OPEN ACCESS Freely available online

PLOS ONE

New Fusion Transcripts Identified in Normal Karyotype Acute Myeloid Leukemia

Hongxiu Wen[1☯], Yongjin Li[2☯], Sami N. Malek[3], Yeong C. Kim[1], Jia Xu[4¤], Peixian Chen[5], Fengxia Xiao[1], Xin Huang[6], Xianzheng Zhou[6], Zhenyu Xuan[2], Shiva Mankala[2], Guihua Hou[4], Janet D. Rowley[7], Michael Q. Zhang[2], San Ming Wang[1,8]*

1 Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, Nebraska, United States of America, **2** Department of Molecular and Cell Biology, The University of Texas at Dallas, Richardson, Texas, United States of America, **3** Department of Medicine, University of Michigan Medical Center, Ann Arbor, Michigan, United States of America, **4** Institute of Experimental Nuclear Medicine, Shandong University School of Medicine, Jinan, China, **5** Department of Medicine, University of Nebraska Medical Center, Omaha, Nebraska, United States of America, **6** School of Medicine, New York Medical College, New York, United States of America, **7** Department of Medicine, University of Chicago Medical Center, Chicago, Illinois, United States of America, **8** Eppley Cancer Institute, University of Nebraska Medical Center, Omaha, Nebraska, United States of America

Abstract

Genetic aberrations contribute to acute myeloid leukemia (AML). However, half of AML cases do not contain the well-known aberrations detectable mostly by cytogenetic analysis, and these cases are classified as normal karyotype AML. Different outcomes of normal karyotype AML suggest that this subgroup of AML could be genetically heterogeneous. But lack of genetic markers makes it difficult to further study this subgroup of AML. Using paired-end RNAseq method, we performed a transcriptome analysis in 45 AML cases including 29 normal karyotype AML, 8 abnormal karyotype AML and 8 AML without karyotype informaiton. Our study identified 134 fusion transcripts, all of which were formed between the partner genes adjacent in the same chromosome and distributed at different frequencies in the AML cases. Seven fusions are exclusively present in normal karyotype AML, and the rest fusions are shared between the normal karyotype AML and abnormal karyotype AML. *CIITA*, a master regulator of MHC class II gene expression and truncated in B-cell lymphoma and Hodgkin disease, is found to fuse with *DEXI* in 48% of normal karyotype AML cases. The fusion transcripts formed between adjacent genes highlight the possibility that certain such fusions could be involved in oncological process in AML, and provide a new source to identify genetic markers for normal karyotype AML.

Citation: Wen H, Li Y, Malek SN, Kim YC, Xu J, et al. (2012) New Fusion Transcripts Identified in Normal Karyotype Acute Myeloid Leukemia. PLoS ONE 7(12): e51203. doi:10.1371/journal.pone.0051203

Editor: William Tse, West Virginia University School of Medicine, United States of America

Received September 5, 2012; **Accepted** October 29, 2012; **Published** December 12, 2012

Funding: The study is supported by the American Recovery and Reinvestment Act of 2009, RC1CA145889 (S.M.W., J.D.R., M.Q.Z.).The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sanming.wang@unmc.edu

☯ These authors contributed equally to this work.

¤ Current address: The First Affiliated Hospital, Zhejiang University School of Medicine, Hongzhou, China

Introduction

Acute myeloid leukemia (AML) is a major type of leukemia, with estimated 13,780 new cases and 10,200 death in the United States in 2012 (American Cancer Society, 2012, http://www.cancer.org/Research/CancerFactsFigures/CancerFactsFigures/cancer-facts-figures-2012). Genetic aberrations including translocation, amplification, inversion, insertion, and deletion, are well-known to contribute to leukemia [1], and multiple recurrent genetic aberrations including t(9;11), inv(16), t(15;17) and t(8;21), have been identified in AML. These aberrations disrupt the normal structure of the affected genes and form fusion genes. The roles of these fusion genes in promoting myeloid leukemogenesis have been extensively studied. The disrupted genetic structures are widely used clinically as specific markers for diagnosis, prognosis, and treatment of AML [2].

Nearly half of AML cases do not contain the known genetic aberrations detectable by cytogenetic techniques [3,4,5,6]. These AML cases are classified into a subgroup named normal karyotype AML as they lack specific markers for further classification. The variation of treatment response and prognosis of normal karyotype AML suggests that this subgroup of AML can be heterogeneous with different genetic aberrations. Indeed, mutations in several genes including *CEBPA*, *FLT3*, *IDH1*, *MLL*, and *NPM1*, and alternated expression of AF1q have been identified in the normal karyotype AML, and the potential value of these changes for clinical applications are under investigation [7,8,9]. Using whole genome sequencing approach, two normal karyotype AML genomes have been sequenced [10,11]. The studies identified heterozygous, non-synonymous somatic mutations in 10 genes in the first case, and in 12 genes in the second case. Except for the known mutations in *FLT3*, *IDH1*, *and NPM1*, however, the rest novel mutations are individual case-specific but not shared within a cohort of 187 additional AML cases, of which 76 are normal karyotype AML [12]. Therefore, identification of recurrent genetic mutations in nomal karyotype AML remains a challenging task.

Genetic aberrations could be examined at the levels of genomic DNA, transcription or translation. As the genomic DNA structure

in normal karyotype AML is intact as revealed by cytogenetic and genome sequencing data, an alternative approach to search for potential aberrations in normal karyotype AML could be at the level of transcription. To investigate this possibility, we used the next generation sequencing-based paired-end RNAseq method [13] to detect potential fusion transcripts in a group of normal karyotype AML cases. Similar approaches have been applied recently in multiple types of cancer with the idenification of many new fusion transcripts [14,15,16,17,18,19]. Our comprehensive sequence data collection, analysis and experimental validation result in the identification of many novel fusion transcripts formed exclusively between the genes adjacent in the same chromosome in normal karyotype AML. Here we report our analyses and observations.

Results

RNAseq sequence collection and analysis

Twenty-nine normal karyotype primary AML cases were used for the study. Each case was diagnosed using the standard criteria and cytogenetic analysis as normal karyotype AML in University of Michigan Medical Center. In addition, eight abnormal karyotype AML cases and eight AML cases without karyotype data were also included for the study. PolyA+ mRNA was extracted from each sample, and subjected to paired-end RNAseq sequencing (2×50). A total of 1,374,219,760 paired-end reads (137,421,976,000 bases), or on average 30,538,217 reads (3,053,821,688 bases) per sample, were collected in the study (Figure 1).

