Published OnlineFirst August 20, 2012; DOI:10.1158/1078-0432.CCR-11-2947

Table 1. Characteristics of subjects positive for *EML4-ALK* by the RT-PCR diagnostic system.

Identification number	Sex/ age, y	Pathologic classification	Specimen type	*EML4-ALK* variant	Smoking history (pack-years)	TNM classification	Clinical stage	iAEP	*EGFR* mutation	*KRAS* mutation
J-#1	M/27	Adenocarcinoma	Sputum (2 different time points)	E13;A20	0	cT4N3M1	4	+	–	–
J-#4	F/39	Adenocarcinoma	Metastatic lymph node	E20;A20	NA	cTxN3M1	4	+	–	–
J-#7	M/74	Adenocarcinoma	Bronchial washing fluid	E13;A20	50	cT4N3M1	4	ND	–	–
J-#12	F/56	Adenocarcinoma	Resected tumor	E13;A20	0	cT1N0M0	1A	+	–	ND
J-#53	M/48	Adenocarcinoma	Tumor biopsy/sputum	E13;A20	0	cT3N2M1	4	+	–	–
J-#88	F/37	Adenocarcinoma	Pleural effusion	E13;A20	0	cT4N3M1	4	ND	–	–
J-#127	F/49	Adenocarcinoma	Tumor biopsy	E6a/b;A20	0.9	cT1N2M1	4	+	–	–
J-#189	F/37	Adenocarcinoma	Resected tumor	E14::ins2; ins56A20	0	cT2N1M1	4	+	–	–
J-#210	F/37	Adenocarcinoma	Resected tumor	E13;A20	0	cT4N2M1	4	ND	–	–
J-#215	F/61	Adenocarcinoma	Sputum	E13;A20	82	cT4N2M1	4	ND	–	–
J-#330	M/72	Adenocarcinoma	Pleural effusion/resected tumor (2 different regions)	E13;A20	0	cT4N1M1	4	+	–	–
J-#350	F/53	Adenocarcinoma	Pleural effusion	E13;A20	0	cT4N2M0	3B	ND	–	–
J-#378	F/78	Adenocarcinoma	Resected tumor	E13;A20	0	cT1N0M0	1A	ND	–	–
J-#385	F/80	Adenocarcinoma	Pleural effusion	E6a/b;A20	0	cT4N3M1	4	ND	–	–
J-#391	F/55	Adenocarcinoma	Tumor biopsy	E13;A20	16.5	cT2N2M1	4	+	–	ND
J-#392	F/38	Adenocarcinoma	Tumor biopsy	E13;A20	34	cT4N2M0	3B	+	–	ND
J-#393	F/42	Adenocarcinoma	Tumor biopsy	E13;A20	0	cT4N3M1	4	–	–	ND
J-#409	F/35	Adenocarcinoma	Tumor biopsy	E13;A20	0	cT4N0M0	3B	+	–	–
J-#422	M/69	Adenocarcinoma	Tumor biopsy	E6a/b;A20	0	cT2N2M0	3A	ND	–	–
J-#450	F/30	Adenocarcinoma	Bronchial washing fluid	E6a/b;A20	0	cT4N2M1	4	+	–	–
J-#530	F/55	Adenocarcinoma	Bronchial washing fluid	E13;A20	0	cT1N1M1	4	+	+	ND
J-#646	F/36	Adenocarcinoma	Bronchial washing fluid	E6a/b;A20	0	cT2N3M0	3B	ND	–	–
J-#657	F/62	Adenocarcinoma	Bronchial washing fluid	E13;A20	15	cT4N2M0	3B	ND	–	–
J-#759	F/32	Adenocarcinoma	Resected tumor	E13;A20	12	cT1N0M0	1A	ND	–	–
J-#771	M/32	Adenocarcinoma	Tumor biopsy	E6a/b;A20	15	cT1N3M1	3B	ND	–	–
J-#817	M/33	Adenocarcinoma	Pleural effusion	E13;A20	0	cT2N1M1	4	ND	–	–
J-#848	M/57	Adenocarcinoma	Bronchial washing fluid	E18;E20	0	cT4N2M0	3B	ND	–	–
J-#887	F/32	Adenocarcinoma	Bronchial washing fluid	E6a/b;A20	0	cTxN3M1	4	ND	–	ND
J-#927	M/36	Adenocarcinoma	Bronchial washing fluid	E6a/b;A20	30	cT4N3M1	4	–	–	–
J-#928	F/71	Adenocarcinoma	Bronchial washing fluid	E6a/b;A20	0	cT4N3M1	4	ND	–	–
J-#996	M/52	Adenocarcinoma	Bronchial washing fluid	E6a/b;A20	0	cT3N3M0	3B	ND	–	–
J-#1001	F/32	Adenocarcinoma	Bronchial washing fluid	E13;A20	6.5	cT2N2M1	4	+	–	–

Abbreviations: F, female; M, male; NA, not available; ND, not determined.

