subclones of the CLL cells before progression begins means that clinicians may now consider treatment at early stages, which might impact on both overall and progression-free survival. This strategy has been dubbed anticipation-based chemotherapy.”

Getz said the team now plans to test these predictions in prospective clinical trials of patients with CLL. “We also plan to extend this work to other types of malignancies and evaluate to what degree the insights found in CLL are generalizable to cancer biology,” he said.

The computational tools used in this study have been patented and are available for licensing from the Broad Institute. Getz said his lab has made the tools freely available to academic and not-for-profit organizations.

Chimeric RNA riddle

Although DNA sequencing efforts have characterized the genetic landscape of CLL, the Michigan team turned to RNA sequencing to identify new transcripts that could drive disease progression but would not be detected solely through genome analysis.

In the last several years, RNA sequencing analyses by the lab of Arul Chinnaiyan, professor of pathology at the University of Michigan Medical School, have led to the identification of many new

cancer-associated chimeric transcripts—RNA species carrying sequences normally found in separate genes—with diagnostic and therapeutic potential in prostate cancer.¹⁰

The most obvious cause of a chimeric transcript would be a chromosomal translocation event, such as the well-known fusion of the *BCR* and *ABL* genes. However, at least one recently identified chimera is not associated with DNA translocation and is likely to be produced by splicing, suggesting that traditional DNA sequencing could miss fusion transcripts or proteins with oncogenic potential.¹¹

To look for such fusions in CLL, Kojo Elenitoba-Johnson, director of translational pathology at the University of Michigan Medical School, teamed up with Chinnaiyan to sequence the transcriptome of seven CLL patient samples. This led to the identification of RNA fusions of two transcripts from different chromosomes, *yippee-like 5* (*YPEL5*) and *protein phosphatase 1 catalytic subunit β -isozyme* (*PPP1CB*), expressed as either *YPEL5-PPP1CB* or *PPP1CB-YPEL5*.

The *PPP1CB-YPEL5* transcript was predicted to produce wild-type *YPEL5*, whereas the *YPEL5-PPP1CB* transcript was predicted to produce a truncated form of *PPP1CB* lacking its first 28 amino acids.

One or both transcripts were detected in 97 of 103 CLL samples (95%). Neither were detected in noncancerous B cells or in a variety of other leukemia or lymphoma samples.

The team tried numerous techniques to detect a DNA fusion between *YPEL5* and *PPP1CB* but found nothing. Elenitoba-Johnson told *SciBX* that result made the team’s finding even more intriguing. “We worked hard at establishing whether there was a genomic origin of the chimera—which we notably did not show—and part of the interest in this study is that it provides further evidence that you can have an RNA chimera that is not driven by a genomic event that may participate in the pathogenesis of some forms of cancer,” he said.

To test the functional significance of *YPEL5-PPP1CB*, the team expressed the transcript in cell culture and confirmed that it produced truncated *PPP1CB* protein. *In vitro*, this purified protein had less phosphatase activity than purified, full-length *PPP1CB*. In cultured leukemia cell lines, small hairpin RNA against *PPP1CB* increased proliferation compared with control shRNA, suggesting that reduced *PPP1CB* function could have an oncogenic effect.

Results were published in the *Proceedings of the National Academy of Sciences*.

Quesada said the results are promising but emphasized that they must be validated by independent labs. “In this study, we have the first genomic hint that there is indeed a common biochemical mechanism at the origin of CLL,” he said. “In my opinion the first order of business is to independently confirm these results. Even though the authors provide ample and satisfying evidence for their case, confirmation from other laboratories is a must for an important discovery like this.”

Elenitoba-Johnson agreed that the result must be independently investigated and said he expects it to be quickly picked up by other labs.

“We suspect that due to the attention this result is getting because of its frequency, other labs will rapidly try to reproduce this finding

“In this study, we have the first genomic hint that there is indeed a common biochemical mechanism at the origin of CLL.”

**—Victor Quesada,
University of Oviedo**

in additional populations. After that, the next step will be to look and see how early this transcript occurs in the evolutionary process of the disease,” he said.

He noted that although the functional evidence is still early, one future direction could be to screen for compounds that activate PPP1CB and test their effects in CLL.

Patent and licensing details were undisclosed.

Cain, C. *SciBX* 6(11); doi:10.1038/scibx.2013.253

Published online March 21, 2013

REFERENCES

1. Landau, D.A. *et al. Cell*; published online Feb. 14, 2013; doi:10.1016/j.cell.2013.01.019
Contact: Catherine J. Wu, Dana-Farber Cancer Institute, Boston, Mass.
e-mail: cwu@partners.org
Contact: Gad Getz, Broad Institute of MIT and Harvard, Cambridge, Mass.
e-mail: gadgetz@broadinstitute.org
2. Velusamy, T. *et al. Proc. Natl. Acad. Sci. USA*; published online Feb. 4, 2013; doi:10.1073/pnas.1214326110
Contact: Kojo S.J. Elenitoba-Johnson, University of Michigan Medical School, Ann Arbor, Mich.
e-mail: kojoelen@med.umich.edu
Contact: Arul M. Chinnaiyan, same affiliation as above
e-mail: arul@umich.edu
3. Puente, X.S. *et al. Nature* **475**, 101–105 (2011)
4. Quesada, V. *et al. Nat. Genet.* **44**, 47–52 (2011)
5. Wang, L. *et al. N. Engl. J. Med.* **365**, 2497–2506 (2011)
6. Cibulskis, K. *et al. Nat. Biotechnol.* **31**, 213–219 (2013)
7. Ding, L. *et al. Nature* **481**, 506–510 (2012)
8. Welch, J.S. *et al. Cell* **150**, 264–278 (2012)
9. Carter, S.L. *et al. Nat. Biotechnol.* **30**, 413–421 (2012)
10. Maher, C.A. *et al. Nature* **458**, 97–101 (2009)
11. Zhang, Y. *et al. Cancer Discov.* **2**, 598–607 (2012)

COMPANIES AND INSTITUTIONS MENTIONED

Broad Institute of MIT and Harvard, Cambridge, Mass.
Dana-Farber Cancer Institute, Boston, Mass.
Harvard Medical School, Boston, Mass.
University of Michigan Medical School, Ann Arbor, Mich.
University of Oviedo, Oviedo, Spain