Supplementary Material:

Dissection of the major hematopoietic quantitative trait locus in chromosome 6q23.3 identifies miR-3662 as a novel player in hematopoiesis and acute myeloid leukemia

Sophia E. Maharry[1], Christopher J. Walker[1], Sandya Liyanarachchi[1], Sujay Mehta[1], Mitra Patel[1], Maryam A. Bainazar[1], Xiaomeng Huang[1], Malori A. Lankenau[1], Kevin W. Hoag[1], Parvathi Ranganathan[1], Ramiro Garzon[1], James S. Blachly[1], Denis C. Guttridge[1], Clara D. Bloomfield[1,#], Albert de la Chapelle[1,#], and Ann-Kathrin Eisfeld[1,#,§]

[1]The Ohio State University Comprehensive Cancer Center, Columbus, Ohio, 43210, USA

[#] all senior authors contributed to the work equally

§ Correspondence:

Ann-Kathrin Eisfeld, MD

The Ohio State University Comprehensive Cancer Center

460W. 12th Ave, BRT Room 850, Columbus, OH 43210, USA

Phone: +1/614-688-4171

Email: ann-kathrin.eisfeld@osumc.edu

Supplementary Table 1. Linkage disequilibrium patterns of the tag-SNPs used for genotyping the 6q23.3 locus. All ten SNPs have previously been identified in independent genome-wide association studies (GWASs) to be associated with various hematologic traits (see Table 1 for details). Because of the high linkage disequilibrium among the SNPs, genotyping of four SNPs was sufficient to determine all genotypes (Tag-SNPs highlighted in grey).

SNP	**Tag-SNP**	**R^2 value SNP/tag-SNP (cut-off: 0.9)**
rs7775698/ rs66650371	rs7775698	1
rs7776054	rs9402686	0.906
rs9399137	rs9402686	0.916
rs9373124	rs9402686	0.947
rs4895441	rs9402686	1
rs9376092	rs9402686	0.99
rs9402686	rs9402686	1
rs9494145	rs9483788	0.908
rs9483788	rs9483788	1
rs6569992	rs6569992	1

Supplementary Table 2. Cytogenetic and molecular information on AML patients used for functional studies (n=12). Endogenous miR-3662 expression levels were determined using qPCR in primary leukemic blasts isolated from all patients. All patients had low endogenous miR-3662 expression compared to mononuclear cells isolated from non-leukemic donors. In addition, the primary leukemic blasts from patients 1-12 were infected with lentiviral miR-3662 for stable expression, which was validated after 48h. Primary leukemic blasts from patients 1-7 successfully overexpressed miR-3662, and displayed levels similar to those observed in differentiated blood cells (positive control). Primary blasts of AML patients 1-4 had a viability ≥60% and were used for further *in vitro* studies (caspase-3/7 and/or TiterGlo assays). RNA isolated from primary leukemic blasts from AML patients 11 and 12 was used to perform the targeted RNA sequencing (Truseq).

AML patient	**rs66650371 genotype**	**FAB type**	**cytogenetics**	***FLT3*-ITD**	***NPM1***	**miR-3662 (endogenous)**	**miR-3662 (forced)**
1	wt/wt	M2	normal	present	wild-type	low	high
2	wt/wt	M1	normal	absent	mutated	low	high
3	wt/wt	n.a.	normal	present	mutated	low	high
4	wt/del	M4	normal	present	wild-type	low	high
5	wt/wt	M2	normal	absent	wild-type	low	high
6	wt/wt	n.a.	normal	absent	wild-type	low	high
7	wt/del	M2	normal	absent	wild-type	low	high
8	wt/wt	M1	normal	absent	mutated	low	low
9	wt/del	M4	normal	present	wild-type	low	low
10	wt/wt	M5	normal	absent	mutated	low	low
11	wt/wt	n.a.	normal	absent	mutated	low	high
12	wt/wt	M2	normal	absent	mutated	low	high