CTGTGGAGGCTGAACTGGATC-3′ and 5′-TCATCAACAAGCTCCACGGTG-3′) specific for the human ribonuclease P (RNase P) gene (GenBank accession number NM_005837). Given that we previously showed that the abundance of RNase P mRNA is similar to that of *EML4-ALK* mRNA in NSCLCs (data not shown), we used the successful amplification of RT-PCR products for RNase P as a threshold for selection of specimens for further analysis. Exclusion of small cell lung cancer specimens and filtering on the basis of RNase P mRNA abundance resulted in the isolation of 808 specimens of primary NSCLCs obtained from 754 individuals.

As shown in Supplementary Fig. S1, bronchial washing fluid, including bronchoalveolar lavage fluid and washing fluid for the brush, needle, forceps, and other implements used in bronchoscopy, constituted 66.3% of the 808 eligible samples, with the remaining specimens including pleural effusion (12.8%); surgically resected tumor (7.05%); sputum (4.33%); tumor biopsy tissue including that obtained by transbronchial lung biopsy and transbronchial needle aspiration (3.71%); peripheral blood (3.71%); cardiac effusion, spinal fluid, or ascites (1.36%); and metastatic lesions of NSCLCs (0.74%).

Multiplex RT-PCR analysis of EML4-ALK and KIF5B-ALK

Each specimen (with the exception of resected tumors) was mixed with an equal volume of RLT buffer at the Institute at which it was harvested. The resulting mixture was sent to Jichi Medical University, where DNA and RNA were extracted with the use of an automated BioRobot EZ1 workstation (Qiagen). The isolated RNA was subjected to RT with a ReverTra Ace qPCR RT kit (Toyobo), and the resulting cDNA was subjected to PCR for 50 cycles of incubation at 94°C for 15 seconds, 60°C for 30 seconds, and 72°C for 1 minute with AmpliTAQ Gold DNA polymerase (Applied Biosystems) and with 2 μmol/L of each of the following

Published OnlineFirst August 20, 2012; DOI:10.1158/1078-0432.CCR-11-2947

primers: F-1, 5′-GCTTTCCCCGCAAGATGGACGG-3′; F-2, 5′-TACCAGTGCTGTCTCAATTGCAGG-3′; F-3, 5′-GTGCAGTGTTTAGCATTCTTGGGG-3′; F-4, 5′-AGCTACATCACACACCTTGACTGG-3′; F-5, 5′-TCAAGCACATCTCAAGAGCAAGTG-3′; F-6, 5′-ATCCTGCGGAACACTATTCAGTGG-3′; F-7, 5′-GACAGTTGGAGGAATCTGTCGATG-3′; F-8, 5′-CAGCTGAGAGAGTGAAAGCTTTGG-3′; and R-1, 5′-TCTTGCCAGCAAAGCAGTAGTTGG-3′. All PCR products were subjected to Sanger sequencing to confirm the presence of *EML4-ALK* or *KIF5B-ALK* cDNA.

Results

Multiplex RT-PCR system

In addition to the original *EML4-ALK* fusion cDNA in which exon 13 of *EML4* is fused to exon 20 of *ALK* in an in-frame manner (designated the E13;A20 variant by analogy with karyotype nomenclature; see http://atlasgeneticsoncology.org/Tumors/inv2p21p23NSCCLungID5667.html), 14 different variants of *EML4-ALK* have been described (1, 14, 21–27). Seven exons of *EML4* are theoretically capable of in-frame fusion with exon 20 of *ALK* (Fig. 1A), and all but the E1;A20 variant would be expected to produce an oncogenic EML4-ALK protein, given that the coiled-coil domain encoded by exon 2 is required for constitutive dimerization of EML4-ALK. In addition, 6 different exons of *KIF5B* are theoretically capable of in-frame fusion with exon 20 of *ALK* (Fig. 1A).