Identification of fusion transcripts

Using the FusionSeq and deFuse programs [20,21], we analyzed the paired-end RNAseq data to identify fusion transcripts. Three well-known fusions in AML were identified in four abnormal karyotype AML cases, including *RUNX1-RUNX1T1* fusion in two cases [46,XXt(8;21)(q22;q22); 46,XYt(8;21)(q22;q22)], *MLL-MLLT1* fusion in one case [48,XX,+8,+8,i(8)(q10)/49,idem,+i(8)(q10)] and *MLL-MLLT3* fusion in one case [46,XY,t(9;11;11)(p22;q23;p11.2)]. The detection of these three known fusions shows the sensitivity of the paired-end RNAseq and the mapping programs used for the analysis. Under the threshold that a fusion must be detected by at least two pairs of fusion sequences and in at least two AML cases, a total of 134 fusions were identified in the 45 AML samples used in the study (Figure 2A). In the 29 normal karyotype AML, 114 fusions were identified, of which 88 are novel fusions detected in all AML, 7 fusions [FAM65A-CTCF (6 cases), KIAA1267-ARL17 (5 cases), *GGCT-BC041636* (3 cases), *R3HDM2-INHBC* (3 cases), LY6G5C-ABHD16A (2 cases), *CRNKL1-NAA20* (2 cases), and *HSPA14-SUV39H2* (2 cases)] were detected only in normal karyotype AML. Twenty-six fusions were only detected in the abnormal karyotype AML and/or the AML without karyotype information (Figure 2A). Each fusion was distributed at different frequencies among the AML cases, with the highest one of 57 fusions detected in a normal karyotype AML case (#25, Table S1).

Validation of fusions

Based on the availability of patient RNA samples, the suitable size and base composition of the fusion sequences for primer design, and the frequency of fusion distribution in the AML cases, 66 fusions were selected for experimental validation by using antisense primers for reverse transcription, PCR amplification and Sanger sequencing. A total of 37 (56%, the relative lower rate is related with the detection of sense transcripts only by using antisense primer for reverse transcription. See below for explanation) were confirmed as true fusion transcripts in the 45 leukemia samples, of which 30 fusions were in the 29 normal karyotype AML cases (Figure 2B, Figure S1, Table S2). Of the 30 fusions, 24 maintain in-frame coding. For example, *NFATC3-PLA2G15* fusion is formed between the upstream gene *NFATC3* and the downstream gene *PLA2G15* in 3′-5′ tail to head orientation. This fusion is in-frame and the V resudue at the fusion junction is shared between the fusion partners (Figure 3). Six fusions truncated their coding structure [*CLN5-FBXL3, CLR-CLEC2D, ATP5I-MFSD7, KIAA1267-ARL17, PLD4-AHNAK2*, and *ZNF789-ZNF394*, Figure 2B].

Genomic origin of fusion transcripts

Mapping the fusions to the human genomic reference sequences (hg18) shows that all the detected novel fusions are formed between the genes adjacent each other in the same chromosome. Complicated orientations of upstream and downstream partner genes are present, including 3′-5′ tail to head, 5′-3′ head to tail, 3′-3′ tail to tail, and 5′-5′ head to head (Figure S2). Considering that there are two partner genes involved in each fusion, a question remains whether only one or both partner genes drove the fusion transcription. In the former case, only one strand should be used as the transcription template resulting in either sense or antisense transcripts; in the latter case, both strands should be used as transcription templates resulting in both sense and antisense transcripts. RNAseq sequences collected by the standard RNAseq method can identify the fusion junction but can not discriminate the original strand used for fusion transcription, and regular RT-PCR using only antisense primer for reverse transcription as used in our validation experiment above can only validate the transcripts from one DNA strand. To address this issue, we performed strand-specific RT-PCR using sense primer and antisense primer separately for reverse transcription. We used RNA samples from 8 myeloid leukemia-derived cell lines (KG-1, THP-1, Molm-13, Molm-14, MV4-11, HL-60, Kasumi-1, and Kasumi-6) to overcome the shortage of patient RNA samples for the analysis, and we selected 12 fusions for the test. The results confirmed that each fusion was expressed in at least one of the cell lines, and determined that 86% of fusion transcripts (68 out of 79 positive reactions in the total of 96 reactions) were transcribed from both sense and antisense strands (Figure 4A). For example, sense and antisense transcripts for one of the two *CIITA-DEXI* fusions were expressed in 7 out of 8 cell lines (Figure 4B). Testing the *CIITA-DEXI* fusion in 10 AML RNA samples show similar pattern of sense and antisense transcription that the sense fusion transcripts were expressed in all and the antisense fusion transcripts were expressed in 7 of the 10 AML samples (Figure 4B).

Possible causes of fusion formation

Both genomic structural changes between the partner genes and post-transcriptional process can cause the formation of fusion transcripts. We used long-range PCR to amplify the mapped genomic region between the partner genes to know if there is a genomic structural change as reflected by the size difference of the amplified region comparing to the reference genome sequences. Of the 11 successfully amplified genomic DNA fragments, 6 have the sizes similar to and 5 have the sizes shorter than their corresponding regions in the reference genome sequences (Figure 5). For example, the size between *IL17RB-ACTR8* fusion is 4,044 bp in the reference genome sequences whereas the size of the amplified genomic fragment in AML is about 4 kb, suggesting that this fusion was generated by post-transcriptional process rather than genomic structural change; in the case of *CLR-*

A. AML case and RNAseq production

No.	Karyotype	Sample type	Blast (%)	Reads (2x50)	Bases
1	46, XY	Bone marrow	90	24,614,703	2,461,470,300
2	46, XY	Bone marrow	88	28,843,603	2,884,360,300
3	46, XY	Bone marrow	43	31,267,542	3,126,754,200
4	46, XX	Bone marrow	38	32,518,678	3,251,867,800
5	46, XY	Bone marrow	90	27,914,265	2,791,426,500
6	46, XX	Bone marrow	42	15,830,586	1,583,058,600
7	46, XX	Peripheral blood	71	30,770,492	3,077,049,200
8	46, XX	Peripheral blood	83	28,597,393	2,859,739,300
9	46, XX	Peripheral blood	66	29,105,962	2,910,596,200
10	46, XY	Peripheral blood	93	32,766,978	3,276,697,800
11	46, XY	Peripheral blood	84	28,074,606	2,807,460,600
12	46, XY	Peripheral blood	65	27,129,917	2,712,991,700
13	46, XX	Peripheral blood	66	32,673,030	3,267,303,000
14	46, XY	Peripheral blood	50	32,629,136	3,262,913,600
15	46, XY	Peripheral blood	55	27,996,366	2,799,636,600
16	46, XX	Peripheral blood	91	30,014,538	3,001,453,800
17	46, XX	Peripheral blood	35	31,305,698	3,130,569,800
18	46, XY	Peripheral blood	47	27,985,862	2,798,586,200
19	46, XY	Peripheral blood	41	33,253,655	3,325,365,500
20	46, XX	Bone marrow	80	21,676,424	2,167,642,400
21	46, XX	Bone marrow	92	34,693,529	3,469,352,900
22	46, XY	Bone marrow	82	20,786,654	2,078,665,400
23	46, XX	Bone marrow	90	26,504,532	2,650,453,200
24	46, XX	Bone marrow	65	44,998,831	4,499,883,100
25	46, XY	Bone marrow	<10	44,645,304	4,464,530,400
26	46, XY	Peripheral blood	92	37,941,973	3,794,197,300
27	46, XY	Peripheral blood	98	35,331,090	3,533,109,000
28	46, XX	Peripheral blood	95	24,535,864	2,453,586,400
29	46, XX	Peripheral blood	80	22,140,260	2,214,026,000
30	46,XY,t(9;11;11)(p22;q23;p11.2)	Bone marrow	94	34,785,818	3,478,581,800
31	46,XX,i(17)(q10)/46,XX	Bone marrow	75	35,756,006	3,575,600,600
32	46,XY,t(8;21)(q22;q22)	Bone marrow	63	16,978,603	1,697,860,300
33	46,XX,t(8;21)(q22;q22)	Bone marrow	58	46,317,335	4,631,733,500
34	46,XY,der(14)t(11;14)(q13;q32)	Bone marrow	92	35,780,585	3,578,058,500
35	48,XY,+8,+15	Bone marrow	NA	46,815,659	4,681,565,900
36	48,XY,+21,+21	Bone marrow	97	29,058,205	2,905,820,500
37	48,XX,+8,+8,i(8)(q10)	Peripheral blood	90	44,234,187	4,423,418,700
38	No available	Bone marrow	33	32,368,840	3,236,884,000
39	No available	Bone marrow	55	22,547,358	2,254,735,800
40	No available	Bone marrow	75	27,173,669	2,717,366,900
41	No available	Bone marrow	NA	44,459,470	4,445,947,000
42	No available	Bone marrow	95	46,567,887	4,656,788,700
43	No available	Peripheral blood	28	8,036,353	803,635,300
44	No available	Peripheral blood	60	20,728,171	2,072,817,100
45	No available	Peripheral blood	80	16,064,143	1,606,414,300

