Cell Reports
Report

FuseFISH: Robust Detection of Transcribed Gene Fusions in Single Cells

Stefan Semrau,[1,2,3,9] Nicola Crosetto,[1,2,3,9,11] Magda Bienko,[1,2,3,10,11] Marina Boni,[4,10] Paolo Bernasconi,[4] Roberto Chiarle,[5,6,7,*] and Alexander van Oudenaarden[1,2,3,8,*]
[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Division of Hematology, Fondazione IRCCS Policlinico San Matteo, University of Pavia, 27100 Pavia, Italy
[5]Department of Pathology, Children's Hospital, Boston, MA 02138, USA
[6]Harvard Medical School, Boston, MA 02138, USA
[7]Department of Molecular Biotechnology and Health Sciences, University of Torino, 10126 Torino, Italy
[8]Hubrecht Institute–KNAW (Royal Netherlands Academy of Arts and Sciences) and University Medical Center Utrecht, 3584 CT Utrecht, The Netherlands
[9]These authors contributed equally to this work
[10]These authors contributed equally to this work
[11]Present address: Hubrecht Institute–KNAW (Royal Netherlands Academy of Arts and Sciences) and University Medical Center Utrecht, 3584 CT Utrecht, The Netherlands
*Correspondence: roberto.chiarle@childrens.harvard.edu (R.C.), a.vanoudenaarden@hubrecht.eu (A.v.O.)
http://dx.doi.org/10.1016/j.celrep.2013.12.002

SUMMARY

Transcribed gene fusions are key biomarkers in many hematologic and solid tumors, often representing the primary oncogenic driver mutation. Here, we report an experimental and computational pipeline for detecting fusion transcripts using single-molecule RNA FISH and unbiased correlation analysis (FuseFISH). We constructed a genome-wide database of optimal oligonucleotide sequences, enabling quick design of FuseFISH probes against known and novel fusions. We implemented FuseFISH in cell lines, tissue sections, and purified RNA, reliably detecting one *BCR-ABL1* positive in 10,000 negative cells. In 34 hematologic samples, we detected *BCR-ABL1* transcripts with high specificity and sensitivity. Finally, we measured *BCR-ABL1* expression heterogeneity and dynamics in single CML cells exposed to the kinase inhibitor Nilotinib. Our resource and methods are ideal for streamlined validation of fusions newly identified by next-generation sequencing, and they pave the way to studying the impact of fusion expression variability on clinical outcome.

INTRODUCTION

Cytogenetic abnormalities such as translocations, inversions, and insertions are characteristic attributes of cancer cells, and they often result in the formation of chimeric genes consisting of segments of two different genes fused together (Fröhling and Döhner, 2008). In most cases, the chimeric gene is transcribed into a fusion transcript encoding parts of a tyrosine kinase or a transcription factor, which become deregulated as a consequence of the fusion (Fröhling and Döhner, 2008). Often a gene fusion represents the primary oncogenic driver mutation in a tumor and hence an ideal pharmacologic target, as demonstrated by the archetypical case of BCR-ABL1 and its selective inhibitor, Imatinib, in chronic myeloid leukemia (CML) (Melo and Barnes, 2007; Ren, 2005; Schiffer, 2007). Therefore, detecting and monitoring in time the expression levels of specific gene fusions in cancer has become common practice in molecular pathology. While recurrent fusions have long been known in hematologic tumors and sarcomas (Mitelman et al., 2007; Rowley, 2009), recent progress in next-generation sequencing technologies has fueled the discovery of new fusions in solid tumors (Maher et al., 2009; Rabbitts, 2009), as exemplified by EML4-ALK in a subset of non-small cell lung cancers (NSCLCs; Soda et al., 2007). Thus, demand for simple and quantitative assays to detect a broad spectrum of fusions will likely emerge in the future.

Detection of fusions is typically achieved at the DNA level by karyotype analysis and DNA fluorescence in situ hybridization (FISH), or at the RNA level by RT-PCR. In spite of the effectiveness and broad use of these techniques, several limitations call for new complementary methods. For instance, even though considerable progress in image processing automation has been done (Alpár et al., 2008; Lerner et al., 2001; Shirley et al., 2011), data analysis of DNA FISH remains difficult to standardize and automate because colocalization of dual-fusion probes or splitting of break-apart probes is usually assessed in a subjective manner. Importantly, DNA FISH is unable to provide information about expression levels of the fusions, which is clinically relevant information (Baccarani et al., 2009). On the other hand, RT-PCR

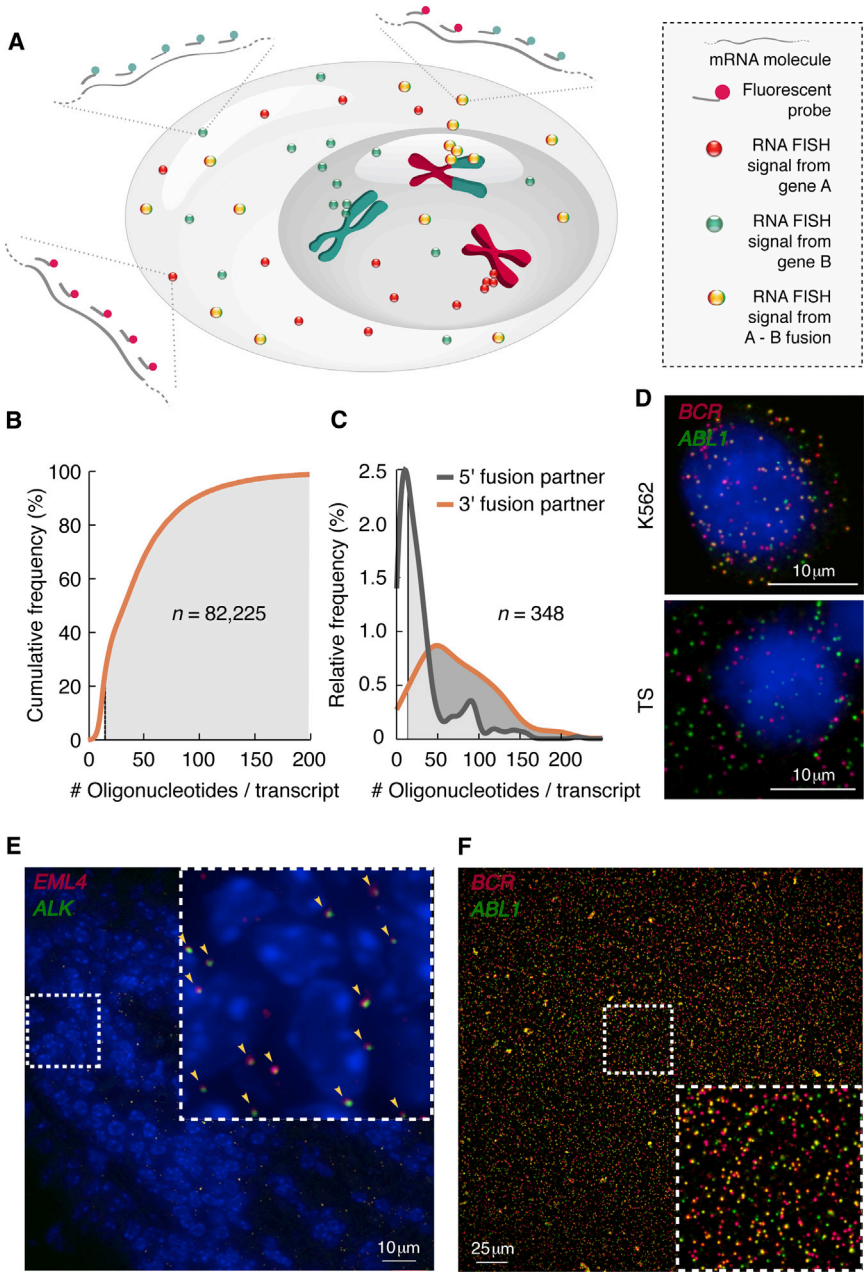


Figure 1. FuseFISH Method

(A) A gene fusion originating from a translocation (magenta-green chromosome) produces fusion transcripts (yellow dots) that are detected as colocalized fluorescent spots by probe sets consisting of differently labeled oligos targeting each fusion partner (zoom-in).

(B) Cumulative frequency of smFISH oligos per coding transcript in the human transcriptome. A total of 79% of transcripts contain ≥15 oligos (gray area). *n*, transcript number.

(C) Relative frequency of smFISH oligos for either the 5′ (gray curve) or 3′ (orange curve) moiety of the shortest variant of 348 unique fusions cataloged in Table S1. Gray area: fraction of transcripts with ≥15 oligos coverage (57% for 5′ and 92% for 3′ partners). *n*, number of fusions analyzed.

(D) *BCR-ABL1* transcripts (yellow spots) in positive (K562) versus negative (TS) cells.

(E) *EML4-ALK* transcripts (yellow spots) in mouse xenografts of lung adenocarcinoma H3122 cells. Yellow arrowheads in the inset indicate examples of fusions.

(F) *BCR-ABL1* fusion transcripts (yellow spots) in spotted purified RNA extracted from K562 cells. See also Figure S1.

is a powerful method to quantify expression, but the development of standardized and reproducible assays for absolute quantification of fusion transcripts can be challenging, especially in formalin-fixed, paraffin-embedded (FFPE) tissue sections. In addition, fusion transcripts often involve different exons in independent clinical samples, thus requiring multiple PCR reactions and controls for their detection. Another limitation relates to the use of RT-PCR in single cells to monitor intratumor expression heterogeneity, which seems clinically informative (La Thangue and Kerr, 2011; Marusyk et al., 2012). Though technically feasible, routine clinical application of single-cell RT-PCR in the clinical context remains challenging, especially in solid tumors, and is associated with high costs for a relatively moderate throughput.

Here, we sought to develop a robust and unbiased experimental and computational framework for detecting specific fusion transcripts in situ or using purified RNA. We demonstrate the feasibility and simplicity of our approach for a variety of fusion transcripts in cell lines, tumor sections, and hematologic specimens. Our resource and methods can be readily applied to biological studies of gene fusions and integrated into clinical cytogenetics.

RESULTS

Method and Probe Resource

In order to detect fusion transcripts at single-molecule resolution, we capitalized on a method for single-molecule RNA FISH (smFISH) previously developed by our group (Raj et al., 2008) based on earlier work (Femino et al., 1998) and on a recent method to detect different mRNA isoforms (Waks et al., 2011). We devised an approach by which each fusion partner is labeled with a set of oligonucleotides coupled to a specific fluorophore, so that the resulting fusion can be detected as two spectrally distinguishable, colocalized diffraction-limited spots (Figure 1A; see Experimental Procedures). We named this approach FuseFISH.

To facilitate smFISH probe design, we compiled a genome-wide list of oligonucleotides with optimal thermodynamic properties targeting 82,225 annotated human protein-coding

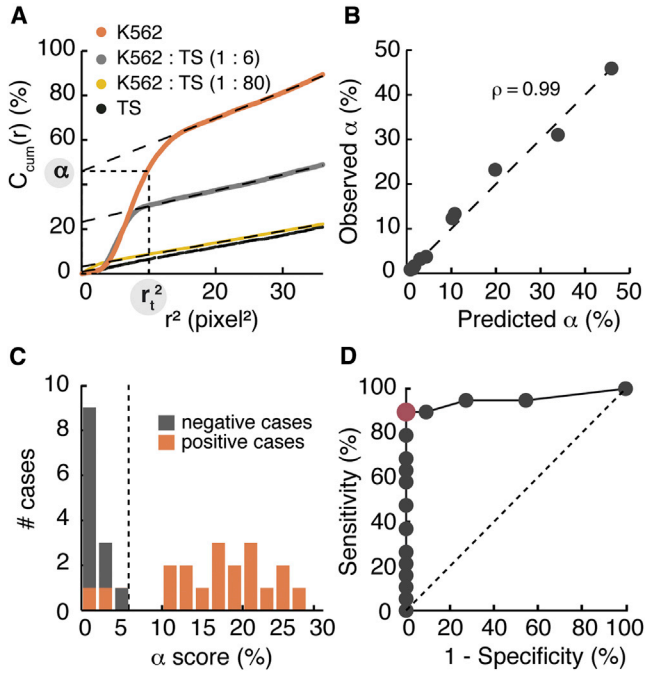


Figure 2. FuseFISH Validation

(A) PICCS correlation analysis of *BCR-ABL1* in K562 cells serially diluted with TS cells. Fitting of the linear portion of each curve yields the correlation fraction α as the offset of the linear fit. The optimal distance r_t threshold balances false positives and negatives such that correlating signals closer than r_t results in the correlation fraction α.

(B) Measured versus predicted scaling of α in the same cell dilutions as in (A).

(C) Distribution of α in hematologic samples. Optimal cut-off is shown as dashed line.

(D) ROC curve of FuseFISH versus DNA FISH and/or RT-PCR for detection of *BCR-ABL1* in hematologic samples. Optimality is achieved for $\alpha \geq 5\%$ (red). Dashed line, line of no discrimination.

See also Figure S2.

transcripts (see Experimental Procedures). For 79% of transcripts, at least 15 oligos can be synthesized, which is sufficient to produce a reliable smFISH signal (Figure 1B). This oligo library is a powerful resource for versatile and rapid design of smFISH probes and is readily accessible at our website (http://www.fusefish.eu).

To estimate coverage for transcript fusions, we separately calculated the number of oligos available for the 5′ and 3′ fusion partner of 348 fused transcripts for which we managed to retrieve the exact breakpoint coordinates by comprehensive literature and database screening (Table S1; see Supplemental Experimental Procedures). To calculate the minimum number of oligos available, we selected the shortest variant when multiple fusion variants were annotated. On average, 30 and 76 oligos are available for 5′ and 3′ fusion partners, respectively, the difference reflecting the fact that 3′ UTRs tend to be longer than 5′ UTRs in most transcripts (Figure 1C). We note that this coverage represents a very stringent estimate, in the sense that it is based on the shortest fusion variants described in the literature, even though their frequency is low in most of cases.