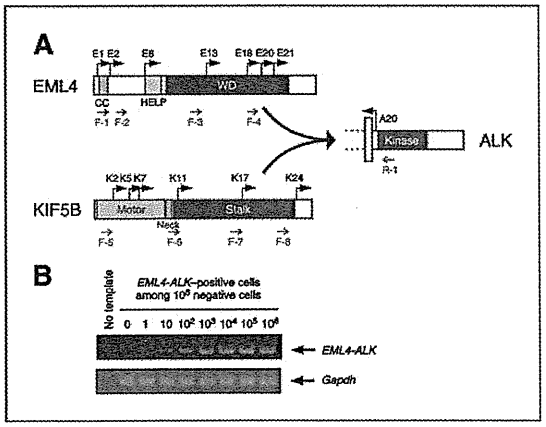

Figure 1. Multiplex RT-PCR system for detection of *EML4-ALK* and *KIF5B-ALK*. A, schematic representation of the structure of EML4, KIF5B, and ALK proteins. The positions of exons (E for *EML4* and K for *KIF5B*) theoretically capable of fusing in-frame to exon 20 (A20) of *ALK* are indicated by arrows. The positions of 8 forward primers (F-1 to F-8) and 1 reverse primer (R-1) for PCR are also indicated below the corresponding proteins. EML4 contains a coiled-coil domain (CC), a hydrophobic EMAP-like protein domain (HELP), and WD repeats (WD). KIF5B consists of an amino-terminal ATP-dependent motor domain, a neck region, and a stalk region. B, various numbers (0 to 1×10^6) of *EML4-ALK* (E13;A20)–positive BA/F3 cells (1) were mixed with a fixed number (1×10^6) of *EML4-ALK*–negative BA/F3 cells, and each mixture was analyzed with our multiplex RT-PCR system. A cDNA for mouse glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) was also amplified by PCR as an internal control with the primers 5′-TGTGTCCGTCGTGGATCTGA-3′ and 5′-CCTGCTTCACCACCTTCTTGA-3′.

To detect any such *EML4-ALK* or *KIF5B-ALK* fusion mRNAs, we developed a multiplex RT-PCR system. We had previously screened our archive of frozen tumors by RT-PCR analysis with 2 forward primers targeted to *EML4* and 1 reverse primer targeted to *ALK* (24), but such PCR conditions resulted in the amplification of products as large as ~1,300 bp for some variants. In this prospective study, we were faced with the analysis of a large number of samples with different levels of RNA quality. If the size of PCR products varied substantially among different *EML4-ALK* or *KIF5B-ALK* variants, some variants with large PCR products might not be amplified efficiently from specimens with low RNA quality. To be able to diagnose all possible fusions even with such samples, we therefore designed 4 forward primers for each of *EML4* and *KIF5B* so that the size variation among all possible RT-PCR products is minimal (Fig. 1A). This new multiplex system faithfully detected all known fusion variants from *EML4-ALK*–positive specimens in our previous archive of NSCLCs (data not shown).

To examine the sensitivity of our RT-PCR system, we mixed *EML4-ALK*–expressing BA/F3 cells (0 to 1×10^6) with *EML4-ALK*–negative cells (1×10^6) and then subjected them to RT-PCR analysis. A fusion cDNA was readily identified even with 10 positive cells (0.001%) among 1×10^6 negative cells (Fig. 1B), showing the high sensitivity of the RT-PCR system.

To confirm the potential of our RT-PCR–based system, we compared it with a sensitive immunohistochemical approach and with FISH for the diagnosis of our archive of surgically resected and freshly frozen tumors with high RNA quality. Fifteen NSCLC specimens that previously stained positive by our sensitive immunohistochemical approach, which is based on an intercalated antibody-enhanced polymer (iAEP) method (14), were analyzed by RT-PCR and FISH together with 96 iAEP-negative specimens in a blinded manner. RT-PCR analysis of all these specimens ($n = 111$) yielded a diagnosis identical to that obtained with the iAEP method ($P = 7.3 \times 10^{-19}$, Fisher exact test; data not shown). Analysis of the same sample set by a split FISH assay with Vysis probes (Abbott Laboratories) revealed that all of the iAEP-positive cases showed a rearranged *ALK* locus, whereas one iAEP-negative sample gave a discordant result (negative by iAEP and RT-PCR but positive by FISH; Supplementary Fig. S2). The reason for this discrepant result remains unclear, but the multiple signals obtained with the 3′-*ALK* probe in the FISH analysis are indicative of amplification of the *ALK* gene or its adjacent region. Despite this discrepancy, the RT-PCR and iAEP data were highly concordant with the FISH results ($P = 1.2 \times 10^{-17}$, Fisher exact test). Compared with the iAEP method, therefore, both the sensitivity and specificity of our RT-PCR system were 100%. In comparison with the Vysis FISH, the sensitivity and specificity of RT-PCR were 93.8% and 100%, respectively.

Detection of EML4-ALK

Screening of the 808 eligible specimens with our multiplex RT-PCR system identified positive products in 36 samples (4.46%) obtained from 32 different individuals

Published OnlineFirst August 20, 2012; DOI:10.1158/1078-0432.CCR-11-2947

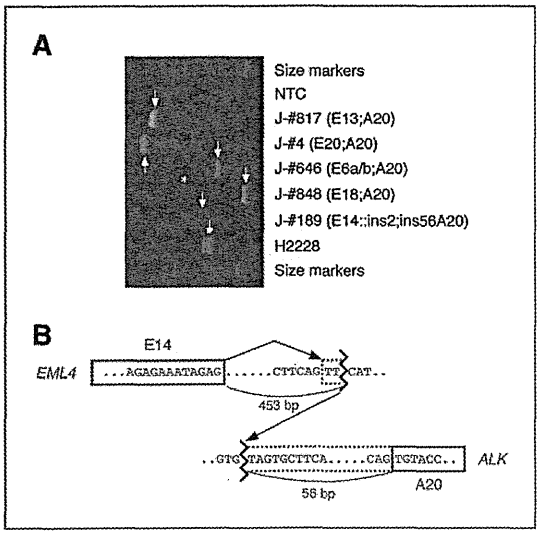

Figure 2. Multiplex RT-PCR detection of *EML4-ALK*–positive NSCLCs. A, RT-PCR products for each of the *EML4-ALK* variants identified in our cohort were separated by agarose gel electrophoresis. RT-PCR products spanning the *EML4-ALK* fusion points are indicated by arrows; the asterisk indicates a nonspecific product. An NSCLC cell line, H2228, harboring the E6a/b;A20 variant of *EML4-ALK* was used as a positive control for the PCR reaction. Size markers include a 50-bp DNA ladder (Invitrogen). NTC, no-template control. B, genomic structure of the fusion point for a novel variant of *EML4-ALK*. Nucleotide sequencing of the genomic PCR and RT-PCR products from patient J-#189 revealed that exon 14 of *EML4* (blue) was spliced to a TT sequence adjacent to the genomic ligation point, with transcription continuing in an in-frame manner into intron 19 and exon 20 of *ALK* (red).

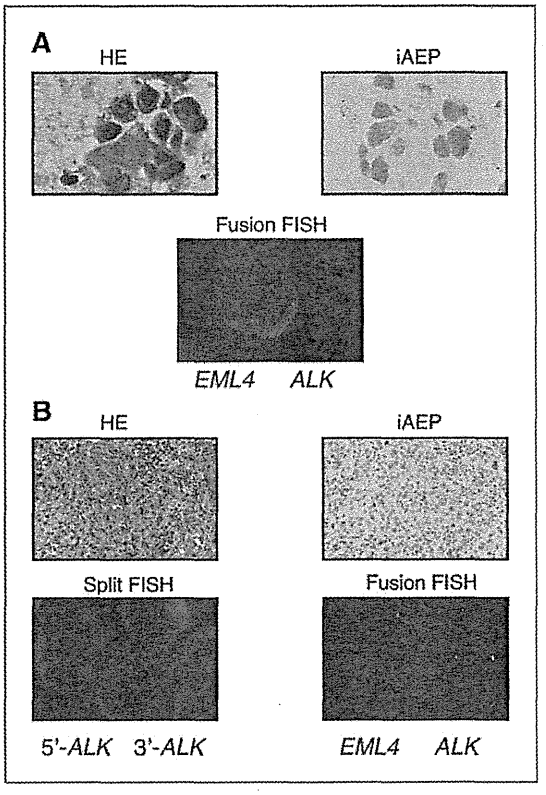

Figure 3. Specimens positive for *EML4-ALK* by RT-PCR but negative by iAEP-based IHC and by FISH. Sections of tumor biopsy specimens for J-#393 tumor (A) and J-#927 (B) were stained with hematoxylin-eosin (HE), subjected to immunohistochemical analysis by the iAEP method, and examined by split or fusion FISH. The color of fluorescence for the probes in each hybridization is indicated below the FISH images. Nuclei are stained blue with 4′,6-diamidino-2-phenylindole (DAPI).

(4.24%; Table 1, Fig. 2A). Nucleotide sequencing of each PCR product identified 19 cases positive for the E13;A20 variant, 10 cases for E6a/b;A20, a single case each for E18;A20, E20;A20, and a novel variant. *EML4-ALK* was detected in a wide range of specimens including bronchial washing fluid ($n = 11$), tumor biopsy ($n = 8$), resected tumor ($n = 7$), pleural effusion ($n = 5$), sputum ($n = 4$), and metastatic lymph node ($n = 1$). We did not detect any *KIF5B-ALK* cDNAs, confirming the rarity of this fusion gene.

Importantly, an E13;A20 product was consistently identified in both of the sputa obtained at different time points from patient J-#1. Likewise, an E13;A20 product was detected in both the tumor biopsy and sputum from patient J-#53 as well as in the pleural effusion and 2 resected tumor specimens from patient J-#330, supporting the reliability of our RT-PCR approach.

Sequence determination for the RT-PCR product from patient J-#189 revealed that exon 14 of *EML4* was fused to exon 20 of *ALK* with an intervening sequence. Genomic PCR analysis of the J-#189 specimen with a forward primer targeted to exon 14 of *EML4* and a reverse primer targeted to exon 20 of *ALK* yielded a specific product, nucleotide sequencing of which revealed that a position 453 bp downstream of *EML4* exon 14 was ligated to a position 56 bp upstream of *ALK* exon 20 (Fig. 2B). In the transcript of this fusion gene, exon 14 of *EML4* is thus spliced to a TT sequence that is located within *EML4* intron 14 and which is directly ligated to intron 19 of *ALK*. This splicing event results in an in-frame fusion between the mRNA sequences derived from *EML4* and *ALK*. Furthermore, a full-length cDNA for this variant, here designated E14::ins2;ins56A20, was isolated by RT-PCR analysis (Supplementary Fig. S3), and the potent transforming ability of the encoded protein was confirmed with an *in vitro* focus formation assay (Supplementary Fig. S4).

Comparison between multiplex RT-PCR and sensitive IHC

Finally, we applied the iAEP method to the *EML4-ALK*–positive cases for which FFPE specimens were also available ($n = 15$). All but 2 cases (J-#393 and J-#927) manifested clear immunoreactivity with antibodies to ALK (Table 1). FISH analysis of these 2 specimens also failed to detect the *EML4-ALK* rearrangements (Fig. 3). Given that genomic DNA was not available for the tumor of patient J-#393, we

Published OnlineFirst August 20, 2012; DOI:10.1158/1078-0432.CCR-11-2947

were not able to determine whether the PCR result was a false-positive. For J-#927, however, PCR analysis of genomic DNA with a forward primer targeted to *EML4* exon 6 and a reverse primer to *ALK* exon 20 resulted in the amplification of an approximately 8.8-kbp genomic fragment, nucleotide sequencing of which revealed a fusion event between intron 6 of *EML4* and intron 19 of *ALK* (Supplementary Fig. S5). Isolation of the genomic fusion point thus indicates that J-#927 indeed harbors an *EML4-ALK*–positive tumor.

Discussion

We have conducted a large-scale, prospective screening for *EML4-ALK* with an RT-PCR–based approach. Whereas RNA extraction and cDNA synthesis add extra labor to the diagnostic procedure, certain introns of *EML4* are too large (intron 6 spans >16 kbp, for instance) for reliable amplification by genomic PCR. We therefore adopted RT-PCR as the method for our prospective screening. Specific PCR products were successfully isolated from different types of specimen, even from sputum (J-#1, J-#53, J-#215) and washing fluid of a tumor biopsy needle (J-#530). Multiple positive results obtained with different specimens of the same individuals further reinforce the reliability of our multiplex RT-PCR system as a diagnostic tool for *EML4-ALK*–positive tumors. Importantly, a subset of *EML4-ALK*–positive individuals diagnosed in the present study entered a clinical trial for crizotinib, and the response rate of the evaluable patients ($n = 9$) was 100% with this drug, again verifying the accuracy of our RT-PCR–based diagnosis.

The frequency of *EML4-ALK* in our cohort was 4.24% for all NSCLC cases and 6.11% for lung adenocarcinoma, values similar to those obtained in previous studies (20, 21). However, our prevalence data might be overestimates because the knowledge of mutual exclusiveness for *EML4-ALK* and *EGFR* mutations may have affected patient selection for our specimen collection. Indeed, *EGFR* mutation frequency among our cohort (23.8%) is slightly lower than that (30.9%) determined in a previous large-scale screening in Japan (28).

The clinicopathologic features of patients with *EML4-ALK*–positive tumors determined in the present study are also in agreement with those previously described, with a bias toward a young age, adenocarcinoma histology, and never or light smoking. Whereas a previous large-scale screening for *EML4-ALK* based on FISH did not detect a sex preference for the fusion gene (7), our cohort revealed a significant female preference. Such a sex difference was evident even among individuals below 40 years of age ($P = 0.03$, Fisher' exact test) and among those with an adenocarcinoma histology ($P = 0.005$, Fisher' exact test). Further large-scale studies are warranted to determine whether this uneven sex distribution of *EML4-ALK* is related to particular clinicopathologic features or ethnic groups.

Given that *EML4-ALK* and *EGFR* mutations are almost mutually exclusive and that the fusion gene is enriched in lung adenocarcinoma with an early onset, it should prove to be clinically beneficial to pay special attention to such subsets of patients. Indeed, *EML4-ALK* was detected in 27.7% of *EGFR* mutation–negative adenocarcinomas in individuals of younger than 50 years and in 50.0% of those in individuals of younger than 40 years in our cohort. Given the marked efficacy of ALK inhibitors in patients with *EML4-ALK*–positive NSCLCs (7), however, physicians should not dismiss the diagnosis in other subsets of patients. For example, *EML4-ALK* was even detected in an 80-year-old woman and in another woman with an intense smoking history (82 pack-years; Table 1).

Multiplex RT-PCR has both advantages and disadvantages compared with other techniques. Importantly, the accuracy of RT-PCR–based diagnosis depends markedly on the RNA quality of specimens. In our cohort, for instance, 71 (7.75%) of the initial 916 specimens were excluded from *EML4-ALK* screening because of a failure to obtain PCR products for RNase P (the other 37 samples were excluded because they were not NSCLCs). Low RNA quality thus clearly hampers reliable RT-PCR–based diagnosis.

Also, as expected, there was a large variation in the PCR cycle number required for successful amplification among specimens. In our cohort, 50 cycles of PCR allowed detection of PCR products for all positive cases, but such extensive amplification may also generate nonspecific products (as shown in Fig. 2A). Further optimization of primer sequences or combinations may minimize the generation of such byproducts. Furthermore, whereas our system should be able to capture all in-frame fusions of *ALK* to *EML4* or *KIF5B*, it is not capable in its present form of detecting *ALK* fusions to other partners, such as *KLC1-ALK*, which was recently shown to be present infrequently in NSCLCs (29).

On the other hand, RT-PCR can be readily applied to specimens such as sputum, bronchial washing fluid, or pleural effusion that may not be suitable for preparation of FFPE samples. Whereas the latter 2 specimen types can be used for the preparation of cell blocks suitable for analysis by FISH or IHC, this procedure may not be as widely adopted in the clinic as is FISH or IHC. More importantly, it is difficult to generate cell blocks or FFPE samples from sputum. Our current prospective screening identified 4 *EML4-ALK*–positive sputa of 35 samples (Table 1, Supplementary Fig. S1), showing that sputum is a suitable specimen for RT-PCR analysis. Indeed, sputum was the only available specimen from patient J-#215 both for the diagnosis of NSCLCs and for the detection of *EML4-ALK*. If RT-PCR had not been applied to this patient's sputum, we would not have been able to identify her tumor as positive for *EML4-ALK*, and she would not have had the chance to receive treatment with an ALK inhibitor in Japan.

Furthermore, PCR-based detection of *EML4-ALK* should have a higher analytic sensitivity compared with IHC or FISH (Fig. 1B). Even with sputum obtained from a patient with chronic bronchitis, RT-PCR was able to readily detect *EML4-ALK* at a concentration of 10 positive cells/mL (1). Thus, provided that RNA is not substantially degraded, RT-PCR–based diagnosis is expected to have a strong advantage

Published OnlineFirst August 20, 2012; DOI:10.1158/1078-0432.CCR-11-2947

with regard to the detection of low numbers of *EML4-ALK*–positive cells.

Ideally, every NSCLC case should be examined for the presence of *EML4-ALK*, with a sensitive and accurate diagnostic strategy for the oncogenic fusion being essential for the adoption of ALK inhibitors in the clinic. Given the reliable detection of *EML4-ALK* mRNA by multiplex RT-PCR shown in the present study, we propose the following scheme for the comprehensive diagnosis of *EML4-ALK*–positive NSCLCs. For sputum, bronchial lavage fluid, pleural effusion, or other specimens that may not be suitable for the preparation of FFPE tissue, multiplex RT-PCR should be applied to detect *ALK* fusion mRNAs. In contrast, given that FFPE specimens usually have fragmented RNA, they should be subjected to FISH and to sensitive immunohistochemical analysis such as that described previously (14, 15). Furthermore, FISH or IHC can be applied to cell blocks prepared from some non-FFPE specimens. No single technique is therefore able to detect *EML4-ALK* in all types of specimen, and appropriate tests should be chosen on the basis of the specimens available for a given patient.

Disclosure of Potential Conflicts of Interest

H. Mano is the CEO of CureGene Co., Ltd.; has commercial research grant from Illumina, Inc. and Astellas Pharma Inc.; has ownership interest (including patents); and is on the consultant/advisory board of Chugai Pharmaceutical, Astellas Pharma Inc., and Daiichi Sankyo Co., Ltd. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: K. Hagiwara, H. Mano
Development of methodology: K. Takeuchi, Y.L. Choi
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Soda, K. Isobe, A. Inoue, S. Oizumi, Y. Fujita, A. Gemma, Y. Yamashita, K. Takeuchi, H. Miyazawa, T. Tanaka, K. Hagiwara
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M. Soda, T. Ueno, H. Mano
Writing, review, and/or revision of the manuscript: S. Oizumi, A. Gemma, K. Hagiwara, H. Mano
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Maemondo, K. Takeuchi, K. Hagiwara
Study supervision: H. Mano

Grant Support

This study was supported in part by a grant for Research on Human Genome Tailor-made from the Ministry of Health, Labor, and Welfare of Japan; by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and by grants from the Japan Society for the Promotion of Science, from Takeda Science Foundation, from Mochida Memorial Foundation for Medical and Pharmaceutical Research, from The Mitsubishi Foundation, and from The Sagawa Foundation for Promotion of Cancer Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 17, 2011; revised June 26, 2012; accepted August 3, 2012; published OnlineFirst August 20, 2012.

References

1. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming *EML4-ALK* fusion gene in non-small-cell lung cancer. Nature 2007;448:561–6.
2. Mano H. ALKoma: a cancer subtype with a shared target. Cancer Discov 2012;2:495–502.
3. Shaw AT, Solomon B. Targeting anaplastic lymphoma kinase in lung cancer. Clin Cancer Res 2011;17:2081–6.
4. Soda M, Takada S, Takeuchi K, Choi YL, Enomoto M, Ueno T, et al. A mouse model for *EML4-ALK*-positive lung cancer. Proc Natl Acad Sci U S A 2008;105:19893–7.
5. Chen Z, Sasaki T, Tan X, Carretero J, Shimamura T, Li D, et al. Inhibition of ALK, PI3K/MEK, and HSP90 in murine lung adenocarcinoma induced by EML4-ALK fusion oncogene. Cancer Res 2010;70:9827–36.
6. Marzec M, Kasprzycka M, Ptasznik A, Wlodarski P, Zhang Q, Odum N, et al. Inhibition of ALK enzymatic activity in T-cell lymphoma cells induces apoptosis and suppresses proliferation and STAT3 phosphorylation independently of Jak3. Lab Invest 2005;85:1544–54.
7. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 2010;363:1693–703.
8. Katayama R, Khan TM, Benes C, Lifshits E, Ebi H, Rivera VM, et al. Therapeutic strategies to overcome crizotinib resistance in non-small cell lung cancers harboring the fusion oncogene EML4-ALK. Proc Natl Acad Sci U S A 2011;108:7535–40.
9. Lovly CM, Heuckmann JM, de Stanchina E, Chen H, Thomas RK, Liang C, et al. Insights into ALK-driven cancers revealed through development of novel ALK tyrosine kinase inhibitors. Cancer Res 2011;71:4920–31.
10. Sakamoto H, Tsukaguchi T, Hiroshima S, Kodama T, Kobayashi T, Fukami TA, et al. CH5424802, a selective ALK inhibitor capable of blocking the resistant gatekeeper mutant. Cancer Cell 2011;19:679–90.
11. Gerber DE, Minna JD. ALK inhibition for non-small cell lung cancer: from discovery to therapy in record time. Cancer Cell 2010;18:548–51.
12. Butrynski JE, D'Adamo DR, Hornick JL, Dal Cin P, Antonescu CR, Jhanwar SC, et al. Crizotinib in ALK-rearranged inflammatory myofibroblastic tumor. N Engl J Med 2010;363:1727–33.
13. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med 2009;361:947–57.
14. Takeuchi K, Choi YL, Togashi Y, Soda M, Hatano S, Inamura K, et al. KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. Clin Cancer Res 2009;15:3143–9.
15. Mino-Kenudson M, Chirieac LR, Law K, Hornick JL, Lindeman N, Mark EJ, et al. A novel, highly sensitive antibody allows for the routine detection of ALK-rearranged lung adenocarcinomas by standard immunohistochemistry. Clin Cancer Res 2010;16:1561–71.
16. Nagai Y, Miyazawa H, Huqun, Tanaka T, Udagawa K, Kato M, et al. Genetic heterogeneity of the epidermal growth factor receptor in non-small cell lung cancer cell lines revealed by a rapid and sensitive detection system, the peptide nucleic acid-locked nucleic acid PCR clamp. Cancer Res 2005;65:7276–82.
17. Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS, et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. J Clin Oncol 2009;27:4247–53.
18. Horn L, Pao W. EML4-ALK: honing in on a new target in non-small-cell lung cancer. J Clin Oncol 2009;27:4232–5.
19. Tiseo M, Gelsomino F, Boggiani D, Bortesi B, Bartolotti M, Bozzetti C, et al. EGFR and EML4-ALK gene mutations in NSCLC: a case report of erlotinib-resistant patient with both concomitant mutations. Lung Cancer (Amsterdam, the Netherlands) 2011;71:241–3.
20. Inamura K, Takeuchi K, Togashi Y, Hatano S, Ninomiya H, Motoi N, et al. EML4-ALK lung cancers are characterized by rare other mutations, a TTF-1 cell lineage, an acinar histology, and young onset. Mod Pathol 2009;22:508–15.
21. Wong DW, Leung EL, So KK, Tam IY, Sihoe AD, Cheng LC, et al. The EML4-ALK fusion gene is involved in various histologic types of lung

Published OnlineFirst August 20, 2012; DOI:10.1158/1078-0432.CCR-11-2947

cancers from nonsmokers with wild-type EGFR and KRAS. Cancer 2009;115:1723–33.
22. Choi YL, Takeuchi K, Soda M, Inamura K, Togashi Y, Hatano S, et al. Identification of novel isoforms of the *EML4-ALK* transforming gene in non-small cell lung cancer. Cancer Res 2008;68:4971–6.
23. Koivunen JP, Mermel C, Zejnullahu K, Murphy C, Lifshits E, Holmes AJ, et al. EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. Clin Cancer Res 2008;14:4275–83.
24. Takeuchi K, Choi YL, Soda M, Inamura K, Togashi Y, Hatano S, et al. Multiplex reverse transcription-PCR screening for *EML4-ALK* fusion transcripts. Clin Cancer Res 2008;14:6618–24.
25. Lin E, Li L, Guan Y, Soriano R, Rivers CS, Mohan S, et al. Exon array profiling detects EML4-ALK fusion in breast, colorectal, and non-small cell lung cancers. Mol Cancer Res 2009;7:1466–76.
26. Takahashi T, Sonobe M, Kobayashi M, Yoshizawa A, Menju T, Nakayama E, et al. Clinicopathologic features of non-small-cell lung cancer with EML4-ALK fusion gene. Ann Surg Oncol 2010;17:889–97.
27. Sanders HR, Li HR, Bruey JM, Scheerle JA, Meloni-Ehrig AM, Kelly JC, et al. Exon scanning by reverse transcriptase-polymerase chain reaction for detection of known and novel EML4-ALK fusion variants in non-small cell lung cancer. Cancer Genet 2011;204:45–52.
28. Toyooka S, Matsuo K, Shigematsu H, Kosaka T, Tokumo M, Yatabe Y, et al. The impact of sex and smoking status on the mutational spectrum of epidermal growth factor receptor gene in non small cell lung cancer. Clin Cancer Res 2007;13:5763–8.
29. Togashi Y, Soda M, Sakata S, Sugawara E, Hatano S, Asaka R, et al. KLC1-ALK: a novel fusion in lung cancer identified using a formalin-fixed paraffin-embedded tissue only. PLoS One 2012;7:e31323.

nature medicine

ARTICLES

A common *BIM* deletion polymorphism mediates intrinsic resistance and inferior responses to tyrosine kinase inhibitors in cancer

King Pan Ng[1,23], Axel M Hillmer[2,23], Charles T H Chuah[1,3,23], Wen Chun Juan[1,23], Tun Kiat Ko[1], Audrey S M Teo[2], Pramila N Ariyaratne[2], Naoto Takahashi[4], Kenichi Sawada[4], Yao Fei[2,5], Sheila Soh[1], Wah Heng Lee[2], John W J Huang[1], John C Allen Jr[6], Xing Yi Woo[2], Niranjan Nagarajan[2], Vikrant Kumar[2], Anbupalam Thalamuthu[2], Wan Ting Poh[2], Ai Leen Ang[3], Hae Tha Mya[3], Gee Fung How[3], Li Yi Yang[3], Liang Piu Koh[7], Balram Chowbay[8], Chia-Tien Chang[1], Veera S Nadarajan[9], Wee Joo Chng[7,10,11], Hein Than[3], Lay Cheng Lim[3], Yeow Tee Goh[3], Shenli Zhang[1], Dianne Poh[1], Patrick Tan[1,2,11], Ju-Ee Seet[12], Mei-Kim Ang[13], Noan-Minh Chau[13], Quan-Sing Ng[13], Daniel S W Tan[13], Manabu Soda[14], Kazutoshi Isobe[15], Markus M Nöthen[16], Tien Y Wong[17], Atif Shahab[2], Xiaoan Ruan[2], Valère Cacheux-Rataboul[2], Wing-Kin Sung[2], Eng Huat Tan[13], