Kinome Sanger Sequencing

Acute Lymphoblastic Leukemia Phase I (ALL P1)

*Protocols performed at British Columbia Cancer Agency. Please refer to *Loh et al.* (Tyrosine kinome sequencing of pediatric acute lymphoblastic leukemia: a report from the Children's Oncology Group TARGET Project, Blood. 2013 Jan 17; 121(3): 485–488).

Patient selection and characteristics

Forty-five cryopreserved diagnostic bone marrow or peripheral blood specimens with at least 80% blasts from children with newly diagnosed ALL were selected for kinome sequencing (Table 1), including 23 from COG P9906[17] that lacked JAK mutations and 22 AALL0232 patients of unknown JAK mutation status. AALL0232 eligibility included age at least 10 years and/or initial peripheral blood white blood cell count of at least 50 000/μL. Minimal residual disease (MRD) burden was determined via flow cytometry in 1 of 2 central reference laboratories at day 29 of induction therapy.[3] All P9906/AALL0232 patients or their patients/guardians provided informed consent for treatment and for banking of specimens for future research in accordance with the Declaration of Helsinki. Institutional review board approval for the laboratory studies was granted by St Jude Children's Research Hospital and the University of New Mexico.

GEP and sample selection

RNA extraction and GEP characterization for P9906 cases have been described previously.[9,11] Affymetrix U133 Plus Version 2.0 gene expression microarray and Affymetrix SNP Version 6.0 microarray profiling were performed on 608 patients consecutively enrolled on AALL0232 with sufficient banked material available; 325 were used as a training set, after modeling the Ph-like GEP on the *BCR-ABL1*$^+$ patients within this training set (n = 21).
We then applied Prediction Analysis for Microarrays (PAM),[18] trained using Ph$^+$ cases to identify all Ph-like cases (supplemental Figure 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).[19] We classified patients in the test set (283 AALL0232 patients) using this Ph-like signature and assessed the outcome of all Ph-like patients enrolled on AALL0232. We further applied this Ph-like PAM algorithm to the COG P9906 samples to identify Ph-like cases in that cohort, and then assessed the prognosis of this group of ALL cases.

We selected 45 P9906 and AALL0232 cases that were either predicted to be Ph-like by PAM (31 cases; 12 of 23 from 9906 and 19 of 22 from AALL0232), or had high *CRLF2* expression or other features suggestive of activated kinase signaling (n = 14) for sequence analysis of 126 genes that encode TKs or mediators of kinase signaling (supplemental Table 1). The entire coding and untranslated regions of each selected gene were subsequently amplified by PCR of whole genome amplified (QIAGEN) genomic DNA and subjected to Sanger sequencing (Beckman Coulter Genomics). A CEPH sample (NA19085) was included as a normal control. Sequence variations were detected using SNPdetector[20] and novel, putative nonsilent coding mutations were selected for validation. Forty-one novel variants that failed in the validation assay or had no matching germline samples were compared with germline variants identified by the National Center for Biotechnology Information Exome Sequencing Project

(http://evs.gs.washington.edu/EVS(link is external)) and 1000 Genomes Project deposited in dbSNP 135 (http://www.ncbi.nlm.nih.gov/projects/SNP). For the 22 patient samples from AALL0232, we performed Sanger sequencing separately for the 5 most commonly mutated exons of *JAK1* and *JAK2*.[8] The gene expression data for COG P9906 have been deposited at the National Center for Biotechnology Information Gene Expression Omnibus (accession no. {"type":"entrez-geo","attrs":{"text":"GSE11877","term_id":"11877","extlink":"1"}}GSE11877). The gene expression data without metadata for COG AALL0232 are deposited at the National Cancer Institute caArray site, project identifier EXP-578 (https://array.nci.nih.gov/caarray/project/EXP-578).