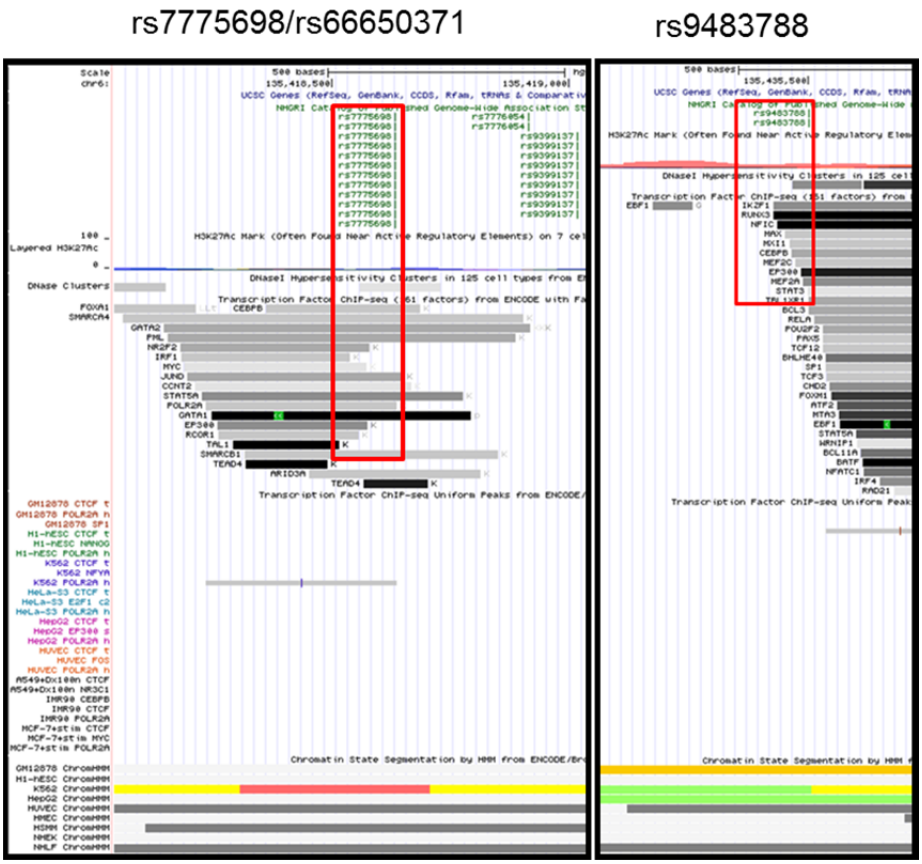
Supplementary Table 3. French-American-British (FAB) classification[35] and molecular information on AML patients used to determine miR-3662's abundance (n=8).

AML patient	rs66650371 genotype	FAB type	*FLT3*-ITD	*NPM1*
1	n.a.	M6	absent	n.a.
2	del/del	M6	n.a.	n.a.
3	wt/wt	M6	n.a.	n.a.
4	wt/del	M6	absent	wild-type
5	wt/wt	M5b	present	n.a.
6	n.a.	M5b	present	wild-type
7	wt/wt	M2	n.a.	n.a.
8	wt/del	AML with MDS changes	present	wild-type

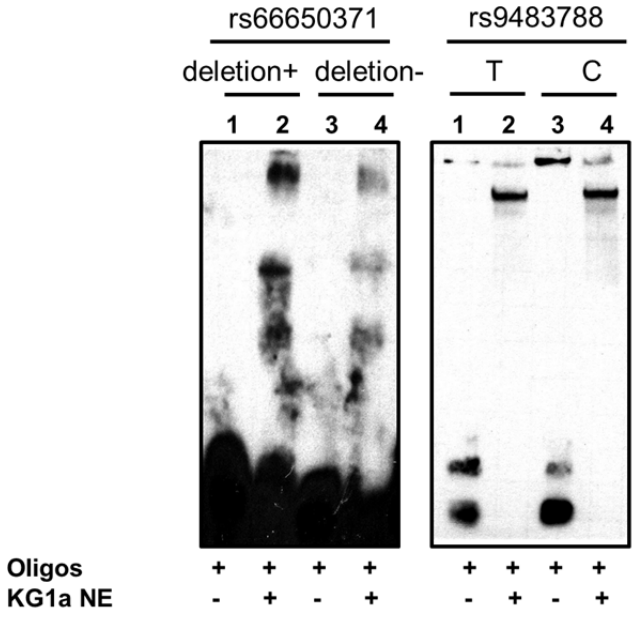
Supplementary Table 4. Canonical pathway analysis of the miR-3662-associated gene expression signature. Listed are the top ranking components of the Global Canonical Pathway Category (GCP) after pathway analysis of the miR-3662-associated gene expression signature (pathway names and involved molecules defined by Ingenuity). The *P*-value associated with each pathway in the Global Canonical Pathways (GCP) is a measure of the likelihood that the association between a set of focus genes in the analyzed experiment and a given pathway is due to random chance. Thus a low *P*-value identifies statistically significant over-representation of the focus genes in the pathway. It is automatically calculated by the Ingenuity program using a right-tailed Fisher's exact test.

Pathway name (Ingenuity)	***P*-value**	**molecules**
Cell Death and Survival	$7.9x10^{-36}$ - 5.49^{-07}	64
Cell Cycle	$2.8x10^{-27}$ - 4.69^{-07}	40
Cellular Growth and Proliferation	$1.02x10^{-23}$ - 5.69^{-07}	62
Cellular Development	$2.79x10^{-19}$ - 2.78^{-07}	55
Gene Expression	$1.82x10^{-17}$ - $5.35E^{-07}$	43

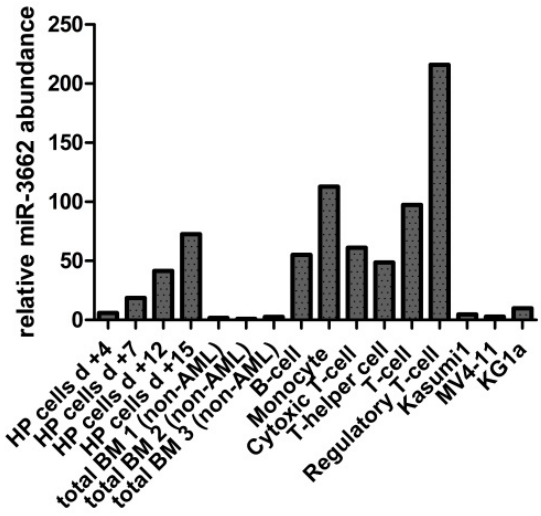
Supplementary Figure 1. Transcription factor binding according to the transcription factor chip data from ENCODE. Of the ten GWAS SNPs, only rs7775698/rs66650371 and rs9483788 are supposed to exhibited transcription factor binding (red boxes).



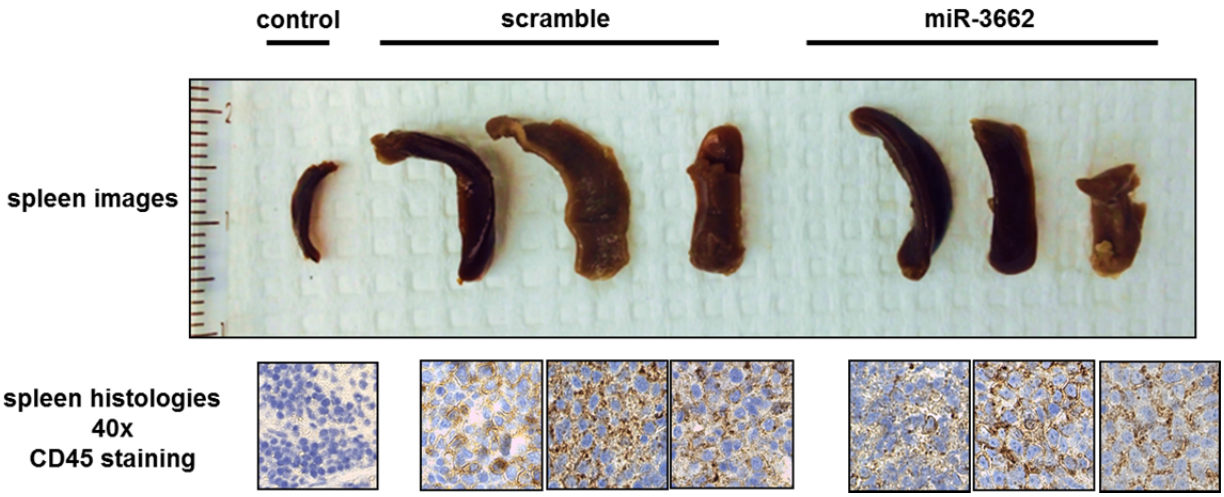
Supplementary Figure 2. Electrophoretic mobility shift assay comparing the binding affinity of the alleles of rs66650371 and rs9483788. While presence of the deletion increased transcription factor binding of rs66650371, no binding differences were seen for the different alleles of rs9483788.



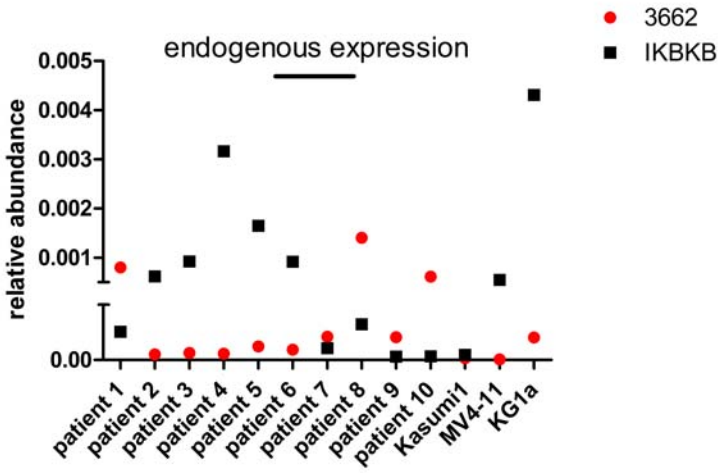
Supplementary Figure 3. Endogenous miR-3662 expression levels of hematopoietic progenitor (HP) cells during differentiation, total bone marrow aspirate of three non-leukemic donors (total BM 1-3), different populations of differentiated peripheral blood cells, and three AML cell lines. The abundance of total BM 2 was used as a reference.



Supplementary Figure 4. Top panel, Macroscopic pictures of the spleens of three mice of the scramble and miR-3662-infected groups (organs harvested post-mortem). All mice had a massive splenomegaly compared to the un-injected, sacrificed control mouse. **Bottom panel,** images of spleen histologies (40x enlargement). Slides were stained for CD45 to proof MV4-11 origin of the leukemia.



Supplementary Figure 5. Endogenous abundance of miR-3662 and IKBKB in patient samples and cell lines.



Supplementary Figure 6. Comparison of the relative miR-3662 abundance of AML patient blasts and AML cell lines before (black) and after (red) forced miR-3662 expression with the lentiviral expression construct.