Supplementary Methods

Patient Samples

Validation of the clinical LSC17 assay was performed using cryopreserved mononuclear cell samples collected at diagnosis from adult AML patients who received 3+7 induction therapy at Princess Margaret (PM) Cancer Centre, Toronto, Ontario between 1999 and 2018. For some assays, freshly collected cells were used in addition to viably frozen cells from the PM Leukemia Tissue Bank. Written informed consent was obtained according to procedures approved by the Research Ethics Board of the University Health Network (UHN; REB# 01-0573-C). Total RNA extracted from cryopreserved peripheral blood (PB) or bone marrow (BM) specimens of AML patients obtained at diagnosis between 2012-2014 at Mayo Clinic (Rochester, MN) was used for interlaboratory technical validation studies (Mayo Clinic IRB# 14-001528).

RNA isolation

RNA was extracted from PB or BM specimens using either the Qiagen RNeasy Extraction Kit (for EDTA tubes) or PAXgene Blood RNA Extraction Kit (for PAXgene tubes). A subset of 52 samples that were received in EDTA tubes were also Ficoll-separated prior to RNeasy kit extraction. RNA was eluted in 60-80 µl of Tris-EDTA (TE) buffer and concentrations were measured using a Qubit fluorometer (Invitrogen, Thermo Fisher). The range of recovered RNA concentration for 52 Ficoll-separated EDTA samples was 7-379 ng /µl, the range for 52 EDTA samples without Ficoll media separation was 4-550 ng/µl, and the range for 27 PAXgene samples was 6-946 ng/µl. EDTA tubes containing PB samples were stored at 4˚C until processed. Except where noted, BM

samples were processed within 8 hours of collection. Ficoll-separated samples were stored in heparin tubes after separation.

Assay Protocols

Quality control and normalization was performed with nSolver analysis software v.4.0, using default settings for gene expression analysis. RCC files containing raw transcript counts from each cartridge and a reporter library file containing codeset probe annotations were imported into nSolver. The nSolver software was used to normalize the captured transcript counts to the geometric mean of the reference genes included in our assay and the TagSet's internal positive controls, and to assess quality metrics for the run. Quality control metrics assessed included: 1) number of fields of view imaged (280 required), 2) binding density, 3) positive control linearity and limit of detection, and 4) normalization quality. The control lane of each cartridge as well as blank lanes were processed in the same manner as the RNA lanes using the nSolver software without normalization to reference genes. In blank lanes in which no target RNA or DNA oligos were loaded, background signal was less than 5 counts per lane in validation studies, and as a result no background subtraction was applied for analysis. Binding Density QC values for samples tested with the redesigned clinical assay probes all fell into the default acceptable range of 0.05-2.25 complexes/micron^2 over 15 runs in the validation set.

Clinical reference cohort

The clinical reference cohort used to establish the median value for classification of LSC17 scores as high versus low has been previously described[1]. 306 adult AML samples were used to build the

clinical reference dataset. Expression values for the LSC17 signature genes were normalized to the expression levels of 12 housekeeping genes measured for each sample. Expression values were multiplied by the ratio of geometric means of housekeeping gene expression in all lanes of each cartridge and the geometric mean of expression values over all cartridges. The fully normalized gene expression counts were log2-transformed after incrementing by 1.

Statistical Analysis

Statistical analyses were performed in R 4.0.3 using the survival package version 3.2-11. The Spearman rank method of correlation was used throughout. Univariate survival analysis was performed using the Kaplan-Meier method. The Cox proportional hazards (CPH) model was used to estimate hazard ratios (HR) and Wald's test was used to evaluate significance in univariate and multivariate time-to-event analyses. Figure were generated using Prism software (version 9.01 for Windows).

To compare the prognostic value of LSC17 scores measured by the research assay versus the LDT, two CPH models were used, where the first included scores from the research assay as a covariate, while the second model also included scores derived from the LDT. The model fits were then evaluated using the likelihood ratio test, where $P < 0.05$ was considered to indicate that the second model and therefore the addition of the LDT scores significantly improved the model fit to patient outcome data.

Data availability

Standard chemistry LSC17 scores were computed by applying the previously published formula to the data deposited at GEO accession code GSE76004[1]. R scripts for data processing, normalization to clinical reference data, and LSC17 score calculation, including a detailed description and associated documentation are available upon request.

References

1. Ng SW, Mitchell A, Kennedy JA, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. 2016;540(7633):433-437.

Supplementary Table 1. Clinical characteristics of Mayo Clinic patients, according to high or low LSC17 scores measured by the LDT.

Characteristic	All patients	High LSC17 score	Low LSC17 score	P-value
	(*n*=45)	(*n*=25)	(*n*=20)	
Male Sex [*n* (%)]	29 (64.4)	19 (76)	10 (50)	0.11‖
Age at AML Diagnosis [years]				
median (range)	68 (0.25-92)	70 (23-92)	61.5 (0.25-80)	0.09†
***De novo* vs. Secondary AML [*n* (%)]**				
De novo	45 (100)	25 (100)	20 (100)	N/A
Secondary / t-AML	0 (0)	0 (0)	0 (0)	
PB WBC count at diagnosis (x10^9/L)				
median (range)	8.4 (0.7-194)	4.2 (0.7-194)	21.4 (0.9-92)	0.06*
Blast % at diagnosis (x10^9/L*)*				
Median BM Blasts (range)	66 (20-99)	59 (20-99)	69.5 (25-97)	0.49*
NCCN risk class at diagnosis [n (%)]	n=39	n=23	n=16	
Favorable	7 (17.9)	1 (4.34)	6 (37.5)	
Intermediate	11 (28.2)	8 (34.7)	3 (18.7)	0.03‖
Adverse	21 (53.8)	14 (60.8)	7 (43.7)	
ELN risk class at diagnosis [n (%)]	n=39	n=23	n=16	
Favorable	8 (20.5)	2 (8.69)	6 (37.5)	
Intermediate-1	8 (20.5)	2 (8.69)	6 (37.5)	0.004‖
Intermediate-2	8 (20.5)	7 (30.4)	1 (6.25)	
Adverse	15 (38.4)	12 (52.1)	3 (18.7)	
Karyotype [n (%)]	n=44	n=25	n=19	
Normal karyotype	16 (36.3)	3 (12)	13 (68.4)	<0.001‖
Abnormal karyotype	28 (63.6)	22 (88)	6 (31.5)	

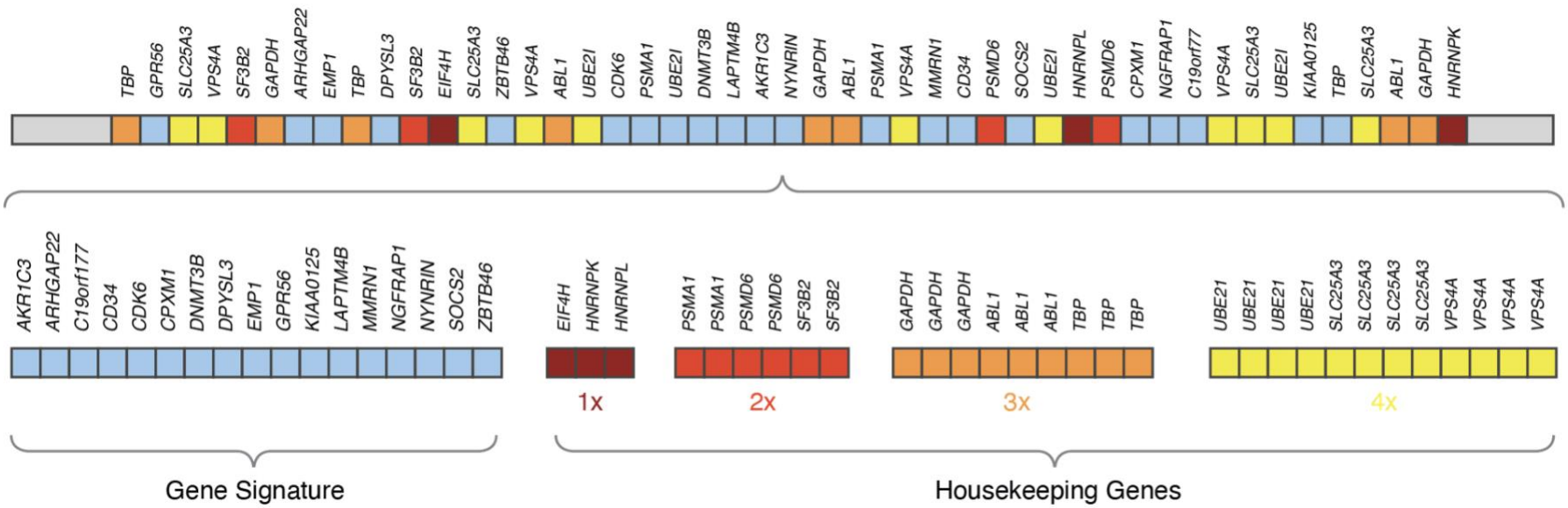
PB, peripheral blood; NCCN, National Comprehensive Cancer Network; ELN, European LeukemiaNet

* P-value calculated using the Wilcoxon rank-sum test

† P-value calculated using the Student's t-test

‖ P-value calculated using Fisher's exact test

Supplementary Figure 1



Schematic diagram illustrating the design of the LSC17 synthetic control. The LSC17 genes (light blue) are each represented at 1 copy. The housekeeping genes are represented at 1 (dark red), 2 (red), 3 (orange) or 4 (yellow) copies.