FuseFISH Implementation

As a proof of principle, we constructed probes targeting the most frequent variants of *BCR-ABL1*, *NPM1-ALK*, and *EML4-ALK* fusions in CML, anaplastic large cell lymphoma, and NSCLC, respectively (Figure S1A; Table S2). In positive cell lines, we observed fusion-specific colocalized transcripts mixed with wild-type transcripts (Figure 1D). Next, we applied FuseFISH to FFPE tissue sections of mouse xenografts of human *EML4-ALK* positive NSCLC cells. Similar to cultured cells, FuseFISH was able to detect fusion transcripts amidst wild-type ones (Figure 1E).

While detection of fusion transcripts in cells and tissues can be powerful in cytogenetic diagnostics of well-characterized fusions, detection in purified RNA would be desirable for rapid validation of newly discovered fusions. To maximize the versatility of FuseFISH, we devised a protocol for smFISH using purified RNA spotted on microscopy slides (Figures S1B and S1C; see Supplemental Experimental Procedures). We spotted RNA extracted from the CML cell line K562 and used probes targeting *BCR* and *ABL1* transcripts. As expected, *BCR-ABL1* fusion transcripts appeared as bright, colocalized fluorescent spots (Figure 1F).

Automated Signal Detection

Though spatial colocalization of fusion transcripts is immediately apparent in the above experiments, unbiased quantification is unfeasible by eye. To achieve fully automated and unbiased fusion detection, we implemented a computational pipeline that determines the average fraction α of signals from a given transcript species A (e.g., *BCR*) with a correlated (i.e., nonrandomly colocalized) signal from a different transcript species B (e.g., *ABL1*), based on particle image cross-correlation spectroscopy (PICCS; Semrau et al., 2011; Figures 2A and S2A; see Experimental Procedures). Briefly, we constructed a cumulative correlation function $C_{cum}(r)$ by counting the average number of B signals neighboring an A signal depending on the distance r of separation. Linear fitting of $C_{cum}(r)$ versus r^2 at distances exceeding the typical correlation length gave the correlation fraction α as the offset of the fitted line. We initially processed 3D image stacks to reduce the incidence of random colocalization. However, at observed transcript densities, we obtained similar correlation fractions and accuracies using maximum projections and 2D analysis (Figure S2B). Since the computational effort is substantially smaller for 2D analysis, we applied this approach in all subsequent analyses. We compared α in cells expressing a particular fusion with negative cells. As expected, α was systematically higher in positive cells (Figure S2C). We also compared formalin fixation to fixation in methanol-acetic acid—the standard fixative in cytogenetics—and obtained very similar α values (Figure S2C). To prove that FuseFISH is quantitative, we calculated $\alpha_{BCR-ABL1}$ for serial dilutions of K562 with *BCR-ABL1*-negative TS cells. We observed strong linear scaling of $\alpha_{BCR-ABL1}$ with our prediction based on pure population measurements (Figure 2B). We then determined the sensitivity of FuseFISH by estimating the accuracy of the measured α by bootstrapping. A mix of K562 and TS cells at 1:80 ratio yielded a significantly higher $\alpha_{BCR-ABL1}$ value compared to TS cells alone (3.2% versus 0.8%, $p < 10^{-3}$). Finally, we estimated the minimal amount of signals that need to be measured to determine α with a given accuracy. We

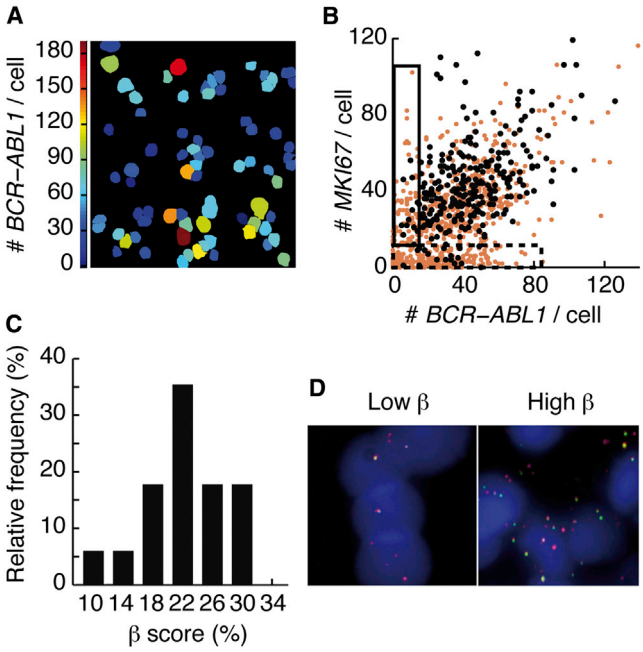


Figure 3. Applications of FuseFISH

(A) Cell-to-cell *BCR-ABL1* expression variability in K562 cells.

(B) *BCR-ABL1* and *MKI67* expression in K562 cells exposed to Nilotinib (orange) or vehicle (black) for 48 hr. Dashed box: heterogeneous *BCR-ABL1* expression in Nilotinib-treated $MKI67^{Low}$ cells. Thick box: $BCR\text{-}ABL1^{Low}$-$MKI67^{High}$ Nilotinib-treated subpopulation.

(C) Intertumor variability of local correlation fraction β.

(D) *BCR-ABL1* expressing cells representative of the left (Low β) and right (High β) portion of the histogram in (C).

See also Figure S3.

modeled PICCS accuracy as a function of several parameters (Figure S2D). For example, if $\alpha = 5\%$ is to be measured with an error margin of 1% at a density of ten transcripts per cell (100 μm^2 cell area), the model predicts that approximately 1,000 signals (100 cells) have to be imaged.

Detection of rare fusion events holds important biological and clinical implications. For example, monitoring of decreasing *BCR-ABL1* levels in CML patients is crucial to assess therapeutic response (Kantarjian et al., 2008; Radich, 2009). However, as the frequency of a fusion decreases in a population of cells, the imaging time required to obtain statistically significant results becomes quickly unpractical. We reasoned that performing FuseFISH in concentrated purified RNA would allow us to detect rare fusions, maximizing the sensitivity of our method. Thus, we devised a protocol to quickly and easily immobilize purified RNA on microscope slides (see Supplemental Experimental Procedures) and imaged thousands of transcripts from a mixture of K562 and TS cells. By this approach, we reliably detected *BCR-ABL1* fusions derived from as little as one K562 cell per 10,000 TS cells (Figure S2E).

Validation of FuseFISH

To validate our approach, we compared FuseFISH with gold standard methods used for *BCR-ABL1* diagnostics (see Experimental Procedures). We retrospectively analyzed 34 hematologic samples fixed in methanol-acetic acid, including 17 cases of CML and 5 cases of acute lymphoblastic leukemia expressing *BCR-ABL1* (age: median, 56 years; range, 52–63 years; see Table S3). We derived a receiver operating characteristic (ROC) curve by comparing FuseFISH α to the diagnosis (*BCR-ABL1* positivity or negativity) previously formulated based on DNA FISH and/or RT-PCR. A score $\alpha \geq 5\%$ correctly identified positive cases with 100% specificity and 89% sensitivity, and α positively correlated ($r^2 = 0.62$, $p = 0.05$) with RT-PCR results (Figures 2C, 2D, and S2F). We also performed systematic cross-validation and determined that our algorithm achieves maximal sensitivity and specificity when trained on $\geq$25 cases (Figure S2G).

BCR-ABL1 Cell-to-Cell Variability

Until now, robust quantification of fusion transcript in single cells has been challenging. Such measurement would advance our understanding of tumor heterogeneity and its implications on disease prognosis and therapeutic response. As a proof of principle, we applied FuseFISH to study fusion transcript expression variability and dynamics in single cancer cells upon fusion-targeted therapy. We first assessed the expression of *BCR-ABL1* in K562 cells and found that it was noisy (coefficient of variation [CV] = 0.48; Figure 3A). Next, we simultaneously measured *BCR-ABL1* and the proliferation marker *MKI67* in hundreds of individual K562 cells undergoing treatment with the specific tyrosine kinase inhibitor, Nilotinib. As expected, upon treatment the fraction of proliferating cells decreased substantially (Figures 3B and S3A). Based on *MKI67* expression upon Nilotinib treatment, we classified cells into $MKI67^{High}$ and $MKI67^{Low}$. After 24 hr exposure to Nilotinib, the $MKI67^{High}$ subpopulation exhibited increased expression variability compared to vehicle control (CV = 0.6, $p = 5 \times 10^{-3}$, F test). Interestingly, we also observed a subpopulation of Nilotinib-treated $MKI67^{High}$ cells with lower *BCR-ABL1* expression compared to untreated cells (Figures 3B and S3A–S3C).

Finally, we measured *BCR-ABL1* variability in our set of hematologic samples. Since cell segmentation can be challenging in these specimens, we computed a local correlation fraction, β, reporting on the *BCR-ABL1*/*BCR* ratio in the neighborhood of *BCR-ABL1* signals (see Experimental Procedures). In serial dilutions of K562 with TS cells, β was constant over a wide range of dilution rates ($\alpha \geq 5\%$) and close to the α value calculated for a pure population of K562 cells (Figure S3D). Thus, β is a faithful proxy of the level of *BCR-ABL1* expression in positive cells. In *BCR-ABL1*-positive specimens the β score varied substantially from cases expressing low *BCR-ABL1* per cell (Low β) to cases expressing high amounts (High β) (Figures 3C and 3D). Interestingly, in seven *BCR-ABL1*-positive cases with known progression-free survival (PFS), we found that PFS tended to be shorter in patients with high α and β; however, without reaching significance ($p > 0.05$) due to low statistical power (Table S3). In the future, ad hoc prospective trials should be designed to clarify the impact of fusion expression and variability on clinical outcome.

DISCUSSION

We have devised a readily applicable and versatile experimental and computational framework that can serve multiple purposes

and complements existing technologies for gene fusion detection. Our genome-scale oligo database is a powerful resource thanks to which FuseFISH probes (and, more generally, smFISH probes) can be easily and rapidly designed, making the method particularly suitable to confirm newly discovered fusions.

Since it can be applied to methanol-acetic acid fixed cells—the standard fixation procedure in cytogenetics—FuseFISH can be integrated in diagnostics practice, accelerating assay time (hybridization, imaging, and automated signal quantification can be performed on the same day) and eliminating subjective biases in scoring colocalized or split-apart DNA FISH signals. We expect that the fully automated and unbiased correlation analysis achievable with our computational pipeline will facilitate fusion transcript detection and help improve interlaboratory reproducibility.

The ability to visualize individual transcripts using purified RNA is particularly beneficial for quick and reliable validation of fusions discovered by next-generation sequencing techniques such as RNA sequencing. Purified RNA FuseFISH might also be applied when the fraction of fusion-expressing cells is low, such as for monitoring *BCR-ABL1* levels in CML patients during targeted therapy. While FuseFISH sensitivity is currently ten times lower than for RT-PCR (Kantarjian et al., 2008; Radich, 2009)—the current standard to monitor minimal residual disease in CML patients—implementation of the method on a high-throughput imaging system could further improve sensitivity.

Investigation of transcriptional heterogeneity in early developing cancers is in its infancy, and experimental and computational frameworks to measure fusion transcript variability are missing. FuseFISH opens the possibility to investigate the dynamics of fusion expression in response to therapeutic agents, relate expression variability to clinical outcomes, and simultaneously assess specific gene expression signatures with defined fusions in individual cells. In the long term, it will be important to design prospective trials to test whether these features predict patients' response to fusion-targeted drugs and whether measuring fusion expression in single cells should be implemented in clinical practice.

EXPERIMENTAL PROCEDURES

smFISH Probe Database

All bioinformatic analyses were performed using custom scripts written in MATLAB. We retrieved the sequence of all human protein-coding transcripts by linear concatenation of the exons of each transcript in the ENSEMBL annotated gene database (release 70, January 2013). We scanned each transcript except the last 19 nt using 20 nt windows sliding in 1 nt steps, computed the guanine-cytosine (GC) content, and then assembled all windows with GC content comprised between 40% and 60%. We designed as many 20 nt oligos as possible separated by at least 2 nt in each region with optimal GC content, starting from the 5′ end of each transcript. We maximized the number of oligos with GC = 45% or GC = 50% by computing the GC content of 20 nt oligos starting from each of five consecutive nucleotide positions and selecting the most 5′ position with GC = 45% or GC = 50%, if available. We screened the uniqueness of designed oligos by performing a local BLAST search against the NCBI RefSeq_RNA database (http://www.ncbi.nlm.nih.gov/refseq/) using *blastn*.

smFISH

Probes targeting *BCR* (48 oligos), *ABL1* (48 oligos), *NPM1* (16 oligos), *ALK* (48 oligos), and *EML4* (48 oligos) transcripts consisted of amine-labeled oligos labeled with Cy5 (GE Healthcare, catalog no. Q15108), Alexa Fluor 594 (Molecular Probes, catalog no. A20004), or 6-TAMRA (Molecular Probes, catalog no. C6123) (Table S2). Hybridizations and washes were done according to modified protocols based on previously described procedures (Bienko et al., 2013; Raj et al., 2008; see Supplemental Experimental Procedures for details).

Microscopy

We performed imaging as described earlier (Bienko et al., 2013) using an inverted epi-fluorescence microscope (Nikon) equipped with a high-resolution charge-coupled device camera (Pixis, Princeton Intruments). Magnification of 100× and 40× oil immersion, high-numerical-aperture Nikon objectives were used for cells and tissues, and purified RNA, respectively. Per region of interest, we typically acquired an image stack consisting of 40 (cells and tissues) or 20 (purified RNA) image planes spaced 0.2 μm apart. Details on image processing are available in the Supplemental Experimental Procedures.

PICCS

Due to chromatic aberration and unavoidable imperfections of microscopy setup, the two signals stemming from the two moieties of a fusion transcript are shifted with respect to each other. A typical solution is defining an arbitrary distance threshold below which two signals are considered colocalized (i.e., correlated). However, especially when expression levels are high (and signals are, therefore, dense), results depend sensitively on the choice of this threshold. To circumvent this limitation, we used PICCS (Semrau et al., 2011) to calculate the global level of correlation, α, between fusion partners in an unbiased way. We constructed the cumulative distribution function, C_{cum}, of distances between signals from differentially labeled species (A and B) by counting the number of all B signals within a distance r of an A signal and dividing it by the number of A signals (Figure S2A). The resulting $C_{cum}(r)$ function reports the average number of B neighbors (of an A signal) closer than r and has the general form $C_{cum}(r) = \alpha_{AB} P_{cum}(r) + \pi c_B r^2$, where α_{AB} is the correlation fraction (i.e., the fraction of A signals which have a correlated B signal), c_B is the average density of B signals, and P_{cum} is the cumulative distribution function of distances between truly correlated signals. When plotted versus r^2, the contribution of uncorrelated B signals $\pi c_B r^2$ can be estimated by linear fitting of $C_{cum}(r)$ at distances exceeding the correlation length. The fit line intercepts the y axis at the correlation fraction α, and its slope is proportional to the density of B signals c_B (Figure 2A).

ROC Analysis and Cross-Validation

We constructed ROC curves by computing the rate of true positives (sensitivity) and false positives (100% − specificity) based on specimen classification by DNA FISH and/or RT-PCR. We obtained separate ROC curves for various signal-to-noise ratio cut-offs. We performed cross-validation by repeated random sampling (without replacement) of a subset of specimens (training set), while the ratio of positive and negative cases was the same as in the full data set. For each sample, we determined an optimal α threshold using three different algorithms and calculated the false and true positive rates for the remaining specimens (test set). Algorithms 1 and 2 force the false positive rate to be below 0.001 while maximizing the true positive rate. While algorithm 1 maximizes the threshold, algorithm 2 tries to minimize the threshold if there is no negative impact on the true positive rate. Algorithm 3 maximizes the absolute difference between the false positive and the true positive rate, which is equivalent to the distance of a point on the ROC curve to the diagonal. Algorithm 1 was chosen because it led to smaller false positive and true positive rates compared to the other algorithms, and it saturated to false positive and true positive rates found for the complete data set (Figure S2G).

Local Correlation Fraction

The optimal threshold r_t is given by $C_{cum}(r) = \alpha_{AB} P_{cum}(r) + \pi c_B r^2$ as indicated in Figure 2A. We considered neighboring *BCR* and *ABL1* signals to be part of the same fusion transcript if they were closer than r_t. If single cells can be identified, *BCR-ABL1* expression can also be determined for individual cells. Unfortunately, abundant cell clumping in clinical samples precluded such analysis. However, having identified *BCR*, *ABL1*, and *BCR-ABL1* separately, we again used PICCS to determine a local correlation fraction, β, between *BCR-ABL1* and all *BCR*-containing transcripts. To do so, we started from *BCR-ABL1*

transcripts and counted the number of *BCR-ABL1* or all *BCR*-containing transcripts in the neighborhood, as shown in Figure S2A. The local correlation fraction, β, is the ratio of the retrieved densities, $\beta = c_{BCR-ABL1}/(c_{BCR-ABL1} + c_{BCR})$, reporting on the ratio of *BCR-ABL1* and total *BCR* expression in *BCR-ABL1*-expressing cells. For a perfectly homogeneous population of cells $\beta = \alpha$, whereas in a mixed population consisting of *BCR-ABL1*-expressing and -nonexpressing cells β is the local correlation fraction of positive cells and, in general, $\beta > \alpha$.

We retrieved hematological specimens in methanol-acetic acid 3/1 v/v from the archive of the Division of Hematology of IRCCS Policlinico San Matteo after approval of the study by the institutional Ethical Committee. Permission to conduct mice experiments was granted by the Italian Ministry of Health (approval number 218-2009-B).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.12.002.

AUTHOR CONTRIBUTIONS

N.C. conceived FuseFISH and the protocol for smFISH on purified RNA, designed the genome-wide probe database, performed experiments, and wrote the manuscript. S.S. developed all computational methods, performed all data analyses, conceived the experiments with Nilotinib, and wrote the manuscript. M. Bienko contributed ideas, performed experiments, and prepared figures. M. Boni and P.B. provided hematologic samples and relevant clinical information. R.C. contributed ideas and provided cell lines and mouse xenotransplants. A.v.O. supported and steered the project, provided ideas, and corrected the manuscript.

ACKNOWLEDGMENTS

This work was supported by MIT NIH/NCI Physical Sciences Oncology Center (U54CA143874 to A.v.O.), NIH Pioneer Award (1DP CA174420 to A.v.O.), NWO Vici Award (to A.v.O.), NWO Rubicon Award (to S.S.), FP7 ERC-2009-StG grant (proposal no. 242965 "Lunely" to R.C.), AIRC IG-12023 and AICR 12-0216 grants (to R.C.), and the Human Frontiers Science Program (to M. Bienko). We thank Sandy Klemm (A.v.O. lab) and Chiara Ambrogio (CNIO, Madrid) for critically reading the manuscript and help with analyses, Maria Stella Scalzo (R.C. lab) for assistance with tissue preparation, and Barbara Rocca and Celeste Calvello (P.B. lab) for RT-PCR data.

Received: August 12, 2013
Revised: November 5, 2013
Accepted: December 3, 2013
Published: December 26, 2013

REFERENCES

Alpár, D., Hermesz, J., Pótó, L., László, R., Kereskai, L., Jáksó, P., Pajor, G., Pajor, L., and Kajtár, B. (2008). Automated FISH analysis using dual-fusion and break-apart probes on paraffin-embedded tissue sections. Cytometry A *73*, 651–657.

Baccarani, M., Cortes, J., Pane, F., Niederwieser, D., Saglio, G., Apperley, J., Cervantes, F., Deininger, M., Gratwohl, A., Guilhot, F., et al.; European LeukemiaNet (2009). Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. J. Clin. Oncol. *27*, 6041–6051.

Bienko, M., Crosetto, N., Teytelman, L., Klemm, S., Itzkovitz, S., and van Oudenaarden, A. (2013). A versatile genome-scale PCR-based pipeline for high-definition DNA FISH. Nat. Methods *10*, 122–124.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts in situ. Science *280*, 585–590.

Fröhling, S., and Döhner, H. (2008). Chromosomal abnormalities in cancer. N. Engl. J. Med. *359*, 722–734.

Kantarjian, H., Schiffer, C., Jones, D., and Cortes, J. (2008). Monitoring the response and course of chronic myeloid leukemia in the modern era of BCR-ABL tyrosine kinase inhibitors: practical advice on the use and interpretation of monitoring methods. Blood *111*, 1774–1780.

La Thangue, N.B., and Kerr, D.J. (2011). Predictive biomarkers: a paradigm shift towards personalized cancer medicine. Nat. Rev. Clin. Oncol. *8*, 587–596.

Lerner, B., Clocksin, W.F., Dhanjal, S., Hultén, M.A., and Bishop, C.M. (2001). Automatic signal classification in fluorescence in situ hybridization images. Cytometry *43*, 87–93.

Maher, C.A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., and Chinnaiyan, A.M. (2009). Transcriptome sequencing to detect gene fusions in cancer. Nature *458*, 97–101.

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer *12*, 323–334.

Melo, J.V., and Barnes, D.J. (2007). Chronic myeloid leukaemia as a model of disease evolution in human cancer. Nat. Rev. Cancer *7*, 441–453.

Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nat. Rev. Cancer *7*, 233–245.

Rabbitts, T.H. (2009). Commonality but diversity in cancer gene fusions. Cell *137*, 391–395.

Radich, J.P. (2009). How I monitor residual disease in chronic myeloid leukemia. Blood *114*, 3376–3381.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat. Methods *5*, 877–879.

Ren, R. (2005). Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. Nat. Rev. Cancer *5*, 172–183.

Rowley, J.D. (2009). Chromosomes in leukemia and beyond: from irrelevant to central players. Annu. Rev. Genomics Hum. Genet. *10*, 1–18.

Schiffer, C.A. (2007). BCR-ABL tyrosine kinase inhibitors for chronic myelogenous leukemia. N. Engl. J. Med. *357*, 258–265.

Semrau, S., Holtzer, L., González-Gaitán, M., and Schmidt, T. (2011). Quantification of biological interactions with particle image cross-correlation spectroscopy (PICCS). Biophys. J. *100*, 1810–1818.

Shirley, J.W., Ty, S., Takebayashi, S.-I., Liu, X., and Gilbert, D.M. (2011). FISH Finder: a high-throughput tool for analyzing FISH images. Bioinformatics *27*, 933–938.

Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S.-I., Watanabe, H., Kurashina, K., Hatanaka, H., et al. (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561–566.

Waks, Z., Klein, A.M., and Silver, P.A. (2011). Cell-to-cell variability of alternative RNA splicing. Mol. Syst. Biol. *7*, 506.

FuseFISH: Robust Detection of Transcribed Gene Fusions in Single Cells

Stefan Semrau, Nicola Crosetto, Magda Bienko, Marina Boni, Paolo Bernasconi, Roberto Chiarle, Alexander van Oudenaarden

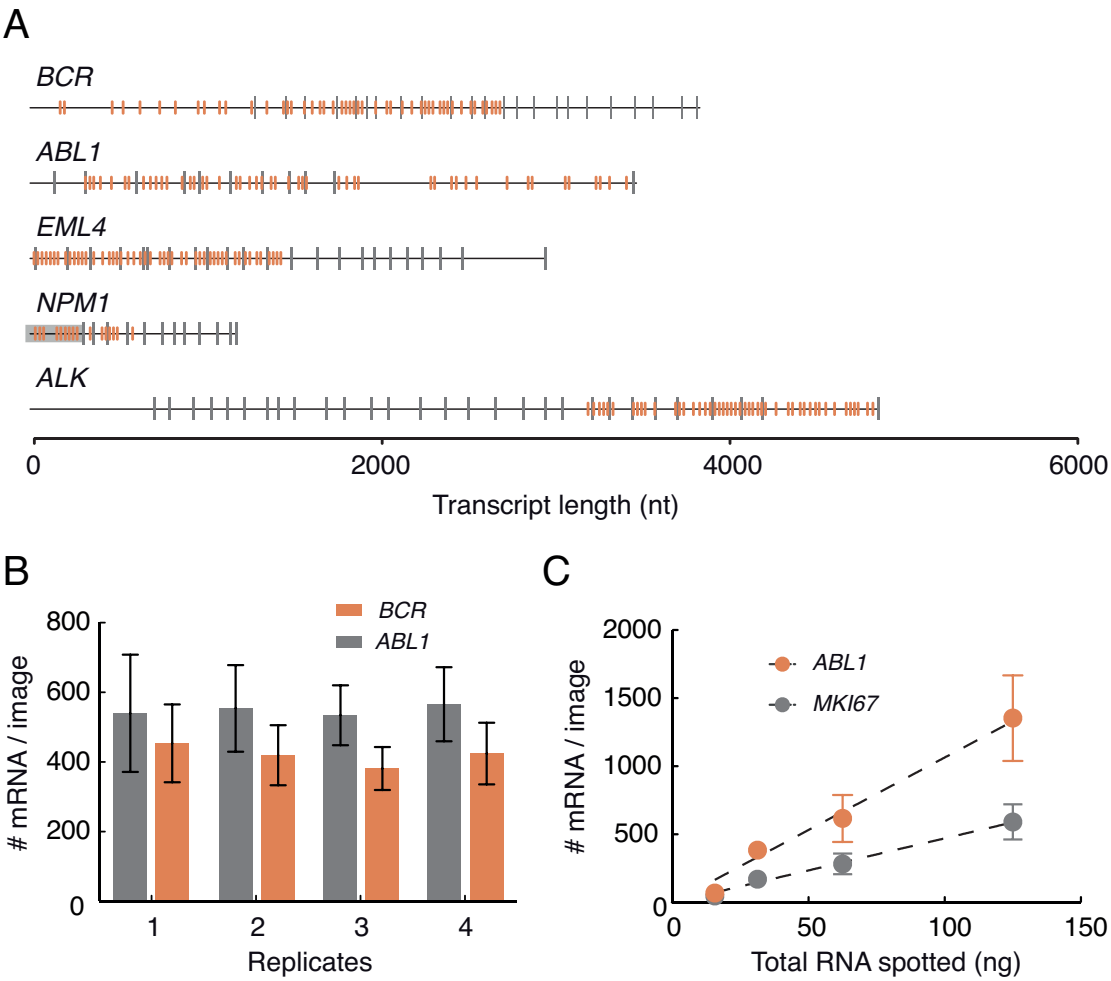
SUPPLEMENTAL INFORMATION

The following supplemental information is included:

- Supplemental Figure S1, Related to Figure 1
- Supplemental Figure S2, Related to Figure 2
- Supplemental Figure S3, Related to Figure 3
- Supplemental Experimental Procedures
- Supplemental Table S1, smFISH Oligo Coverage for Known Fused Transcripts, Related to Figure 1C
- Supplemental Table S2, List of smFISH probes, Related to Figure S1A
- Supplemental Table S3, Characteristics of Hematologic Samples and Donors, Related to Figures 2C and 2D
- Supplemental References

SUPPLEMENTAL FIGURES

Figure S1. FuseFISH Method, Related to Figure 1

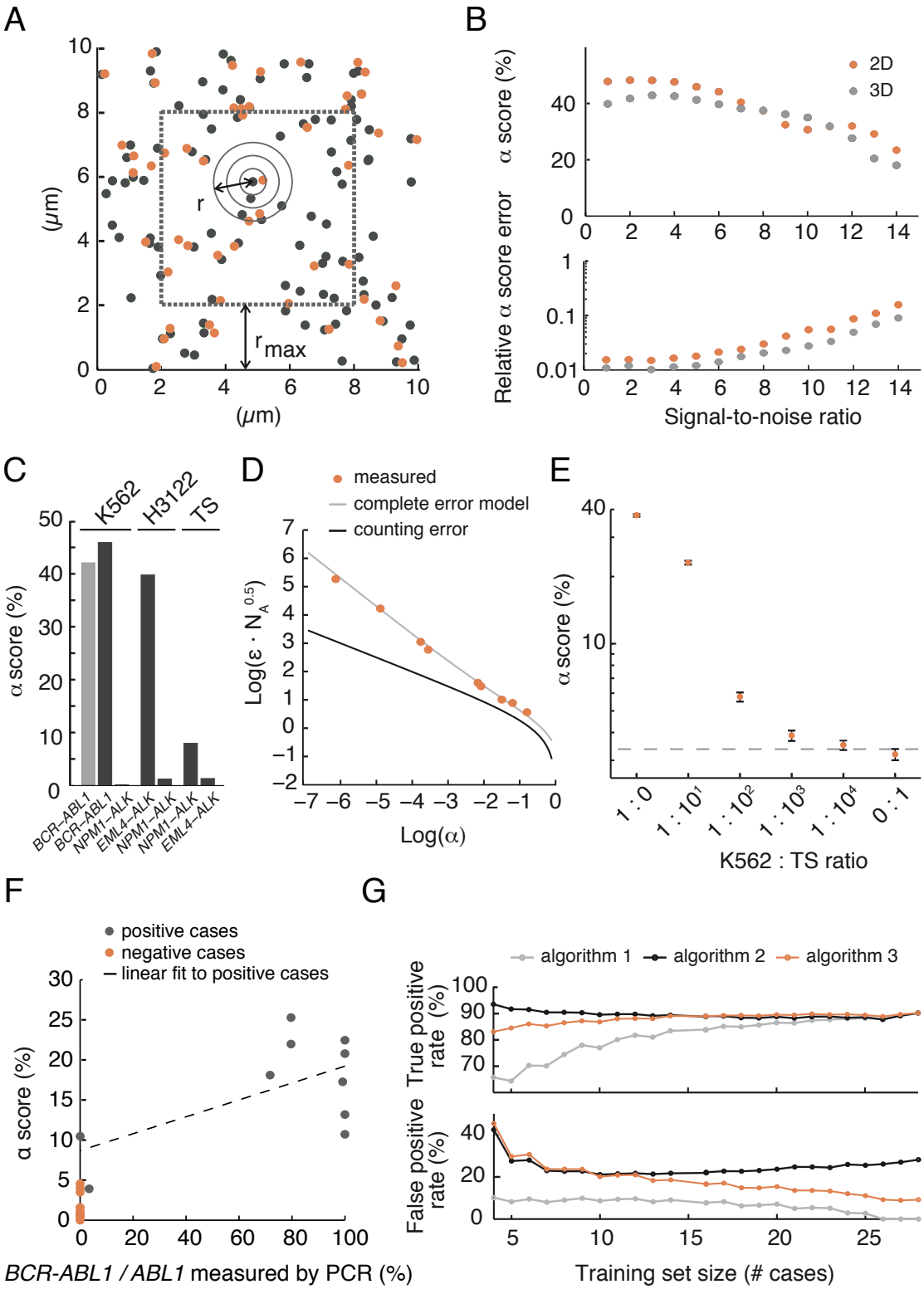


(**A**) Location of FuseFISH probes (orange marks) in *BCR-ABL1*, *NPM1-ALK*, and *EML4-ALK* fusion transcripts. Gray marks: exon-exon boundaries. Gray highlight: 5′ UTR.

(**B**) RNA spotting robustness. We spotted 125 ng of total RNA in four different experimental replicates. Mean ± s.d. is shown for each replicate.

(**C**) RNA spotting linearity. We spotted the indicated amounts of total RNA purified from K562 cells. Mean ± s.d. is shown for each amount. Dashed lines: linear fits. All data in (**B**) and (**C**) were acquired using a 100× magnification oil-immersion microscope objective.

Figure S2. FuseFISH validation, Related to Figure 2



(**A**) Illustration of the PICCS algorithm (see **Experimental Procedures** for details).

(**B**) Comparison of PICCS analysis on 2D maximum image projections versus full 3D images.

(**C**) α score for various fusion transcripts in different cell lines. Gray: cells fixed in formalin. Black: cells fixed in methanol-acetic acid.

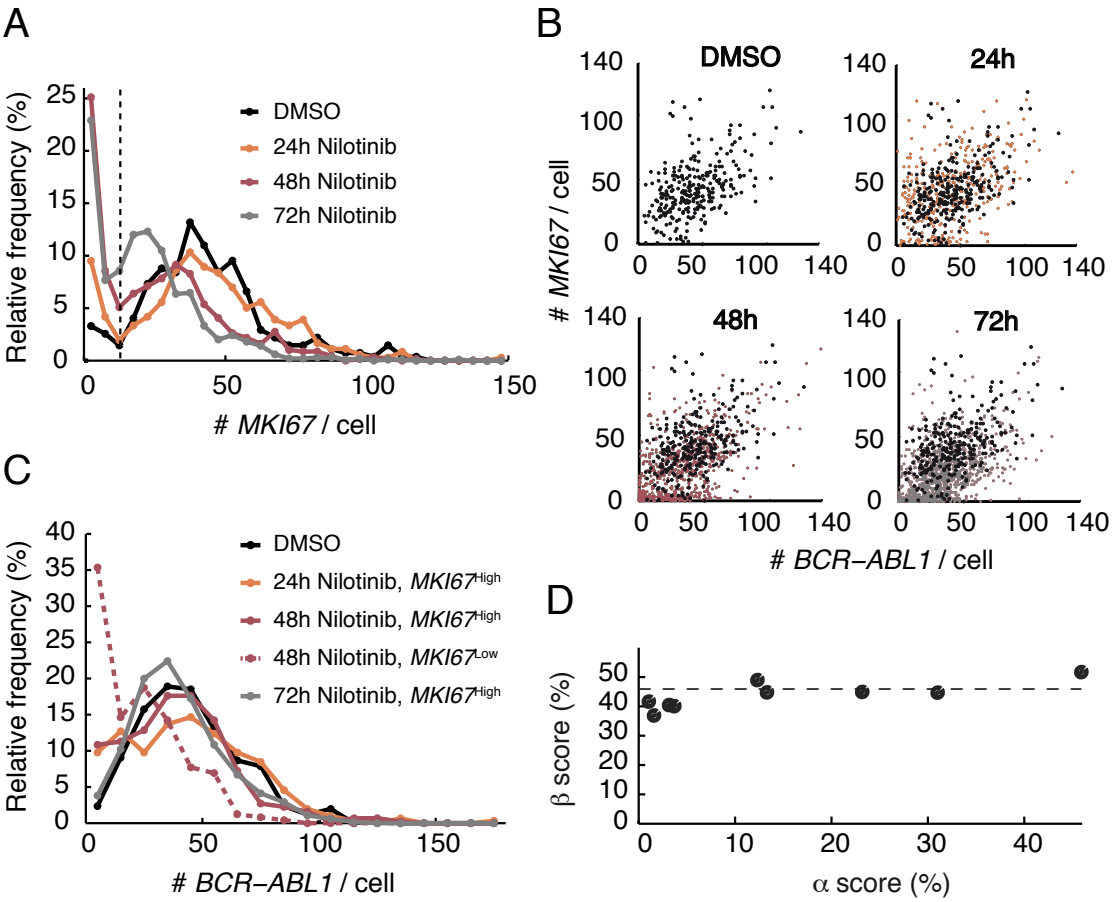
(**D**) Validation of the PICCS error model based on serial dilutions of K562 with TS cells (see **Supplemental Experimental Procedures** for details).

(**E**) Dynamic range for *BCR-ABL1* detection in purified RNA. For each cell dilution, we spotted 125 ng of total purified RNA. Mean $\pm$ s.d. is shown for each dilution.

(**F**) Comparison of FuseFISH α with the (*BCR-ABL1*)/*ABL1* RT-PCR ratio available in 21 cases.

(**G**) Systematic cross validation of the dataset in **Figures 2C** and **2D** using the algorithms described in **Experimental Procedures**.

Figure S3. FuseFISH applications, Related to Figure 3



(**A**) Distribution of *MKI67* mRNA in single cells. The dash line marks an arbitrary cutoff between $MKI67^{Low}$ (non-proliferating) and $MKI67^{High}$ (proliferating) cells.

(**B**) *MKI67* vs. *BCR-ABL1* expression in single cells treated with 30 nM Nilotinib for the indicated time (hours). DMSO: cells treated with dimethyl sulfoxide for 24 hours. For each time point, DMSO-treated cells (black) are superimposed onto Nilotinib-treated cells.

(**C**) Distribution of *BCR-ABL1* expression in single cells exposed to 30 nM Nilotinib for the indicated times (hours).

(**D**) Local correlation fraction β in serial dilutions of K562 with TS cells. Dashed line: linear fitting.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Fusion Transcript Coverage

To estimate the coverage of fusion transcripts by oligos in our database (http://www.fusefish.eu), we compiled a comprehensive list of expressed gene fusions based on the COSMIC (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/) and OMIM (http://www.omim.org) databases, and on publications reporting fusions identified by next-generation sequencing. For each unique pair of fusion transcripts, we retrieved the ENSEMBL longest protein-coding transcript (http://www.ensembl.org), and then sought to identify the shortest described variant together with the number of the exons flanking the fusion point using OMIM and COSMIC. If this information was unavailable, we searched the web with Google using a combination of keywords comprising the gene name followed by "fusion RT-PCR". 348 unique fusions for which we managed to retrieve complete information as well the number of oligos complementary to the 5′ and 3′ moiety of each fusion partner are listed in **Table S1**. We note that this coverage represents a very stringent estimate in the sense that it is based on the shortest fusion variants described in the literature, even though their frequency is low in most of cases.

Cells and Tissues

We obtained cells from ATCC unless otherwise stated. We retrieved hematologic specimens in methanol-acetic acid 3/1 vol./vol. from the archive of the Division of Hematology of IRCCS Policlinico San Matteo after approval of the study by the institutional Ethical Committee. We found that highly efficient and durable attachment of cells to both uncoated microscope slides and coverglasses can be achieved by either spotting cells in methanol-acetic acid 3/1 vol./vol. (MAA) manually or by cytospin, as well as by spotting cells suspended in an aqueous buffer by cytospin, followed by fixation with MAA. This procedure avoids time-consuming and inefficient attachment of suspended cells to coated coverglasses (e.g. with poly-L-lysine) as well as exposure to the carcinogen formaldehyde. Moreover, cells in cytogenetics laboratories are almost universally stored in MAA fixative or its variants (e.g.

Carnoy's fixative), making our method readily applicable to archive material and to samples routinely collected by diagnostics laboratories. For CML specimens, we spotted 5 µl of cell suspension in MAA on No. 1 coverglasses (VWR VistaVision™, catalogue no. 16004-098) mounted on microscope slides (VWR VistaVision™, catalogue no. 16004-422), and let the cells spontaneously dry at room temperature. We stored spotted cells dry in ambient air up to 4 months before analysis.

Permission to conduct mice experiments was granted by the Italian Ministry of Health (approval number 218-2009-B). We prepared mouse xenografts by injections in the flank (10^7 cells per mouse) or in the tail vein (10^7 cells per mouse) of NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice. We sacrificed mice when the subcutaneous tumors reached 10 mm in diameter or after 35 days for lung tumors. Tumors were excised, immediately fixed in formalin and paraffin embedded following standard procedures. Tissue sections were de-paraffinized and post-fixed, and subjected to mild tissue digestion with pepsin and auto-fluorescence quenching with $NaBH_4$.

smFISH

We prepared deparaffinized tissue sections in D-Limonene Clearant (VWR, catalogue no. 95057-818), followed by re-hydration, post-fixation in 4% formalin in 1× PBS for 10 min at room temperature, and RNA retrieval for 45 min at 80 °C in 0.01 M sodium citrate pH 6. Afterwards, we de-hydrated samples in ethanol and covered the tissue with a hybridization chamber as described below. We digested tissue with 0.025% pepsin in 0.01 M HCl, and then quenched auto-fluorescence by repeatedly flushing the chamber with $NaBH_4$ 1% in 1× PBS solution over a period of 10 min at room temperature. We then washed the tissue three times 10 min each in RNAse-free water, and stored samples in Ribonucleoside Vanadyl Complex (RVC, NEB, catalogue no. S1402S) diluted 1:20 vol./vol. in 2× SSC buffer (Ambion, catalogue no. AM9765) (SSC-RVC buffer) at 4 °C until hybridization was performed.

We covered cell spots or tissue sections mounted on coverglasses with 22×22 mm SecureSeal hybridization chambers (EMS, catalogue no. 70333-10). All subsequent steps consisted of manual injections/aspirations of approx. 100 µl of a given solution into/from the hybridization chamber. Before hybridization, we briefly re-hydrated cells by filling the chamber twice with SSC-RVC at room temperature. We quickly solubilized cellular lipids by filling the chamber with SSC-RVC supplemented with Triton X-100 0.5% and incubating 10 min at room temperature. We washed the cells twice with SSC-RVC, and equilibrated them for 5 min at room temperature with RNA Wash buffer (RWB) containing 2× SSC buffer and Formamide 25% (Ambion, catalogue no. AM9342). Afterwards, we filled the chamber with RNA Hybridization buffer (RHB) containing 2× SSC buffer, Formamide 25%, Dextran Sulphate 10% (Sigma, catalogue no. D8906), *E. coli* tRNA (Sigma, catalogue no. R4251), Bovine Serum Albumin (Ambion, catalogue no. AM2616), RVC and 40 ng of the desired probes (the mass refers only to pooled oligonucleotides, excluding fluorophores, and is based on absorbance measurements at 260 nm). We preformed hybridization for 16-18 h at 30 ˚C, after which we quickly washed cells twice in RWB, incubated them in RWB for 1 h at 30 ˚C, and finally incubated them in RWB supplemented with DAPI 20 ng/ml for 30 min at 30 ˚C. For microscopy, we filled the hybridization chamber with a mounting solution containing 2× SSC buffer, 10 mM Tris, 0.4% Glucose, 100 µg/ml Catalase, 37 µg/ml Glucose Oxidase, and 2 mM Trolox.

For smFISH in purified RNA, we extracted total RNA from cultured cell lines using the QIAshredder and RNeasy Plus Mini kits (Qiagen, catalogue no. 79654 and 74134, respectively). We prepared 1:1 vol./vol. mixes of Proteinase K 20 mg/ml (Ambion, catalogue no. AM2548) and purified RNA 500 µg/ml in nuclease-free water, and manually spotted 0.5 µl of the mix in the center of No. 1 coverglasses (VWR VistaVision™, catalogue no. 16004-098) attached with tape onto microscope slides (VWR VistaVision™, catalogue no. 16004-422). We dried RNA-protein spots for 20 min at 80 °C in a hybridization thermoblock. We covered the spots with 22×22 mm SecureSeal hybridization chambers (EMS, catalogue no.

70333-10). All subsequent steps consisted of manual injections/aspirations of approx. 100 µl of a given solution into/from the hybridization chamber. We fixed RNA-protein spots with MAA for 5 min at room temperature, and quickly washed the chamber twice with SSC-RVC at room temperature. Afterwards, we filled the chamber with SSC-RVC and 40 ng of the desired probes. We performed hybridization for 5 min at 80 ˚C, followed by gradual cooling down to 55 ˚C for 10 min. Finally, we quickly washed the chamber twice in RWB, and filled it with an imaging solution containing 2× SSC buffer, 10 mM Tris, 0.4% Glucose, 100 µg/ml Catalase, 37 µg/ml Glucose Oxidase, and 2 mM Trolox.

Karyotype analysis, RT-PCR, and DNA FISH

Karyotyping, RT-PCR, and DNA FISH assessment of the clinical specimens listed in **Table S3** were performed as part of routine cytogenetic diagnostics at the Division of Hematology of IRCCS Policlinico San Matteo. *BCR-ABL1* DNA FISH was done with the Vysis LSI BCR ABL ES dual color translocation probe (Abbot Molecular, catalogue no. 08L55-020 CE), whereas the *ipsogen* BCR-ABL1 Mbcr Controls Kit (Qiagen, catalogue no. 670191) was used for RT-PCR.

Image processing

We filtered individual images with a high-pass Fast Fourier transform filter (cutoff $= 9$ pixels), preserving diffraction-limited signals (full width at half maximum, $w = 2.3$ pixels). After maximum projection, we repeated filtering of the resulting image. To separate unspecific signals caused by autofluorescence or non-specifically bound probes from real mRNA signals, we first estimated the noise by computing the power spectrum of linearized images. We constructed a linearized image by concatenating the rows of each image to a one-dimensional vector. We then estimated the noise level, n as the mean of the power spectrum over spatial frequencies $> 1/w$. We used a multiple of the noise level n as threshold. $5n$ was sufficient to reject most sporadic signals while accepting most real signals. We identified

signals iteratively by first finding the brightest pixel in the projected image and then setting an area of size 3×3 around this pixel to zero (pixels in the 3×3 area belong to the same diffraction-limited signal). We continued this scheme until no more pixels exceeded the threshold. To determine the position and intensity of the signals, we fitted 2D Gaussians approximating the point-spread function (PSF) of the microscope. Fit parameters were x and y position, signal width w, and intensity I. We first fitted x and y using the positions of the brightest pixels determined in the previous step as initial values. We assumed $w = 2.3$ pixels and approximated signal intensity by a local integration weighted with a Gaussian of width $w = 2.3$ pixels and unit intensity. We used an iterative method (Thompson et al., 2002) to determine signal positions with sub-pixel accuracy. Subsequently, we determined intensity and width by least squares fitting keeping x and y fixed. We only considered signals with widths between 1.8 and 2.8 pixels. To exclude bright debris outside of cells, we segmented DAPI-stained nuclei, and discarded signals that were further than 20 pixels away from the edge of a nucleus. From the integrated intensity determined by least squares fitting, we calculated the maximum of the Gaussian that approximates the PSF. We only considered signals for which the ratio of this maximum and the noise (signal-to-noise ratio, SNR) exceeded a certain threshold T.

Scaling of α with the number of positive cells

Using a mixing ratio of one K562 cell in x TS cells, the expected PICCS correlation fraction is: $\alpha_{BCR-ABL1}(x) = (c_{BCR}^{pos}\alpha_{BCR-ABL1}^{pos} + xc_{BCR}^{neg}\alpha_{BCR-ABL1}^{neg}) / (c_{BCR}^{pos} + xc_{BCR}^{neg})$, where $c_{BCR}^{pos/neg}$ and $\alpha_{BCR-ABL1}^{pos/neg}$ are the density of *BCR* and *BCR-ABL1* correlation fraction in K562 (pos) and TS (neg) cells, respectively. We derived c_{BCR}^{pos}, c_{BCR}^{neg}, $\alpha_{BCR-ABL1}^{pos}$ and $\alpha_{BCR-ABL1}^{neg}$ from pure populations of K562 or TS cells, and used them to predict $\alpha_{BCR-ABL1}(x)$ for different mixing ratios.

Estimation of PICCS accuracy and calculation of *P* values

To determine the accuracy of α_{AB} we performed bootstrapping by randomly picking (with replacement) N_A out of N total signals, and performing standard PICCS analysis. We repeated this procedure 200 times, and estimated the error as the standard deviation of the obtained α_{AB} values. By bootstrapping the negative control samples, we created a null distribution for the null hypothesis that there are no fusion transcripts present. We calculated *P* values for correlation fractions based on this null distribution.

The error of the computed correlation fraction has two independent sources: 1) counting correlated signals; and 2) counting uncorrelated, randomly colocalized signals.

1) Counting correlated signals

We modeled the process of counting correlated signals of species B as a Bernoulli trial. The number of successful trials (i.e. correlated signals of species B) is governed by a binomial distribution. Given a correlation fraction α_{AB} and N_A measured signals of species A, we expect an average of $\alpha_{AB}N_A$ correlated signals of species B and variability $N_A\alpha_{AB}(1-\alpha_{AB})$. Therefore, we expect a relative error $\varepsilon_1 = \frac{1}{\sqrt{N_A}}\sqrt{\frac{1-\alpha_{AB}}{\alpha_{AB}}}$ from counting statistics alone.

Figure S2D shows the relative error estimated by bootstrapping in a dilution series of K562 with TS cells. Variability due to counting correlated signals does not sufficiently explain the observed relative error. However, the above stated expression for ε_1 is a useful lower theoretical limit for the achievable relative error.

2) Counting randomly colocalized signal

The contribution of randomly colocalized signals is determined by linear fitting of C_{cum} at length scales exceeding the typical correlation length r_c. While the error for counting the number of signals of species B (which have density c_B) in a circle with area r_c around a signal of species A scales with $\sqrt{c_B}$, the pre-factor depends on details of the fitting.

Therefore this contribution of the error is modeled as $\varepsilon_2 = \frac{k}{\alpha_{AB}}\sqrt{\frac{c_B}{N_A}}$ where k is a constant that depends on the details of the linear fitting. Since both error sources are independent, the total relative error is $\varepsilon = \sqrt{\varepsilon_1^2 + \varepsilon_2^2} = \sqrt{\frac{1}{N_A}}\sqrt{\frac{kc_B}{\alpha_{AB}^2} + \frac{(1-\alpha_{AB})}{\alpha_{AB}}}$. We empirically determined k from the dilution series of K562 with TS cells where α_{AB} is considered to be the difference to the α_{AB} measured in the negative control. After estimation of ε by bootstrapping we calculated $k = \frac{N_A\varepsilon^2\alpha_{AB}^2 - (1-\alpha_{AB})\alpha_{AB}}{c_B}$. k was approximately 30 pixels^2 for the fitting parameters used. As shown in **Figure S2D** this model describes the observed error well. To measure α_{AB} with a relative error ε the number of signals N_A has to satisfy $N_A > \frac{1}{\varepsilon^2}\left(\frac{kc_B}{\alpha_{AB}^2} + \frac{(1-\alpha_{AB})}{\alpha_{AB}}\right)$.

SUPPLEMENTAL TABLES

Supplemental Table S1, smFISH Oligo Coverage for Known Fused Transcripts, Related to Figure 1C

This table summarizes the oligo coverage for the shortest variant of known fusion transcripts with annotated information about which exons are retained in the fusion. This table can be downloaded as a single. xlsx file entitled "**Table S1.smFISH Oligo Coverage for Known Fused Transcripts, Related to Figure 1C**".

Supplemental Table S2, List of smFISH probes, Related to Figure S1A

This table contains the transcript accession number and smFISH oligo sequences used to visualize the transcripts described in this study. This table can be downloaded as a single. xlsx file entitled "**Table S2. List of smFISH probes, Related to Figure S1A**".

Supplemental Table S3, Characteristics of Hematologic Samples and Donors, Related to Figures 2C and 2D

This table contains disease type, tissue source, cytogenetic, and clinical outcome information related to the hematologic samples and donors used to obtain the data summarized in **Figures 2C** and **2D**. This table can be downloaded as a single. xlsx file entitled "**Table S3. Characteristics of Hematologic Samples and Donors, Related to Figures 2C and 2D**".

SUPPLEMENTAL REFERENCES

Thompson, R.E., Larson, D.R., Webb, W.W., 2002. Precise nanometer localization analysis for individual fluorescent probes. Biophys. J. 82, 2775–2783.

References in Table S1, Related to Table S1

Below is a detailed list of papers used for compiling **Table S1**, and cited in column "I" of **Table S1**.

Abe, A., Emi, N., Tanimoto, M., Terasaki, H., Marunouchi, T., Saito, H., 1997. Fusion of the platelet-derived growth factor receptor beta to a novel gene CEV14 in acute myelogenous leukemia after clonal evolution. Blood 90, 4271–4277.

An, Q., Wright, S.L., Konn, Z.J., Matheson, E., Minto, L., Moorman, A.V., Parker, H., Griffiths, M., Ross, F.M., Davies, T., Hall, A.G., Harrison, C.J., Irving, J.A., Strefford, J.C., 2008. Variable breakpoints target PAX5 in patients with dicentric chromosomes: a model for the basis of unbalanced translocations in cancer. Proc. Natl. Acad. Sci. U.S.A. 105, 17050–17054.

An, Q., Wright, S.L., Moorman, A.V., Parker, H., Griffiths, M., Ross, F.M., Davies, T., Harrison, C.J., Strefford, J.C., 2009. Heterogeneous breakpoints in patients with acute lymphoblastic leukemia and the dic(9;20)(p11-13;q11) show recurrent involvement of genes at 20q11.21. Haematologica 94, 1164–1169.

Antonescu, C.R., Dal Cin, P., Nafa, K., Teot, L.A., Surti, U., Fletcher, C.D., Ladanyi, M., 2007. EWSR1-CREB1 is the predominant gene fusion in angiomatoid fibrous histiocytoma. Genes Chromosomes Cancer 46, 1051–1060.

Aplan, P.D., Lombardi, D.P., Reaman, G.H., Sather, H.N., Hammond, G.D., Kirsch, I.R., 1992. Involvement of the putative hematopoietic transcription factor SCL in T-cell acute lymphoblastic leukemia. Blood 79, 1327–1333.

Arai, Y., Hosoda, F., Kobayashi, H., Arai, K., Hayashi, Y., Kamada, N., Kaneko, Y., Ohki, M., 1997. The inv(11)(p15q22) chromosome translocation of de novo and therapy-related myeloid malignancies results in fusion of the nucleoporin gene, NUP98, with the putative RNA helicase gene, DDX10. Blood 89, 3936–3944.

Argani, P., Lui, M.Y., Couturier, J., Bouvier, R., Fournet, J.-C., Ladanyi, M., 2003. A novel CLTC-TFE3 gene fusion in pediatric renal adenocarcinoma with t(X;17)(p11.2;q23). Oncogene 22, 5374–5378.

Asmann, Y.W., Hossain, A., Necela, B.M., Middha, S., Kalari, K.R., Sun, Z., Chai, H.-S., Williamson, D.W., Radisky, D., Schroth, G.P., Kocher, J.-P.A., Perez, E.A., Thompson, E.A., 2011. A novel bioinformatics pipeline for identification and characterization of fusion transcripts in breast cancer and normal cell lines. Nucleic Acids Research 39, e100.

Asp, J., Persson, F., Kost-Alimova, M., Stenman, G., 2006. CHCHD7-PLAG1 and TCEA1-PLAG1 gene fusions resulting from cryptic, intrachromosomal 8q rearrangements in pleomorphic salivary gland adenomas. Genes Chromosomes Cancer 45, 820–828.

Aström, A.K., Voz, M.L., Kas, K., Röijer, E., Wedell, B., Mandahl, N., Van de Ven, W., Mark, J., Stenman, G., 1999. Conserved mechanism of PLAG1 activation in salivary gland tumors with and without chromosome 8q12 abnormalities: identification of SII as a new fusion partner gene. Cancer Res. 59, 918–923.

Attard, G., Clark, J., Ambroisine, L., Mills, I.G., Fisher, G., Flohr, P., Reid, A., Edwards, S., Kovacs, G., Berney, D., Foster, C., Massie, C.E., Fletcher, A., De Bono, J.S., Scardino, P., Cuzick, J., Cooper, C.S., Transatlantic Prostate Group, 2008. Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. Br. J. Cancer 99, 314–320.

Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K.K., Carter, S.L., Frederick, A.M.,

Lawrence, M.S., Sivachenko, A.Y., Sougnez, C., Zou, L., Cortes, M.L., Fernandez-Lopez, J.C., Peng, S., Ardlie, K.G., Auclair, D., Bautista-Piña, V., Duke, F., Francis, J., Jung, J., Maffuz-Aziz, A., Onofrio, R.C., Parkin, M., Pho, N.H., Quintanar-Jurado, V., Ramos, A.H., Rebollar-Vega, R., Rodriguez-Cuevas, S., Romero-Cordoba, S.L., Schumacher, S.E., Stransky, N., Thompson, K.M., Uribe-Figueroa, L., Baselga, J., Beroukhim, R., Polyak, K., Sgroi, D.C., Richardson, A.L., Jimenez-Sanchez, G., Lander, E.S., Gabriel, S.B., Garraway, L.A., Golub, T.R., Melendez-Zajgla, J., Toker, A., Getz, G., Hidalgo-Miranda, A., Meyerson, M., 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 486, 405–409.

Barbouti, A., Höglund, M., Johansson, B., Lassen, C., Nilsson, P.-G., Hagemeijer, A., Mitelman, F., Fioretos, T., 2003. A novel gene, MSI2, encoding a putative RNA-binding protein is recurrently rearranged at disease progression of chronic myeloid leukemia and forms a fusion gene with HOXA9 as a result of the cryptic t(7;17)(p15;q23). Cancer Res. 63, 1202–1206.

Bizarro, S., Cerveira, N., Correia, C., Lisboa, S., Peixoto, A., Norton, L., Teixeira, M.R., 2007. Molecular characterization of a rare MLL-AF4 (MLL-AFF1) fusion rearrangement in infant leukemia. Cancer Genet. Cytogenet. 178, 61–64.

Bodmer, D., Schepens, M., Eleveld, M.J., Schoenmakers, E.F.P.M., Geurts van Kessel, A., 2003. Disruption of a novel gene, DIRC3, and expression of DIRC3-HSPBAP1 fusion transcripts in a case of familial renal cell cancer and t(2;3)(q35;q21). Genes Chromosomes Cancer 38, 107–116.

Bongarzone, I., Butti, M.G., Coronelli, S., Borrello, M.G., Santoro, M., Mondellini, P., Pilotti, S., Fusco, A., Porta, Della, G., Pierotti, M.A., 1994. Frequent activation of ret protooncogene by fusion with a new activating gene in papillary thyroid carcinomas. Cancer Res. 54, 2979–2985.

Borkhardt, A., Bojesen, S., Haas, O.A., Fuchs, U., Bartelheimer, D., Loncarevic, I.F., Bohle, R.M., Harbott, J., Repp, R., Jaeger, U., Viehmann, S., Henn, T., Korth, P., Scharr, D., Lampert, F., 2000. The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q. Proc. Natl. Acad. Sci. U.S.A. 97, 9168–9173.

Bousquet, M., Broccardo, C., Quelen, C., Meggetto, F., Kuhlein, E., Delsol, G., Dastugue, N., Brousset, P., 2007. A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5. Blood 109, 3417–3423.

Catalano, A., Dawson, M.A., Somana, K., Opat, S., Schwarer, A., Campbell, L.J., Iland, H., 2007. The PRKAR1A gene is fused to RARA in a new variant acute promyelocytic leukemia. Blood 110, 4073–4076.

Cazzaniga, G., Daniotti, M., Tosi, S., Giudici, G., Aloisi, A., Pogliani, E., Kearney, L., Biondi, A., 2001. The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. Cancer Res. 61, 4666–4670.

Chan, E.M., Comer, E.M., Brown, F.C., Richkind, K.E., Holmes, M.L., Chong, B.H., Shiffman, R., Zhang, D.-E., Slovak, M.L., Willman, C.L., Noguchi, C.T., Li, Y., Heiber, D.J., Kwan, L., Chan, R.J., Vance, G.H., Ramsey, H.C., Hromas, R.A., 2005. AML1-FOG2 fusion protein in myelodysplasia. Blood 105, 4523–4526.

Chase, A., Reiter, A., Burci, L., Cazzaniga, G., Biondi, A., Pickard, J., Roberts, I.A., Goldman, J.M., Cross, N.C., 1999. Fusion of ETV6 to the caudal-related homeobox gene CDX2 in acute myeloid leukemia with the t(12;13)(p13;q12). Blood 93, 1025–1031.

Ciampi, R., Giordano, T.J., Wikenheiser-Brokamp, K., Koenig, R.J., Nikiforov, Y.E., 2007. HOOK3-RET: a novel type of RET/PTC rearrangement in papillary thyroid carcinoma. Endocr. Relat. Cancer 14, 445–452.

Coenen, E.A., Zwaan, C.M., Meyer, C., Marschalek, R., Pieters, R., van der Veken, L.T., Beverloo, H.B., van den Heuvel-Eibrink, M.M., 2011. KIAA1524: A novel MLL translocation partner in acute myeloid leukemia. Leuk. Res. 35, 133–135.

Cools, J., DeAngelo, D.J., Gotlib, J., Stover, E.H., Legare, R.D., Cortes, J., Kutok, J., Clark, J., Galinsky, I., Griffin, J.D., Cross, N.C.P., Tefferi, A., Malone, J., Alam, R., Schrier,

S.L., Schmid, J., Rose, M., Vandenberghe, P., Verhoef, G., Boogaerts, M., Wlodarska, I., Kantarjian, H., Marynen, P., Coutre, S.E., Stone, R., Gilliland, D.G., 2003. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. N. Engl. J. Med. 348, 1201–1214.

Corral, J., Forster, A., Thompson, S., Lampert, F., Kaneko, Y., Slater, R., Kroes, W.G., van der Schoot, C.E., Ludwig, W.-D., Karpas, A., 1993. Acute leukemias of different lineages have similar MLL gene fusions encoding related chimeric proteins resulting from chromosomal translocation. Proc. Natl. Acad. Sci. U.S.A. 90, 8538–8542.

Coyaud, E., Struski, S., Prade, N., Familiades, J., Eichner, R., Quelen, C., Bousquet, M., Mugneret, F., Talmant, P., Pages, M.-P., Lefebvre, C., Penther, D., Lippert, E., Nadal, N., Taviaux, S., Poppe, B., Luquet, I., Baranger, L., Eclache, V., Radford, I., Barin, C., Mozziconacci, M.-J., Lafage-Pochitaloff, M., Antoine-Poirel, H., Charrin, C., Perot, C., Terre, C., Brousset, P., Dastugue, N., Broccardo, C., 2010. Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. Blood 115, 3089–3097.

Dean, M., Park, M., Vande Woude, G.F., 1987. Characterization of the rearranged tpr-met oncogene breakpoint. Mol. Cell. Biol. 7, 921–924.

Denk, D., Nebral, K., Bradtke, J., Pass, G., Möricke, A., Attarbaschi, A., Strehl, S., 2012. PAX5-AUTS2: a recurrent fusion gene in childhood B-cell precursor acute lymphoblastic leukemia. Leuk. Res. 36, e178–81.

Dreyling, M.H., Martinez-Climent, J.A., Zheng, M., Mao, J., Rowley, J.D., Bohlander, S.K., 1996. The t(10;11)(p13;q14) in the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family. Proc. Natl. Acad. Sci. U.S.A. 93, 4804–4809.

Edgren, H., Murumagi, A., Kangaspeska, S., Nicorici, D., Hongisto, V., Kleivi, K., Rye, I.H., Nyberg, S., Wolf, M., Borresen-Dale, A.-L., Kallioniemi, O., 2011. Identification of fusion genes in breast cancer by paired-end RNA-sequencing. Genome Biol. 12, R6.

Errani, C., Zhang, L., Sung, Y.S., Hajdu, M., Singer, S., Maki, R.G., Healey, J.H., Antonescu, C.R., 2011. A novel WWTR1-CAMTA1 gene fusion is a consistent abnormality in epithelioid hemangioendothelioma of different anatomic sites. Genes Chromosomes Cancer 50, 644–653.

Fehr, A., Röser, K., Heidorn, K., Hallas, C., Löning, T., Bullerdiek, J., 2008. A new type of MAML2 fusion in mucoepidermoid carcinoma. Genes Chromosomes Cancer 47, 203–206.

Francis, R.W., Thompson-Wicking, K., Carter, K.W., Anderson, D., Kees, U.R., Beesley, A.H., 2012. FusionFinder: a software tool to identify expressed gene fusion candidates from RNA-Seq data. PLoS ONE 7, e39987.

French, C.A., Miyoshi, I., Kubonishi, I., Grier, H.E., Perez-Atayde, A.R., Fletcher, J.A., 2003. BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. Cancer Res. 63, 304–307.

Fu, J.-F., Hsu, J.-J., Tang, T.-C., Shih, L.-Y., 2003. Identification of CBL, a proto-oncogene at 11q23.3, as a novel MLL fusion partner in a patient with de novo acute myeloid leukemia. Genes Chromosomes Cancer 37, 214–219.

Fuchs, U., Rehkamp, G., Haas, O.A., Slany, R., Kōnig, M., Bojesen, S., Bohle, R.M., Damm-Welk, C., Ludwig, W.-D., Harbott, J., Borkhardt, A., 2001. The human formin-binding protein 17 (FBP17) interacts with sorting nexin, SNX2, and is an MLL-fusion partner in acute myelogeneous leukemia. Proc. Natl. Acad. Sci. U.S.A. 98, 8756–8761.

Fujino, T., Suzuki, A., Ito, Y., Ohyashiki, K., Hatano, Y., Miura, I., Nakamura, T., 2002. Single-translocation and double-chimeric transcripts: detection of NUP98-HOXA9 in myeloid leukemias with HOXA11 or HOXA13 breaks of the chromosomal translocation t(7;11)(p15;p15). Blood 99, 1428–1433.

Galiègue-Zouitina, S., Quief, S., Hildebrand, M.P., Denis, C., Detourmignies, L., Laï, J.L., Kerckaert, J.P., 1999. Nonrandom fusion of L-plastin(LCP1) and LAZ3(BCL6) genes by t(3;13)(q27;q14) chromosome translocation in two cases of B-cell non-Hodgkin lymphoma. Genes Chromosomes Cancer 26, 97–105.

Gamou, T., Kitamura, E., Hosoda, F., Shimizu, K., Shinohara, K., Hayashi, Y., Nagase, T., Yokoyama, Y., Ohki, M., 1998. The partner gene of AML1 in t(16;21) myeloid malignancies is a novel member of the MTG8(ETO) family. Blood 91, 4028–4037.
Ge, H., Liu, K., Juan, T., Fang, F., Newman, M., Hoeck, W., 2011. FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. Bioinformatics 27, 1922–1928.
Geurts, J.M., Schoenmakers, E.F., Röijer, E., Stenman, G., Van de Ven, W.J., 1997. Expression of reciprocal hybrid transcripts of HMGIC and FHIT in a pleomorphic adenoma of the parotid gland. Cancer Res. 57, 13–17.
Grand, E.K., Grand, F.H., Chase, A.J., Ross, F.M., Corcoran, M.M., Oscier, D.G., Cross, N.C.P., 2004. Identification of a novel gene, FGFR1OP2, fused to FGFR1 in 8p11 myeloproliferative syndrome. Genes Chromosomes Cancer 40, 78–83.
Guasch, G., Mack, G.J., Popovici, C., Dastugue, N., Birnbaum, D., Rattner, J.B., Pébusque, M.J., 2000. FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33). Blood 95, 1788–1796.
Hallor, K.H., Mertens, F., Jin, Y., Meis-Kindblom, J.M., Kindblom, L.-G., Behrendtz, M., Kalén, A., Mandahl, N., Panagopoulos, I., 2005. Fusion of the EWSR1 and ATF1 genes without expression of the MITF-M transcript in angiomatoid fibrous histiocytoma. Genes Chromosomes Cancer 44, 97–102.
Han, B., Mehra, R., Dhanasekaran, S.M., Yu, J., Menon, A., Lonigro, R.J., Wang, X., Gong, Y., Wang, L., Shankar, S., Laxman, B., Shah, R.B., Varambally, S., Palanisamy, N., Tomlins, S.A., Kumar-Sinha, C., Chinnaiyan, A.M., 2008. A fluorescence in situ hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. Cancer Res. 68, 7629–7637.
Hermans, K.G., Bressers, A.A., van der Korput, H.A., Dits, N.F., Jenster, G., Trapman, J., 2008. Two unique novel prostate-specific and androgen-regulated fusion partners of ETV4 in prostate cancer. Cancer Res. 68, 3094–3098.
Hosokawa, Y., Maeda, Y., Ichinohasama, R., Miura, I., Taniwaki, M., Seto, M., 2000. The Ikaros gene, a central regulator of lymphoid differentiation, fuses to the BCL6 gene as a result of t(3;7)(q27;p12) translocation in a patient with diffuse large B-cell lymphoma. Blood 95, 2719–2721.
Hunger, S.P., Ohyashiki, K., Toyama, K., Cleary, M.L., 1992. Hlf, a novel hepatic bZIP protein, shows altered DNA-binding properties following fusion to E2A in t(17;19) acute lymphoblastic leukemia. Genes Dev. 6, 1608–1620.
Hussey, D.J., Nicola, M., Moore, S., Peters, G.B., Dobrovic, A., 1999. The (4;11)(q21;p15) translocation fuses the NUP98 and RAP1GDS1 genes and is recurrent in T-cell acute lymphocytic leukemia. Blood 94, 2072–2079.
Impera, L., Albano, F., Cunsolo, Lo, C., Funes, S., Iuzzolino, P., Laveder, F., Panagopoulos, I., Rocchi, M., Storlazzi, C.T., 2008. A novel fusion 5“AFF3/3”BCL2 originated from a t(2;18)(q11.2;q21.33) translocation in follicular lymphoma. Oncogene 27, 6187–6190.
Jaju, R.J., Fidler, C., Haas, O.A., Strickson, A.J., Watkins, F., Clark, K., Cross, N.C., Cheng, J.F., Aplan, P.D., Kearney, L., Boultwood, J., Wainscoat, J.S., 2001. A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. Blood 98, 1264–1267.
Janssen, J.W., Schleithoff, L., Bartram, C.R., Schulz, A.S., 1998. An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subunit and HUMORF8, a putative deubiquitinating enzyme. Oncogene 16, 1767–1772.
Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet.
Karathanasis, S.K., Ferris, E., Haddad, I.A., 1987. DNA inversion within the apolipoproteins AI/CIII/AIV-encoding gene cluster of certain patients with premature atherosclerosis. Proc. Natl. Acad. Sci. U.S.A. 84, 7198–7202.
Kawamata, N., Ogawa, S., Zimmermann, M., Niebuhr, B., Stocking, C., Sanada, M., Hemminki, K., Yamatomo, G., Nannya, Y., Koehler, R., Flohr, T., Miller, C.W., Harbott, J., Ludwig, W.-D., Stanulla, M., Schrappe, M., Bartram, C.R., Koeffler, H.P., 2008.

Cloning of genes involved in chromosomal translocations by high-resolution single nucleotide polymorphism genomic microarray. Proc. Natl. Acad. Sci. U.S.A. 105, 11921–11926.

Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., Gotoh, T., Motoi, T., Fukayama, M., Aburatani, H., Takizawa, T., Nakamura, T., 2006. Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. Hum. Mol. Genet. 15, 2125–2137.

Kazmierczak, B., Hennig, Y., Wanschura, S., Rogalla, P., Bartnitzke, S., Van de Ven, W., Bullerdiek, J., 1995. Description of a novel fusion transcript between HMGI-C, a gene encoding for a member of the high mobility group proteins, and the mitochondrial aldehyde dehydrogenase gene. Cancer Res. 55, 6038–6039.

Kim, D., Salzberg, S.L., 2011. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. 12, R72.

Kourlas, P.J., Strout, M.P., Becknell, B., Veronese, M.L., Croce, C.M., Theil, K.S., Krahe, R., Ruutu, T., Knuutila, S., Bloomfield, C.D., Caligiuri, M.A., 2000. Identification of a gene at 11q23 encoding a guanine nucleotide exchange factor: evidence for its fusion with MLL in acute myeloid leukemia. Proc. Natl. Acad. Sci. U.S.A. 97, 2145–2150.

Kuefer, M.U., Chinwalla, V., Zeleznik-Le, N.J., Behm, F.G., Naeve, C.W., Rakestraw, K.M., Mukatira, S.T., Raimondi, S.C., Morris, S.W., 2003. Characterization of the MLL partner gene AF15q14 involved in t(11;15)(q23;q14). Oncogene 22, 1418–1424.

Kulkarni, S., Heath, C., Parker, S., Chase, A., Iqbal, S., Pocock, C.F., Kaeda, J., Cwynarski, K., Goldman, J.M., Cross, N.C., 2000. Fusion of H4/D10S170 to the platelet-derived growth factor receptor beta in BCR-ABL-negative myeloproliferative disorders with a t(5;10)(q33;q21). Cancer Res. 60, 3592–3598.

Kurose, K., Mine, N., Doi, D., Ota, Y., Yoneyama, K., Konishi, H., Araki, T., Emi, M., 2000. Novel gene fusion of COX6C at 8q22-23 to HMGIC at 12q15 in a uterine leiomyoma. Genes Chromosomes Cancer 27, 303–307.

Lahortiga, I., Vizmanos, J.L., Agirre, X., Vázquez, I., Cigudosa, J.C., Larrayoz, M.J., Sala, F., Gorosquieta, A., Perez-Equiza, K., Calasanz, M.J., Odero, M.D., 2003. NUP98 is fused to adducin 3 in a patient with T-cell acute lymphoblastic leukemia and myeloid markers, with a new translocation t(10;11)(q25;p15). Cancer Res. 63, 3079–3083.

Lee, C.-H., Ou, W.-B., Mariño-Enriquez, A., Zhu, M., Mayeda, M., Wang, Y., Guo, X., Brunner, A.L., Amant, F., French, C.A., West, R.B., McAlpine, J.N., Gilks, C.B., Yaffe, M.B., Prentice, L.M., McPherson, A., Jones, S.J.M., Marra, M.A., Shah, S.P., van de Rijn, M., Huntsman, D.G., Dal Cin, P., Debiec-Rychter, M., Nucci, M.R., Fletcher, J.A., 2012. 14-3-3 fusion oncogenes in high-grade endometrial stromal sarcoma. Proc. Natl. Acad. Sci. U.S.A. 109, 929–934.

Lindern, von, M., Fornerod, M., van Baal, S., Jaegle, M., de Wit, T., Buijs, A., Grosveld, G., 1992. The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, results in the fusion of two genes, dek and can, and the expression of a chimeric, leukemia-specific dek-can mRNA. Mol. Cell. Biol. 12, 1687–1697.

Magnusson, M.K., Meade, K.E., Brown, K.E., Arthur, D.C., Krueger, L.A., Barrett, A.J., Dunbar, C.E., 2001. Rabaptin-5 is a novel fusion partner to platelet-derived growth factor beta receptor in chronic myelomonocytic leukemia. Blood 98, 2518–2525.

Maher, C.A., Palanisamy, N., Brenner, J.C., Cao, X., Kalyana-Sundaram, S., Luo, S., Khrebtukova, I., Barrette, T.R., Grasso, C., Yu, J., Lonigro, R.J., Schroth, G., Kumar-Sinha, C., Chinnaiyan, A.M., 2009. Chimeric transcript discovery by paired-end transcriptome sequencing. Proc. Natl. Acad. Sci. U.S.A. 106, 12353–12358.

Megonigal, M.D., Cheung, N.K., Rappaport, E.F., Nowell, P.C., Wilson, R.B., Jones, D.H., Addya, K., Leonard, D.G., Kushner, B.H., Williams, T.M., Lange, B.J., Felix, C.A., 2000a. Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. Proc. Natl. Acad. Sci. U.S.A. 97, 2814–2819.

Megonigal, M.D., Rappaport, E.F., Jones, D.H., Williams, T.M., Lovett, B.D., Kelly, K.M., Lerou, P.H., Moulton, T., Budarf, M.L., Felix, C.A., 1998. t(11;22)(q23;q11.2) In acute

myeloid leukemia of infant twins fuses MLL with hCDCrel, a cell division cycle gene in the genomic region of deletion in DiGeorge and velocardiofacial syndromes. Proc. Natl. Acad. Sci. U.S.A. 95, 6413–6418.

Megonigal, M.D., Rappaport, E.F., Wilson, R.B., Jones, D.H., Whitlock, J.A., Ortega, J.A., Slater, D.J., Nowell, P.C., Felix, C.A., 2000b. Panhandle PCR for cDNA: a rapid method for isolation of MLL fusion transcripts involving unknown partner genes. Proc. Natl. Acad. Sci. U.S.A. 97, 9597–9602.

Mercher, T., Coniat, M.B., Monni, R., Mauchauffe, M., Nguyen Khac, F., Gressin, L., Mugneret, F., Leblanc, T., Dastugue, N., Berger, R., Bernard, O.A., 2001. Involvement of a human gene related to the Drosophila spen gene in the recurrent t(1;22) translocation of acute megakaryocytic leukemia. Proc. Natl. Acad. Sci. U.S.A. 98, 5776–5779.

Modena, P., Testi, M.A., Facchinetti, F., Mezzanzanica, D., Radice, M.T., Pilotti, S., Sozzi, G., 2003. UQCRH gene encoding mitochondrial Hinge protein is interrupted by a translocation in a soft-tissue sarcoma and epigenetically inactivated in some cancer cell lines. Oncogene 22, 4586–4593.

Moore, S.D.P., Offor, O., Ferry, J.A., Amrein, P.C., Morton, C.C., Dal Cin, P., 2006. ELF4 is fused to ERG in a case of acute myeloid leukemia with a t(X;21)(q25-26;q22). Leuk. Res. 30, 1037–1042.

Morerio, C., Acquila, M., Rosanda, C., Rapella, A., Dufour, C., Locatelli, F., Maserati, E., Pasquali, F., Panarello, C., 2004. HCMOGT-1 is a novel fusion partner to PDGFRB in juvenile myelomonocytic leukemia with t(5;17)(q33;p11.2). Cancer Res. 64, 2649–2651.

Mullighan, C.G., Collins-Underwood, J.R., Phillips, L.A.A., Loudin, M.G., Liu, W., Zhang, J., Ma, J., Coustan-Smith, E., Harvey, R.C., Willman, C.L., Mikhail, F.M., Meyer, J., Carroll, A.J., Williams, R.T., Cheng, J., Heerema, N.A., Basso, G., Pession, A., Pui, C.-H., Raimondi, S.C., Hunger, S.P., Downing, J.R., Carroll, W.L., Rabin, K.R., 2009. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. Nat. Genet. 41, 1243–1246.

Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.-H., Relling, M.V., Evans, W.E., Shurtleff, S.A., Downing, J.R., 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 446, 758–764.

Nakamura, T., Alder, H., Gu, Y., Prasad, R., Canaani, O., Kamada, N., Gale, R.P., Lange, B., Crist, W.M., Nowell, P.C., 1993. Genes on chromosomes 4, 9, and 19 involved in 11q23 abnormalities in acute leukemia share sequence homology and/or common motifs. Proc. Natl. Acad. Sci. U.S.A. 90, 4631–4635.

Nakamura, T., Largaespada, D.A., Lee, M.P., Johnson, L.A., Ohyashiki, K., Toyama, K., Chen, S.J., Willman, C.L., Chen, I.M., Feinberg, A.P., Jenkins, N.A., Copeland, N.G., Shaughnessy, J.D., 1996. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat. Genet. 12, 154–158.

Nakamura, T., Yamazaki, Y., Hatano, Y., Miura, I., 1999. NUP98 is fused to PMX1 homeobox gene in human acute myelogenous leukemia with chromosome translocation t(1;11)(q23;p15). Blood 94, 741–747.

Nakanishi, G., Lin, S.-N., Asagoe, K., Suzuki, N., Matsuo, A., Tanaka, R., Makino, E., Fujimoto, W., Iwatsuki, K., 2007. A novel fusion gene of collagen type I alpha 1 (exon 31) and platelet-derived growth factor B-chain (exon 2) in dermatofibrosarcoma protuberans. Eur J Dermatol 17, 217–219.

Nakata, T., Kitamura, Y., Shimizu, K., Tanaka, S., Fujimori, M., Yokoyama, S., Ito, K., Emi, M., 1999. Fusion of a novel gene, ELKS, to RET due to translocation t(10;12)(q11;p13) in a papillary thyroid carcinoma. Genes Chromosomes Cancer 25, 97–103.

Nebral, K., Denk, D., Attarbaschi, A., Kōnig, M., Mann, G., Haas, O.A., Strehl, S., 2009. Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. Leukemia 23, 134–143.

Nebral, K., König, M., Harder, L., Siebert, R., Haas, O.A., Strehl, S., 2007. Identification of PML as novel PAX5 fusion partner in childhood acute lymphoblastic leukaemia. Br. J.

Haematol. 139, 269–274.
Nucifora, G., Begy, C.R., Erickson, P., Drabkin, H.A., Rowley, J.D., 1993. The 3;21 translocation in myelodysplasia results in a fusion transcript between the AML1 gene and the gene for EAP, a highly conserved protein associated with the Epstein-Barr virus small RNA EBER 1. Proc. Natl. Acad. Sci. U.S.A. 90, 7784–7788.
Ono, R., Taki, T., Taketani, T., Taniwaki, M., Kobayashi, H., Hayashi, Y., 2002. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res. 62, 4075–4080.
Osaka, M., Rowley, J.D., Zeleznik-Le, N.J., 1999. MSF (MLL septin-like fusion), a fusion partner gene of MLL, in a therapy-related acute myeloid leukemia with a t(11;17)(q23;q25). Proc. Natl. Acad. Sci. U.S.A. 96, 6428–6433.
Panagopoulos, I., Fioretos, T., Isaksson, M., Samuelsson, U., Billström, R., Strömbeck, B., Mitelman, F., Johansson, B., 2001. Fusion of the MORF and CBP genes in acute myeloid leukemia with the t(10;16)(q22;p13). Hum. Mol. Genet. 10, 395–404.
Panagopoulos, I., Isaksson, M., Billström, R., Strömbeck, B., Mitelman, F., Johansson, B., 2003a. Fusion of the NUP98 gene and the homeobox gene HOXC13 in acute myeloid leukemia with t(11;12)(p15;q13). Genes Chromosomes Cancer 36, 107–112.
Panagopoulos, I., Isaksson, M., Lindvall, C., Hagemeijer, A., Mitelman, F., Johansson, B., 2003b. Genomic characterization of MOZ/CBP and CBP/MOZ chimeras in acute myeloid leukemia suggests the involvement of a damage-repair mechanism in the origin of the t(8;16)(p11;p13). Genes Chromosomes Cancer 36, 90–98.
Panagopoulos, I., Kerndrup, G., Carlsen, N., Strömbeck, B., Isaksson, M., Johansson, B., 2007. Fusion of NUP98 and the SET binding protein 1 (SETBP1) gene in a paediatric acute T cell lymphoblastic leukaemia with t(11;18)(p15;q12). Br. J. Haematol. 136, 294–296.
Panagopoulos, I., Storlazzi, C.T., Fletcher, C.D.M., Fletcher, J.A., Nascimento, A., Domanski, H.A., Wejde, J., Brosjö, O., Rydholm, A., Isaksson, M., Mandahl, N., Mertens, F., 2004. The chimeric FUS/CREB3l2 gene is specific for low-grade fibromyxoid sarcoma. Genes Chromosomes Cancer 40, 218–228.
Pegram, L.D., Megonigal, M.D., Lange, B.J., Nowell, P.C., Rowley, J.D., Rappaport, E.F., Felix, C.A., 2000. t(3;11) translocation in treatment-related acute myeloid leukemia fuses MLL with the GMPS (GUANOSINE 5' MONOPHOSPHATE SYNTHETASE) gene. Blood 96, 4360–4362.
Pflueger, D., Terry, S., Sboner, A., Habegger, L., Esgueva, R., Lin, P.-C., Svensson, M.A., Kitabayashi, N., Moss, B.J., MacDonald, T.Y., Cao, X., Barrette, T., Tewari, A.K., Chee, M.S., Chinnaiyan, A.M., Rickman, D.S., Demichelis, F., Gerstein, M.B., Rubin, M.A., 2011. Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. Genome Res. 21, 56–67.
Pierron, G., Tirode, F., Lucchesi, C., Reynaud, S., Ballet, S., Cohen-Gogo, S., Perrin, V., Coindre, J.-M., Delattre, O., 2012. A new subtype of bone sarcoma defined by BCOR-CCNB3 gene fusion. Nat. Genet. 44, 461–466.
Prasad, R., Gu, Y., Alder, H., Nakamura, T., Canaani, O., Saito, H., Huebner, K., Gale, R.P., Nowell, P.C., Kuriyama, K., 1993. Cloning of the ALL-1 fusion partner, the AF-6 gene, involved in acute myeloid leukemias with the t(6;11) chromosome translocation. Cancer Res. 53, 5624–5628.
Preudhomme, C., Roumier, C., Hildebrand, M.P., Dallery-Prudhomme, E., Lantoine, D., Laï, J.L., Daudignon, A., Adenis, C., Bauters, F., Fenaux, P., Kerckaert, J.P., Galiègue-Zouitina, S., 2000. Nonrandom 4p13 rearrangements of the RhoH/TTF gene, encoding a GTP-binding protein, in non-Hodgkin's lymphoma and multiple myeloma. Oncogene 19, 2023–2032.
Przybylski, G.K., Dik, W.A., Wanzeck, J., Grabarczyk, P., Majunke, S., Martin-Subero, J.I., Siebert, R., Dölken, G., Ludwig, W.-D., Verhaaf, B., van Dongen, J.J.M., Schmidt, C.A., Langerak, A.W., 2005. Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-TRDC fusion transcripts and is associated with the

absence of wild-type BCL11B transcripts in T-ALL. Leukemia 19, 201–208.
Qiao, Y., Ogawa, S., Hangaishi, A., Yuji, K., Izutsu, K., Kunisato, A., Imai, Y., Wang, L., Hosoya, N., Nannya, Y., Sato, Y., Maki, K., Mitani, K., Hirai, H., 2003. Identification of a novel fusion gene, TTL, fused to ETV6 in acute lymphoblastic leukemia with t(12;13)(p13;q14), and its implication in leukemogenesis. Leukemia 17, 1112–1120.
Quentmeier, H., Schneider, B., Röhrs, S., Romani, J., Zaborski, M., Macleod, R.A.F., Drexler, H.G., 2009. SET-NUP214 fusion in acute myeloid leukemia- and T-cell acute lymphoblastic leukemia-derived cell lines. J Hematol Oncol 2, 3.
Raffini, L.J., Slater, D.J., Rappaport, E.F., Nigro, Lo, L., Cheung, N.-K.V., Biegel, J.A., Nowell, P.C., Lange, B.J., Felix, C.A., 2002. Panhandle and reverse-panhandle PCR enable cloning of der(11) and der(other) genomic breakpoint junctions of MLL translocations and identify complex translocation of MLL, AF-4, and CDK6. Proc. Natl. Acad. Sci. U.S.A. 99, 4568–4573.
Raza-Egilmez, S.Z., Jani-Sait, S.N., Grossi, M., Higgins, M.J., Shows, T.B., Aplan, P.D., 1998. NUP98-HOXD13 gene fusion in therapy-related acute myelogenous leukemia. Cancer Res. 58, 4269–4273.
Reader, J.C., Meekins, J.S., Gojo, I., Ning, Y., 2007. A novel NUP98-PHF23 fusion resulting from a cryptic translocation t(11;17)(p15;p13) in acute myeloid leukemia. Leukemia 21, 842–844.
Rickman, D.S., Pflueger, D., Moss, B., VanDoren, V.E., Chen, C.X., la Taille, de, A., Kuefer, R., Tewari, A.K., Setlur, S.R., Demichelis, F., Rubin, M.A., 2009. SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. Cancer Res. 69, 2734–2738.
Rosati, R., La Starza, R., Veronese, A., Aventin, A., Schwienbacher, C., Vallespi, T., Negrini, M., Martelli, M.F., Mecucci, C., 2002. NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). Blood 99, 3857–3860.
Rosebeck, S., Madden, L., Jin, X., Gu, S., Apel, I.J., Appert, A., Hamoudi, R.A., Noels, H., Sagaert, X., Van Loo, P., Baens, M., Du, M.-Q., Lucas, P.C., McAllister-Lucas, L.M., 2011. Cleavage of NIK by the API2-MALT1 fusion oncoprotein leads to noncanonical NF-kappaB activation. Science 331, 468–472.
Ross, T.S., Bernard, O.A., Berger, R., Gilliland, D.G., 1998. Fusion of Huntingtin interacting protein 1 to platelet-derived growth factor beta receptor (PDGFbetaR) in chronic myelomonocytic leukemia with t(5;7)(q33;q11.2). Blood 91, 4419–4426.
Sagawa, M., Shimizu, T., Shimizu, T., Awaya, N., Mitsuhashi, T., Ikeda, Y., Okamoto, S., Kizaki, M., 2006. Establishment of a new human acute monocytic leukemia cell line TZ-1 with t(1;11)(p32;q23) and fusion gene MLL-EPS15. Leukemia 20, 1566–1571.
Salassidis, K., Bruch, J., Zitzelsberger, H., Lengfelder, E., Kellerer, A.M., Bauchinger, M., 2000. Translocation t(10;14)(q11.2:q22.1) fusing the kinetin to the RET gene creates a novel rearranged form (PTC8) of the RET proto-oncogene in radiation-induced childhood papillary thyroid carcinoma. Cancer Res. 60, 2786–2789.
Salomon-Nguyen, F., Della-Valle, V., Mauchauffe, M., Busson-Le Coniat, M., Ghysdael, J., Berger, R., Bernard, O.A., 2000. The t(1;12)(q21;p13) translocation of human acute myeloblastic leukemia results in a TEL-ARNT fusion. Proc. Natl. Acad. Sci. U.S.A. 97, 6757–6762.
Sasaki, T., Rodig, S.J., Chirieac, L.R., Janne, P.A., 2010. The biology and treatment of EML4-ALK non-small cell lung cancer. Eur. J. Cancer 46, 1773–1780.
Schoenmakers, E.F., Huysmans, C., Van de Ven, W.J., 1999. Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas. Cancer Res. 59, 19–23.
Shozu, M., Sebastian, S., Takayama, K., Hsu, W.-T., Schultz, R.A., Neely, K., Bryant, M., Bulun, S.E., 2003. Estrogen excess associated with novel gain-of-function mutations affecting the aromatase gene. N. Engl. J. Med. 348, 1855–1865.
Singh, B.N., Ansari, A., Hampsey, M., 2009. Detection of gene loops by 3C in yeast. Methods 48, 361–367.
Sjögren, H., Meis-Kindblom, J., Kindblom, L.G., Aman, P., Stenman, G., 1999. Fusion of the

EWS-related gene TAF2N to TEC in extraskeletal myxoid chondrosarcoma. Cancer Res. 59, 5064–5067.
Smedley, D., Hamoudi, R., Clark, J., Warren, W., Abdul-Rauf, M., Somers, G., Venter, D., Fagan, K., Cooper, C., Shipley, J., 1998. The t(8;13)(p11;q11-12) rearrangement associated with an atypical myeloproliferative disorder fuses the fibroblast growth factor receptor 1 gene to a novel gene RAMP. Hum. Mol. Genet. 7, 637–642.
So, C.W., Caldas, C., Liu, M.M., Chen, S.J., Huang, Q.H., Gu, L.J., Sham, M.H., Wiedemann, L.M., Chan, L.C., 1997. EEN encodes for a member of a new family of proteins containing an Src homology 3 domain and is the third gene located on chromosome 19p13 that fuses to MLL in human leukemia. Proc. Natl. Acad. Sci. U.S.A. 94, 2563–2568.
Stevens-Kroef, M.-J.P.L., Schoenmakers, E.F.P.M., van Kraaij, M., Huys, E., Vermeulen, S., van der Reijden, B., van Kessel, A.G., 2006. Identification of truncated RUNX1 and RUNX1-PRDM16 fusion transcripts in a case of t(1;21)(p36;q22)-positive therapy-related AML. Leukemia 20, 1187–1189.
Strehl, S., Borkhardt, A., Slany, R., Fuchs, U.E., König, M., Haas, O.A., 2003. The human LASP1 gene is fused to MLL in an acute myeloid leukemia with t(11;17)(q23;q21). Oncogene 22, 157–160.
Strehl, S., Kōnig, M., Boztug, H., Cooper, B.W., Suzukawa, K., Zhang, S.-J., Chen, H.-Y., Attarbaschi, A., Dworzak, M.N., 2013. All-trans retinoic acid and arsenic trioxide resistance of acute promyelocytic leukemia with the variant STAT5B-RARA fusion gene. Leukemia 27, 1606–1610.
Such, E., Cervera, J., Valencia, A., Barragán, E., Ibañez, M., Luna, I., Fuster, O., Perez-Sirvent, M.L., Senent, L., Sempere, A., Martinez, J., Martín-Aragonés, G., Sanz, M.A., 2011. A novel NUP98/RARG gene fusion in acute myeloid leukemia resembling acute promyelocytic leukemia. Blood 117, 242–245.
Suzuki, A., Ito, Y., Sashida, G., Honda, S., Katagiri, T., Fujino, T., Nakamura, T., Ohyashiki, K., 2002. t(7;11)(p15;p15) Chronic myeloid leukaemia developed into blastic transformation showing a novel NUP98/HOXA11 fusion. Br. J. Haematol. 116, 170–172.
Tagawa, H., Miura, I., Suzuki, R., Suzuki, H., Hosokawa, Y., Seto, M., 2002. Molecular cytogenetic analysis of the breakpoint region at 6q21-22 in T-cell lymphoma/leukemia cell lines. Genes Chromosomes Cancer 34, 175–185.
Taketani, T., Taki, T., Sako, M., Ishii, T., Yamaguchi, S., Hayashi, Y., 2008. MNX1-ETV6 fusion gene in an acute megakaryoblastic leukemia and expression of the MNX1 gene in leukemia and normal B cell lines. Cancer Genet. Cytogenet. 186, 115–119.
Taketani, T., Taki, T., Shibuya, N., Kikuchi, A., Hanada, R., Hayashi, Y., 2002. Novel NUP98-HOXC11 fusion gene resulted from a chromosomal break within exon 1 of HOXC11 in acute myeloid leukemia with t(11;12)(p15;q13). Cancer Res. 62, 4571–4574.
Tanabe, S., Bohlander, S.K., Vignon, C.V., Espinosa, R., Zhao, N., Strissel, P.L., Zeleznik-Le, N.J., Rowley, J.D., 1996. AF10 is split by MLL and HEAB, a human homolog to a putative Caenorhabditis elegans ATP/GTP-binding protein in an invins(10;11)(p12;q23q12). Blood 88, 3535–3545.
Tiulpakov, A., Kalintchenko, N., Semitcheva, T., Polyakov, A., Dedov, I., Sverdlova, P., Kolesnikova, G., Peterkova, V., Rubtsov, P., 2005. A potential rearrangement between CYP19 and TRPM7 genes on chromosome 15q21.2 as a cause of aromatase excess syndrome. J. Clin. Endocrinol. Metab. 90, 4184–4190.
Tkachuk, D.C., Kohler, S., Cleary, M.L., 1992. Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. Cell 71, 691–700.
Tonon, G., Modi, S., Wu, L., Kubo, A., Coxon, A.B., Komiya, T., O'Neil, K., Stover, K., El-Naggar, A., Griffin, J.D., Kirsch, I.R., Kaye, F.J., 2003. t(11;19)(q21;p13) translocation in mucoepidermoid carcinoma creates a novel fusion product that disrupts a Notch signaling pathway. Nat. Genet. 33, 208–213.
Totoki, Y., Tatsuno, K., Yamamoto, S., Arai, Y., Hosoda, F., Ishikawa, S., Tsutsumi, S., Sonoda, K., Totsuka, H., Shirakihara, T., Sakamoto, H., Wang, L., Ojima, H., Shimada,

K., Kosuge, T., Okusaka, T., Kato, K., Kusuda, J., Yoshida, T., Aburatani, H., Shibata, T., 2011. High-resolution characterization of a hepatocellular carcinoma genome. Nat. Genet. 43, 464–469.

Ueda, C., Akasaka, T., Kurata, M., Maesako, Y., Nishikori, M., Ichinohasama, R., Imada, K., Uchiyama, T., Ohno, H., 2002. The gene for interleukin-21 receptor is the partner of BCL6 in t(3;16)(q27;p11), which is recurrently observed in diffuse large B-cell lymphoma. Oncogene 21, 368–376.

van Dongen, J.J., Macintyre, E.A., Gabert, J.A., Delabesse, E., Rossi, V., Saglio, G., Gottardi, E., Rambaldi, A., Dotti, G., Griesinger, F., Parreira, A., Gameiro, P., Diáz, M.G., Malec, M., Langerak, A.W., San Miguel, J.F., Biondi, A., 1999. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. Leukemia 13, 1901–1928.

Veltman, I.M., Vreede, L.A., Cheng, J., Looijenga, L.H.J., Janssen, B., Schoenmakers, E.F.P.M., Yeh, E.T.H., van Kessel, A.G., 2005. Fusion of the SUMO/Sentrin-specific protease 1 gene SENP1 and the embryonic polarity-related mesoderm development gene MESDC2 in a patient with an infantile teratoma and a constitutional t(12;15)(q13;q25). Hum. Mol. Genet. 14, 1955–1963.

Vizmanos, J.L., Novo, F.J., Román, J.P., Baxter, E.J., Lahortiga, I., Larrayoz, M.J., Odero, M.D., Giraldo, P., Calasanz, M.J., Cross, N.C.P., 2004. NIN, a gene encoding a CEP110-like centrosomal protein, is fused to PDGFRB in a patient with a t(5;14)(q33;q24) and an imatinib-responsive myeloproliferative disorder. Cancer Res. 64, 2673–2676.

Walz, C., Haferlach, C., Hänel, A., Metzgeroth, G., Erben, P., Gosenca, D., Hochhaus, A., Cross, N.C.P., Reiter, A., 2009. Identification of a MYO18A-PDGFRB fusion gene in an eosinophilia-associated atypical myeloproliferative neoplasm with a t(5;17)(q33-34;q11.2). Genes Chromosomes Cancer 48, 179–183.

Wheway, J.M., Yau, S.C., Nihalani, V., Ellis, D., Irving, M., Splitt, M., Roberts, R.G., 2003. A complex deletion-inversion-deletion event results in a chimeric IL1RAPL1-dystrophin transcript and a contiguous gene deletion syndrome. J. Med. Genet. 40, 127–131.

Wilkinson, K., Velloso, E.R.P., Lopes, L.F., Lee, C., Aster, J.C., Shipp, M.A., Aguiar, R.C.T., 2003. Cloning of the t(1;5)(q23;q33) in a myeloproliferative disorder associated with eosinophilia: involvement of PDGFRB and response to imatinib. Blood 102, 4187–4190.

Yang, J.J., Cho, S.Y., Suh, J.-T., Lee, H.J., Lee, W.-I., Yoon, H.-J., Baek, S.K., Park, T.S., 2012. Detection of RUNX1-MECOM fusion gene and t(3;21) in a very elderly patient having acute myeloid leukemia with myelodysplasia-related changes. Ann Lab Med 32, 362–365.

Yoshida, S., Kaneita, Y., Aoki, Y., Seto, M., Mori, S., Moriyama, M., 1999. Identification of heterologous translocation partner genes fused to the BCL6 gene in diffuse large B-cell lymphomas: 5'-RACE and LA - PCR analyses of biopsy samples. Oncogene 18, 7994–7999.

Zani, V.J., Asou, N., Jadayel, D., Heward, J.M., Shipley, J., Nacheva, E., Takasuki, K., Catovsky, D., Dyer, M.J., 1996. Molecular cloning of complex chromosomal translocation t(8;14;12)(q24.1;q32.3;q24.1) in a Burkitt lymphoma cell line defines a new gene (BCL7A) with homology to caldesmon. Blood 87, 3124–3134.