B. Summary of squence data and fusion detection

AML types	Number of cases	Reads	Bases
Normal karyotype	29	866,547,471	86,654,747,100
Abnormal karyotype	8	289,726,398	28,972,639,800
No karyotype	8	217,945,891	21,794,589,100
Total	45	1,374,219,760	137,421,976,000

Figure 1. AML sample list and RNAseq data collection. A. A total of 45 AML samples were used for the analysis, including 29 normal karyotype AML, 8 abnormal karyotype AML and 8 AML without karyotype information. B. RNAseq data collected from the 45 AML cases.
doi:10.1371/journal.pone.0051203.g001

CLEC2D fusion, the size in the reference genome sequences is 31 kb but the size of the amplified genomic fragment has only 8 kb. This fusion transcript is likely caused by genomic DNA deletion between the two partner genes. The results indicate that both genomic structural change and post-transcriptional processing can contribute to the fomation of the detected fusion transcripts.

Specificity of fusions for AML

We used several filters to determine the specificity of fusion transcripts detected in AML. A recent comprehensive study identified 800 "Conjoined genes", which are formed between the genes adjacent in the same chromosome in the human genome [22]. As the data set is largely derived from normal human genome data, we used these as a filter to eliminate the fusions also present in normal human cells. The comparison identified and eliminated 16 such fusions from the AML fusion list (*ADSL-*

A. Total fusions detected by RNAseq

Fusions	Normal karyotype AML	Abnormal karyotype AML	AML without karyotype	Total
In all 3 types of AML	88	78	69	88
Only in normal karyotype AML	7	-	-	7
Only in abnormal karyotypes and unknown karyotype	-	17	9	
Shared with prostate cancer	5	5	4	
Also in the Conjoined gene database	14	13	12	
Total	114	113	94	134*

*The fusions detected in different types of AML counted only once

B. Fusions validated in the 29 normal karyotype AML

Fusion	Case number*																								Case/Fusion**	Coding change
	25	24	10	26	8	11	4	2	13	5	15	19	29	20	6	23	14	18	17	7	22	28	1	12		
CIITA-DEXI (1)																									10	in-frame
CIITA-DEXI (2)																									8	in-frame
CIITA-DEXI (3)																									3	in-frame
FAM65B-BC070382																									9	in-frame
BC067907-MSN																									6	in-frame
G3BP2-AK311578																									5	in-frame
BC042152-AK024119																									5	in-frame
IL17RB-ACTR8																									4	in-frame
HERC3-FAM13AOS																									3	in-frame
HSP90B1-DKFZp547P055																									3	in-frame
AGK-KIAA1147																									2	in-frame
C2orf56-PRKD3																									2	in-frame
CLN5-FBXL3																									2	truncated
CLR-CLEC2D																									2	truncated
CNNM3-ANKRD23																									2	in-frame
HARS2-ZMAT2																									2	in-frame
LY6G5C-ABHD16A																									2	in-frame
PPCS-LOC728621																									2	in-frame
ZNRF2-DKFZP586I1420																									2	in-frame
ARL13A-TRMT2B																									1	in-frame
ATP5I-MFSD7																									1	truncated
BC016361-EREG																									1	in-frame
DHX8-ETV4																									1	in-frame
GGCT-BC041636																									1	in-frame
KIAA1267-ARL17																									1	truncated
NFATC3-PLA2G15																									1	in-frame
NUMB-AK055876																									1	in-frame
PAQR6-SMG5																									1	in-frame
PLD4-AHNAK2																									1	truncated
ZNF789-ZNF394																									1	truncated
Fusion/Case	13	12	9	7	6	5	4	4	3	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	85	

* Case order is sorted by the frequency of fusions in each case

**These detected in single case were also detected in abnormal karyptype or non karyotype AML cases

Figure 2. Fusion transcript information. A. Fusion transcripts identified in different types of AML. B. Validated fusion transcripts identified in normal karyotype AML.
doi:10.1371/journal.pone.0051203.g002

*SGSM3, ARHGAP27-LOC201175, C11orf79-C11orf66, CS-CNPY2, ETFB-CLDND2, FLOT2-DHRS13, GALT-IL11RA, GPN3-ARPC3, MGC72080-ASNS, RAB24-MXD3, RRM2-C2orf48, SCO2-TYMP, STX16-NPEPL1, TNFAIP8L2-SCNM1, VAMP8-VAMP5, and VPS72-TMOD*4, Table S1). Fusion transcripts between adjacent genes have also been identified in multiple types of solid cancer including prostate cancer [13], ovarian cancer [15], breast cancer [16], lung cancer and colorectal cancer [18]. We compared the fusions identified in normal karyotype AML with the fusions identified by these studies. Five fusions (*XPA-NCBP1*, *HARS2-ZMAT2*, *HSP90B1-DKFZp547P055*, *IL17RB-ACTR8*, *ANKRD23-ANKRD39*) are identified to share with the fusions identified in prostate cancer but no fusions are shared with breast cancer, colorectal cancer, ovarian cancer, and lung cancer. Further classification of the remaining 113 fusions shows that 7 fusions (*FAM65A-CTCF, KIAA1267-ARL17, BC041636-GGCT, R3HDM2-INHBC, ABHD16A-LY6G5C, CRNKL1-NAA20, HSPA14-SUV39H2*) are present only in normal karyotype AML, 18 are present only in abnormal karyotype AML or AML without karyotype information, and 88 fusioins are shared between normal, abnormal karyotype, and no karyotype AML groups (Table S1). The fusions present only in each group have lower frequencies [2–6 (7–21%) for the 7 fusions only in the 29 normal karyotype AML, 1 to 3 (6–19%) for the 18 fusions only in the 16 abnormal karyotype AML and AML without karyotype information], but the frequencies for the fusions shared in the three AML groups are from low to high [2–30 (4–67%) for the 88 fusions in all 45 AML samples]. The frequencies of the fusion distribution in these AML groups imply that the majority of the fusions identified in AML are not related with the karyotype status of the AML cases. Although the common fusions may not be much value as markers for abnormal karyotype AML as they already have well-defined genetic markers, these fusions can be potentially very useful markers for the normal karyotype AML.

CIITA-DEXI fusions

Three *CIITA-DEXI* fusion isoforms were detected in 14 of the 29 normal karyotype AML cases [Figure 2B, Figure S2]. They were formed between *CIITA* and *DEXI* in 3′–3′ tail to tail orientation in chromosome 16 [Figure 6B]. The three fusions contain the same *CIITA* part after the stop codon and fused to different locations of the same intron of *DEXI*. CIITA is a master regulator of *MHCII* gene expression, which is important for antigen-presenting cells to maintain their proper immune function. CIITA functions through interaction with multiple proteins including RFX complex (RFXANK, RFX5, and RFXAP) bound at *MHCII* promoters and other transcriptional factors (Figure 6A). Mutation of *CIITA* causes Bare Lymphocyte Syndrome, a disease unable to fight against infection [23]. DEXI is a calcium binding protein with two exons of which only one is for coding. The

A. NFATC3

MFIGTADDRYLRPHAFYQVHRITGKTVATASQEIIIASTKVLEIPLLPENNMSASIDCAGILKL
RNSDIELRKGETDIGRKNTRVRLVFRVHIPQPSGKVLSLQIASIPVECSQRSAQELPHIEKYSI
NSCSVNGGHEMVVTGSNFLPESKIIFLEKGQDGRPQWEVEGKIIREKCQGAHIVLEVPPYHNPA
VTAAVQVHFYLCNGKRKKSQSQRFTYTPVLMKQEHREEIDLSSVPSLPVPHPAQTQRPSSDSGC
SHDSVLSGQRSLICSIPQTYASMVTSSHLPQLQCRDESVSKEQHMIPSPIVHQPFQVTPTPPVG
SSYQPMQTNVVYNGPTCLPINAASSQEFDSVLFQQDATLSGLVNLGCQPLSSIPFHSSNSGSTG
HLLAHTPHSVHTLPHLQSMGYHCSNTGQRSLSSPVADQITGQPSSQLQPITYGPSHSGSATTAS
PAASHPLASSPLSGPPSPQLQPMPYQSPSSGTASSPSPATRMHSGQHSTQAQSTGQGGLSAPSS
LICHSLCDPASFPPDGATVSIKPEPEDREPNFATIGLQDITLDD**V**NEIIGRDMSQISVSQGAGV
SRQAPLPSPESLDLGRSDGL

B. PLA2G15

MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPPVVL**V**PGDLGNQLEAKLDKPTVVHYL
CSKKTESYFTIWLNLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDVRVPGFGKTFSLEFLD
PSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQLYGGPV
VLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRIPVIGPLKI
REQQRSAVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLMRQDTEGLVEA
TMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQAWQSRQEHQVLL
QELPGSEHIEMLANATTLAYLKRVLLGP-

C. NFATC3-PLA2G15 fusion

MFIGTADDRYLRPHAFYQVHRITGKTVATASQEIIIASTKVLEIPLLPENNMSASIDCAGILKL
RNSDIELRKGETDIGRKNTRVRLVFRVHIPQPSGKVLSLQIASIPVECSQRSAQELPHIEKYSI
NSCSVNGGHEMVVTGSNFLPESKIIFLEKGQDGRPQWEVEGKIIREKCQGAHIVLEVPPYHNPA
VTAAVQVHFYLCNGKRKKSQSQRFTYTPVLMKQEHREEIDLSSVPSLPVPHPAQTQRPSSDSGC
SHDSVLSGQRSLICSIPQTYASMVTSSHLPQLQCRDESVSKEQHMIPSPIVHQPFQVTPTPPVG
SSYQPMQTNVVYNGPTCLPINAASSQEFDSVLFQQDATLSGLVNLGCQPLSSIPFHSSNSGSTG
HLLAHTPHSVHTLPHLQSMGYHCSNTGQRSLSSPVADQITGQPSSQLQPITYGPSHSGSATTAS
PAASHPLASSPLSGPPSPQLQPMPYQSPSSGTASSPSPATRMHSGQHSTQAQSTGQGGLSAPSS
LICHSLCDPASFPPDGATVSIKPEPEDREPNFATIGLQDITLDD**V**PGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDVRVPGFGKTFSLEF
LDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQLYGG
PVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRIPVIGPL
KIREQQRSAVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLMRQDTEGLV
EATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQAWQSRQEHQV
LLQELPGSEHIEMLANATTLAYLKRVLLGP-

Figure 3. NFATC3-PLA2G15 fusion. The fusion is formed between upstream gene *NFATC3* and downstream gene *PLA2G15* in 3′-5′ tail to head orientation. In this fusion, amino acid V (GTC) is shared at the fusion point (G from *NFATC3* and TC from *PLA2G15*). A. Wild-type NFATC3 protein sequence; B. Wild-type PLA2G15 protein sequence; C. NFATC3-PLA2G15 fusion protein sequences. The bold V residue marks the fusion junction.
doi:10.1371/journal.pone.0051203.g003

CIITA-DEXI fusion maintains the entire coding exons of *CIITA* including its stop codon but removes the entire 3′ UTR of *CIITA* (Figure 6B). This is very different from the *CIITA*-fusions identified in lymphoma, in which the *CIITA* parts are truncated in the coding region [24]. Functional means of *CIITA-DEXI* fusion in myeloleukemogenesis remains to be elucidated.

Discussion

Fusion genes are frequently present in hematopoietic cancer, often caused by structural changes of translocation, inversion, deletion and insertion involving different chromosomes or distal regions in the same chromosome. While normal karyotype AML lacks the well-known genetic aberrations, our study shows that fusion transcripts, predominantly derived between the genes adjacent in the same chromosome, are widely present in this subgroup of AML. It is also interesting to note that the majority of the fusion transcripts detected in the normal karyotype AML are also present in the abnormal karyotype AML cases. This fact indicates that the expression of the fusion transcripts are independent events regardless the karyotype status of the AML cells.

Unlike the fusions between the adjacent genes in normal human genome, in which the partner genes are unilaterally orientated along the same chromosome and expressed most likely by the ''read-through'' mechanism [22], the partner genes of the fusions identified in AML consist of very heterogeneous patterns of partner gene orientation. The presence of antisense transcripts for the majority of the genomic fusion loci further complicates the determination of the original strands used for the fusion transcription.

To identify the fusion transcripts present in AML, it is important to exclude the fusion transcripts present in the normal cells of the same individual and the normal cells mixed with the AML samples (the blast rate of the AML samples used in this study is between 95% to <10%, see Figure 1). Obtaining normal cells from the same leukemic individual, e.g., skin tissue, may not be a perfect choice as evidence shows that the normal skin tissue from leukemic patients can be highly contaminated with leukemic cells [10]. To overcome this technical difficulty, we used the ''conjoined genes'' as the filter to exclude the AML fusions also present in

A. Summary of the results

Fusion gene	Myeloid cell lines (sense transcripts /antisense transcripts)							
	KG-1	THP-1	Molm-13	Molm-14	MV4-11	HL-60	Kasumi-1	Kasumi-6
CIITA-DEXI	+/-	+/+	+/+	+/+	+/-	-/-	-/-	+/+
*CIITA-DEXI**	+/+	+/+	+/+	+/+	+/+	+/+	+/-	+/+
*FAM65B-BC070382**	+/+	-/-	-/-	-/-	-/-	-/-	+/+	+/+
TRMT2B-ARL13A	+/-	+/+	+/+	+/+	+/+	+/+	+/+	+/+
G3BP2-AK311578	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+
C2orf56-PRKD3	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+
BC042152-AK024119	+/+	+/+	+/+	+/+	+/+	+/+	+/-	-/-
BC067907-MSN	+/+	+/-	+/+	+/+	+/+	+/+	+/+	+/-
KIAA1267-ARL17	-/-	-/-	-/-	-/-	-/-	+/+	-/-	-/-
HARS2-ZMAT2	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+
HSP90B1-DKFZp547P05	+/-	+/-	+/+	+/+	+/+	+/+	+/+	+/+
IL17RB-ACTR8	-/+	-/-	-/+	+/+	+/+	+/+	-/-	+/+

*These two fusions were shown in B,

B. Examples of *CIITA* sense and antisense transcripts in myeloid cell lines and AML cases

cDNA was generated by either sense or antisense primers. Positive control: beta-actin, negative control: reaction with a degraded RNA to rule out genomic DNA contamination

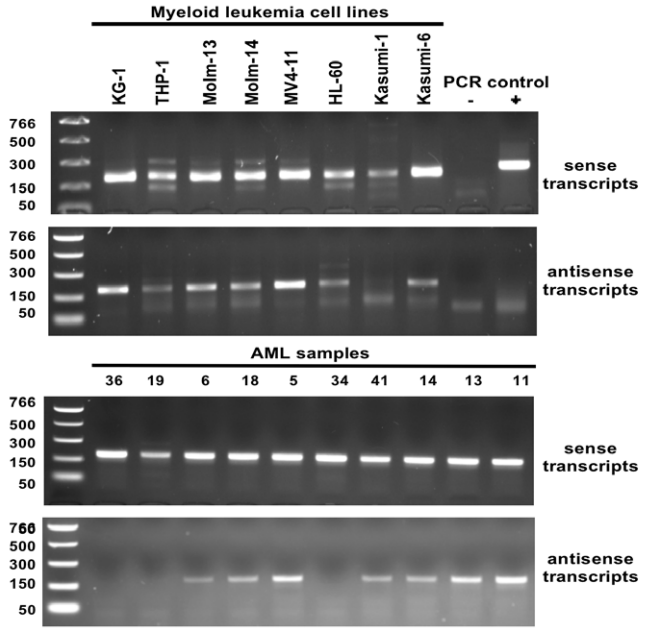


Figure 4. Validation of sense and antisense fusion transcripts by strand-specific RT-PCR. A. Summary for RNA samples from 8 myeloid cell lines. B. *CIITA-DEXI* sense and antisense fusion transcripts detected in myeloid cell lines and AML samples. +: positive control with beta-actin; -: netative control without RNA templates.
doi:10.1371/journal.pone.0051203.g004

normal human cells. Although this filter is not perfect solution, it is helpful to narrow down the fusion candidates associated with AML.

There are seven fusions detected only in the normal karyotype AML (*FAM65A-CTCF, KIAA1267-ARL17, BC041636-GGCT, R3HDM2-INHBC, ABHD16A-LY6G5C, CRNKL1-NAA20, HSPA14-SUV39H2*). The involved gene *CTCF, INHBC and SUV39H2* are particularly interesting. *CTCF* encodes a zinc-finger transcriptional regulator protein with 11 zinc fingers. It regulates transcription both positively and negatively. Mutations in *CTCF* have been observed in breast, prostate, and Wilms' cancers [25]. *INHBC* encodes the beta C chain of inhibin, a member of the TGF-beta superfamily involving hormone secretion [26]. The fusions of *FAM65A-CTCF* and *R3HDM2-INHBC* both removed the first exon and translational starting site of CTCF and *INHBC*, likely caused loss of the function of *CTCF and INHBC*; *SUV39H2* encodes a histone-lysine N-methyltransferase [27]. The *HSPA14-SUV39H2* fusion removed the last 3 exons of *HSPA14* and connected the remaining *HSPA14* part to the first exon of *SUV39H2*.

Five fusions (*XPA-NCBP1*, *HARS2-ZMAT2*, *HSP90B1-DKFZp547P055*, *IL17RB-ACTR8*, *ANKRD23-ANKRD39)* are shared between AML and prostate cancer. This raises an interesting question whether certain fusions could be common genetic aberrations contributing to different types of cancer. The *HARS2-ZMAT2* fusion was formed in-frame between the 3′ UTR of *HARS2* and the 3rd exon of *ZMAT2* in 3′-5′ tail to head, sense orientation. *HARS2* (histidyl-tRNA synthetase 2) is a nuclear-encoded mitochondrial protein involving in the synthesis of histidyl-transfer RNA [28]. Mutations in *HARS2* cause Perrault

A. Summary

Fusion gene	Chromosome	Size in hg18	Size amplified	Result
ZNRF2-DKFZP586I1420	7	5,052 bp	5 kb	Similar
IL17RB-ACTR8	3	4,044 bp	4 kb	Similar
CLN5-FBXL3	13	13,612 bp	8 kb	Shorter
KIAA1147-AGK	7	15,718 bp	10 kb	Shorter
CLR-CLEC2D	12	31 kb	8 kb	Shorter
HARS2-ZMAT2	5	3,393 bp	3.3 kb	Similar
HERC3-FAM13AOS	4	3,672 bp	3.6 kb	Similar
C2orf56-PRKD3	2	4,997 bp	3 kb	Shorter
HSP90B1-DKFZp547P055	12	6,599 bp	6 kb	Similar
PAQR6-SMG5	1	3,633 bp	3.6 kb	Similar
GGCT-BC041636	7	24 kb	10 kb	Shorter

B. Sizes of the amplified genomic DNA fragments

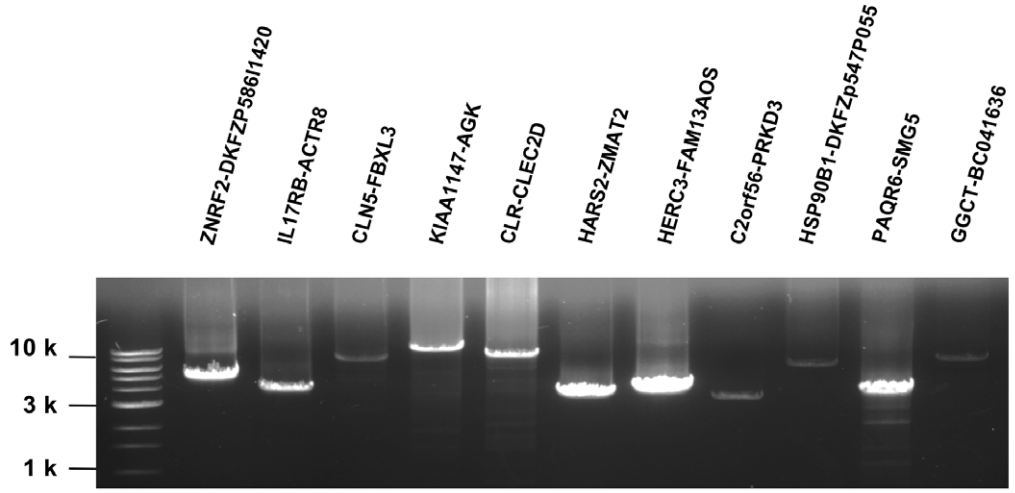
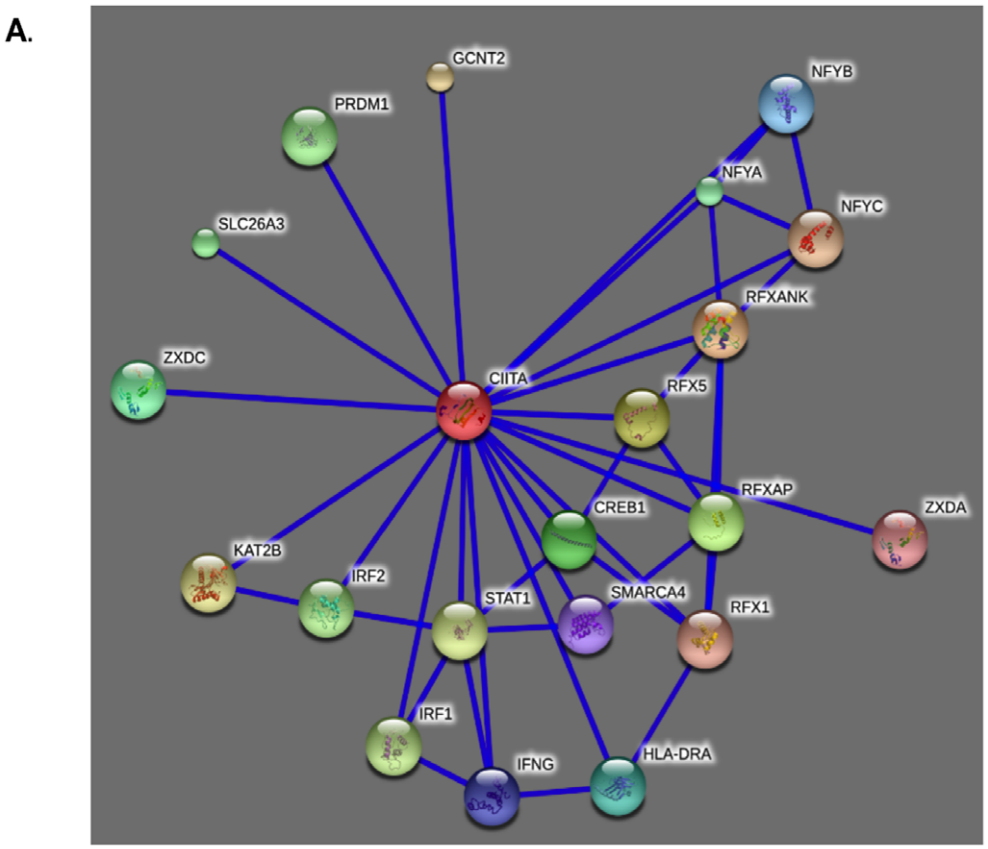


Figure 5. Long-range PCR results. A. Summary of the results from 11 fusion candidates. B. Size distribution of the amplified genomic DNA fragments.
doi:10.1371/journal.pone.0051203.g005

syndrome with ovarian dysgenesis and hearing loss [29]. *ZMAT2* is a matrin-type 2 zinc finger protein with unknown function. The *HSP90B1-DKFZp547P055* fusion was formed between the 17th of 18 exons of *HSP90B1* before its stop codon and the 2nd exon of a functionally unknown gene *DKFZp547P055*, fused in a 3′-3′ tail to tail orientation between the two genes. HSP90B1 is a member of HSP90 proteins located in the endoplasmic reticulum. It is involved in folding Toll-like receptors involving in regulating innate and adaptive immunity, maintaining hematopoietic stem cell development, and potentially to be used as a vaccine for cancer treatment [30,31,32,33]. The *HSP90B1-DKFZp547P055* fusion disrupted the normal structure of *HSP90B1*, and possibly led to loss of function of HSP90B1. *IL17RB-ACTR8* fusion is formed in a 3′-3′ tail to tail opposite orientation before the stop codon of *IL17RB* and the 3′ UTR of *ACTR8*. IL17RB is involved in the activation of the NF-kappaB pathway [34]. This in-frame fusion maintains the normal codon of *IL17RB*, suggesting that the fusion could be translated into the two original proteins. Prostate cells are derived from endoderm and hematopoietic cells are derived from mesoderm. Therefore, the shared fusions between prostate cancer and AML are unlikely attributed to the same developmental origin. It will be interesting to further study the biological basis of the shared fusion transcripts between the two very different cancer types. *MHCII* is generally expressed in T cells, B cells and dendritic cells involving antigen presentation [35]. While myeloid cells are not directly involved in antigen presentation, it has been observed that AML cells express many B-cell specific genes including *CIITA* [36]. It will be interesting to know if the altered *CIITA* in normal karyotype AML could contribute to its escape of the host's immune-surveillance.

Sense and antisense transcripts are commonly present for the majority of human genes [37]. Our data show that sense and antisense transcripts are also present for most of the fusions detected. The orientations of the fusion partners at genomic DNA level are very different. It will be expected that only the 3′ to 3′, tail to tail fused partners can drive the sense and antisene transcription. However, of the 12 fusions used for sense antisense detection, 6 are not in the 3′ to 3′ orientation but sense antisense transcripts were detected in each case. How were the sense and antisense transcripts expressed from these fusion loci remain to be understood.

While the biological/pathological function of many fusion transcripts identified in this study remains to be elucidated, a recent study shows that fusion transcripts formed by post-transcriptional process can have functional relevance [38]. Importantly, the high recurrence of the fusion transcripts formed between adjacent genes in AML provides a new targeting area to study disease mechanism and a rich resource to search for potential markers for clinical applications, particularly for the


A.
GCNT2
NFYB
PRDM1
NFYA
NFYC
SLC26A3
RFXANK
ZXDC
CIITA
RFX5
RFXAP
CREB1
ZXDA
KAT2B
IRF2
STAT1
SMARCA4
RFX1
IRF1
IFNG
HLA-DRA

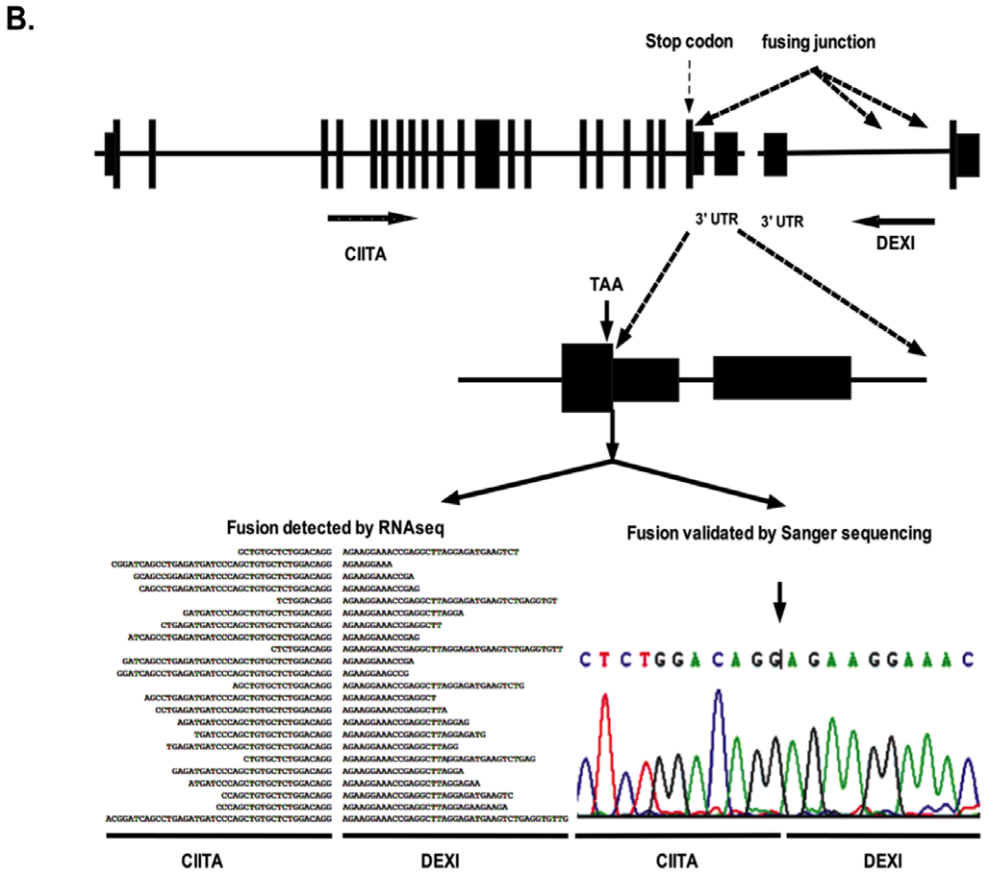

B.
Stop codon
fusing junction
3' UTR
3' UTR
CIITA
DEXI
TAA
Fusion detected by RNAseq
Fusion validated by Sanger sequencing
CIITA
DEXI
CIITA
DEXI

Figure 6. *CIITA-DEXI* fusion. A. CIITA-involved protein-protein interaction network. B. Mapping and validation of *CIITA-DEXI* fusion. Three *CIITA-DEXI* fusions were detected. In the fusion, *CIITA* preserved the coding exone till the stop codon but losed its 3′ untranslated region. This fusion was detected by 23 paired-end RNAseq sequences and validated by Sanger sequencing.
doi:10.1371/journal.pone.0051203.g006

cancer types with intact genomic structure, as examplified by the normal karyotype AML.

Methods

AML samples

The study used 29 primary samples of normal karyotype AML, eight AML primary samples with abnormal karyotype AML and eight AML primary samples without karyotype information [Figure 1], of which 38 were from University of Michigan Comprehensive Cancer Center and 7 were from Allcells Inc (AllCells, CA). The use of samples from University of Michigan was approved by the University of Michigan Institutional Review Board (IRBMED #2004-1022), and written informed consent was obtained from all patients prior to enrollment.

RNAseq process

Total RNA was extracted from each AML sample using TRIZOL reagent (Invitrogen). RNAseq library was prepared using PolyA+ mRNA isolated from each RNA sample following Illumina RNASeq protocol. Paired-end sequencing at 2×50 was performed by using Illumina GAIIx sequencer. Raw sequence data were processed by the standard GAIIx quality control program. FusionSeq [20] and deFuse [21] programs were used for fusion detection following the instructions. In FusionSeq mapping, reads from each end of a pair were mapped to the human reference genome sequences (hg18) independently. A set of filters, including PCR filter, homology filter, annotation filter, repeat filter, abnormal insert size filter and distance filter with default parameters in the program were used to remove low quality fusion candidates. Conditions were further set that a candidate fusion must be supported by at least two paired-end reads crossing the same boundary and the mapped sequences in both partner genes must contain at least 10 bp. In deFuse mapping, paired-end reads were mapped to both spliced and unspliced transcripts from Ensembl (version 62). Those mapped to the same transcript, mitochondrion and ribosome RNA were removed. The remaining paired-reads were trimmed and mapped again to the spliced and unspliced transcript sequences to remove those mapped to the same transcript. The remaining reads were clustered into contigs. Paired-end reads with only one end mapped were used to find breakpoint. The resulting fusions were further filtered based on the mapping of exon junction sequence (e.g., GT or AG) to generate the final fusion candidates. The Conjoined genes identified from the human genome were obtained from the ConjoinG database [22] (http://metasystems.riken.jp/conjoing/). The fusions identified from prostate cancer, breast cancer, ovarian cancer, lung cancer and colorectal cancer were obtained from the references [14,15,16,17,18,19].

AML cell lines

Eight cell lines of AML origin were used for validation study. Cell lines of KG-1, THP-1, MV4-11, HL-60, Kasumi-1 and Kasumi-6 were obtained from the American Type Culture Collection; Molm-13 and Molm-14 were provided by John Kersey (University of Minnesota). All cell lines were cultured in RPMI-1640, 10% fetal calf serum, 2 mM L-glutamine, 50 units/ml penicillin and 50 μg/ml streptomycin (Invitrogen). Total RNA was isolated from each cell line (5×10^6 cells) using the RNeasy Protect Cell Mini Kit (QIAGEN).

Experimental validation

Strand-specific RT-PCR and Sanger sequencing were used to validate the fusion transcripts identified by sequence mapping analysis. For each fusion selected for the test, at least one AML case was used for the validation. For primer design, sequences longer than the paired-end sequences covering the fusion were generated by extending the fusion–mapped location upstream and downstream within the exon region. Sense and antisense primers were designed using Primer3 program (http://frodo.wi.mit.edu/). Strand-specific reverse transcription was performed with either antisense primer or sense primer (2 pmole), total RNA (100 ng), M-MLV Reverse Transcriptase (2 units/reaction, Invitrogen) in a 20 μl volume at 65°C for 5 minutes, 37°C for 50 minutes, and 70°C for 15 minutes. Five μl of the resulting cDNA were used for PCR amplification in 50 μl containing sense primer (10 pmole), antisense primer (10 pmole), Taq polymerase (1.25 U, Promega), $MgCl_2$ (1.5 mM) in an ABI 9700 thermal cycler under cycling conditions of 95°C 7 minutes, 38 cycles of 95°C 30 seconds, 56°C 30 seconds, and 72°C 30 seconds, and final extension at 72°C 7 minutes. PCR products were loaded on 2% agarose gels. The bands with expected size were excised from gel, purified with MinElute Gel Extraction Kit (QIAGEN), and sequenced by BigDye reagents (Life Technologies). Sequences were mapped to hg18 by Genetyx program (version 7.03, Genetyx).

Long-range PCR

The same sense/antisense primer pairs used for RT-PCR were used to amplify the genomic regions by using a LongRange PCR kit (QIAGEN) with cycling conditions of 93°C 3 minutes, 35 cycles of 93°C 15 seconds, 55°C 30 seconds, and 68°C 4 minutes, and final extension at 68°C 7 minutes. The amplified products were checked on 0.8% agarose gels and the sizes were estimated by comparing with 1 kb DNA ladder (BioLabs).

Protein-protein interaction analysis

String 9.0 program was used to identify the interaction proteins for CIITA [39, http://string-db.org/newstring_cgi/show_input_page.pl?UserId = RUj0Zcmw26Xa&sessionId = cVwSUCEMLeFG]. Confidence score was set at the highest level (0.900), and the interactors shown were set at no more than 20.

Supporting Information

Figure S1 RNAseq mapped fusion.
(PDF)

Figure S2 Genomic patterns of fusion partner genes.
(PDF)

Table S1 Full fusion transcript list.
(XLS)

Table S2 Validated fusion transcript list.
(XLS)

Acknowledgments

The authors thank Markus Nabolz for comments on data analysis and intepretation, Cheng Wang and Xiangmin Lv for technical assistance, Tulika Prakash and Todd Taylor for providing the Conjoined gene data.

Author Contributions

Conceived and designed the experiments: SMW MZ JR. Performed the experiments: HW JX PC FX XH XZ. Analyzed the data: YL YCK ZX SM MQZ SMW. Contributed reagents/materials/analysis tools: SNM. Wrote the paper: SMW HW JR.

References

1. Mitelman F, Johansson B, Mertens F (2004) Fusion genes and rearranged genes as a linear function of chromosome aberrations in cancer. Nat Genet 36: 331–334.
2. Rowley JD (1998) The critical role of chromosome translocations in human leukemia. Annu Rev Genet 32: 495–519.
3. Grimwade D, Moorman A, Hills R, Wheatley K, Walker H, et al. (2004) Impact of karyotype on treatment outcome in acute myeloid leukemia. Ann Hematol 83 Suppl 1: S45–48.
4. Grimwade D, Walker H, Harrison G, Oliver F, Chatters S, et al. (2001) The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. Blood 98: 1312–1320.
5. Nimer SD (2008) Is it important to decipher the heterogeneity of "normal karyotype AML"? Best Pract Res Clin Haematol 21: 43–52.
6. Marcucci G, Mrozek K, Bloomfield CD (2005) Molecular heterogeneity and prognostic biomarkers in adults with acute myeloid leukemia and normal cytogenetics. Curr Opin Hematol 12: 68–75.
7. Mrózek K, Marcucci G, Paschka P, Whitman SP, Bloomfield CD (2007) Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification? Blood 109: 431–448.
8. Schlenk RF, Döhner K, Krauter J, Fröhling S, Corbacioglu A, et al. (2008) Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. N Engl J Med 358: 1909–1918.
9. Strunk CJ, Platzbecker U, Thiede C, Schaich M, Illmer T, et al. (2009) Elevated AF1q expression is a poor prognostic marker for adult acute myeloid leukemia patients with normal cytogenetics. Am J Hematol 84: 308–309.
10. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, et al. (2008) DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 456: 66–72.
11. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, et al. (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med 361: 1058–1066.
12. Link DC, Kunter G, Kasai Y, Zhao Y, Miner T, et al. (2007) Distinct patterns of mutations occurring in de novo AML versus AML arising in the setting of severe congenital neutropenia. Blood 110: 1648–1655.
13. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57–63.
14. Pflueger D, Terry S, Sboner A, Habegger L, Esgueva R, et al. (2010) Discovery of non-ETS gene fusions in human prostate cancer using next generation RNA sequencing. Genome Res 21: 56–67.
15. Kannan K, Wang L, Wang J, Ittmann MM, Li W, et al. (2011) Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. Proc Natl Acad Sci U S A 108: 9172–9177.
16. Salzman J, Marinelli RJ, Wang PL, Green AE, Nielsen JS, et al. (2011) ESRRA-C11orf20 is a recurrent gene fusion in serous ovarian carcinoma. PLoS Biol 9: e1001156.
17. Edgren H, Murumagi A, Kangaspeska S, Nicorici D, Hongisto V, et al. (2011) Identification of fusion genes in breast cancer by paired-end RNA-sequencing. Genome Biol 12: R6.
18. Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, et al. (2012) RET, ROS1 and ALK fusions in lung cancer. Nat Med 18: 378–381.
19. Lipson D, Capelletti M, Yelensky R, Otto G, Parker A, et al. (2012) Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med 18: 382–384.
20. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, et al. (2011) deFuse: An Algorithm for Gene Fusion Discovery in Tumor RNA-Seq Data. PLoS Comput Biol 7: e1001138.
21. Snel B, Lehmann G, Bork P, Huynen MA (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res 28: 3442–3444.
22. Prakash T, Sharma VK, Adati N, Ozawa R, Kumar N, et al. (2010) Expression of conjoined genes: another mechanism for gene regulation in eukaryotes. PLoS One 5: e13284.
23. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, et al. (2004) Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 34: 1513–1525.
24. Horsman DE, Staudt LM, Steidl U, Marra MA, Gascoyne RD (2011) MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 471: 377–381.
25. Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. Cell 137: 1194–1211.
26. Mathews LS, Vale WW (1991) Expression cloning of an activin receptor, a predicted transmembrane serine kinase. Cell 65: 973–982.
27. O'Carroll D, Scherthan H, Peters AH, Opravil S, Haynes AR, et al. (2000) Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. Mol Cell Biol 20: 9423–9433.
28. O'Hanlon TP, Raben N, Miller FW (1995) A novel gene oriented in a head-to-head configuration with the human histidyl-tRNA synthetase (HRS) gene encodes an mRNA that predicts a polypeptide homologous to HRS. Biochem Biophys Res Commun 210: 556–566.
29. Pierce SB, Chisholm KM, Lynch ED, Lee MK, Walsh T, et al. (2011) Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proc Natl Acad Sci U S A 108: 6543–6548.
30. Liu B, Yang Y, Qiu Z, Staron M, Hong F, Li Y, et al. (2010) Folding of Toll-like receptors by the HSP90 paralogue gp96 requires a substrate-specific cochaperone. Nat Commun 1: 79.
31. Luo B, Lam BS, Lee SH, Wey S, Zhou H, et al. (2011) The endoplasmic reticulum chaperone protein GRP94 is required for maintaining hematopoietic stem cell interactions with the adult bone marrow niche. PLoS One 6: e20364.
32. Audouard C, Le Masson F, Charry C, Li Z, Christians ES (2011) Oocyte-targeted deletion reveals that hsp90b1 is needed for the completion of first mitosis in mouse zygotes. PLoS One 6: e17109.
33. Wood CG, Mulders P (2009) Vitespen: a preclinical and clinical review. Future Oncol 5: 763–774.
34. Hata K, Andoh A, Shimada M, Fujino S, Bamba S, et al. (2002) IL-17 stimulates inflammatory responses via NF-kappaB and MAP kinase pathways in human colonic myofibroblasts. Am J Physiol Gastrointest Liver Physiol 282: G1035–1044.
35. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, et al. (2004) Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 34: 1513–1525.
36. Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, et al. (2011) MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 471: 377–381.
37. Ge X, Wu Q, Jung YC, Chen J, Wang SM (2006) A large quantity of novel human antisense transcripts detected by LongSAGE. Bioinformatics 22: 2475–2479.
38. Zhang Y, Gong M, Yuan H, Park HG, Frierson HF, et al. (2012) Chimeric Transcript Generated by cis-Splicing of Adjacent Genes Regulates Prostate Cancer Cell Proliferation. Cancer Discov 2: 598–607.
39. Sboner A, Habegger L, Pflueger D, Terry S, Chen DZ, et al. (2010) FusionSeq: a modular framework for finding gene fusions by analyzing paired-end RNA-sequencing data. Genome Biol 11: R104.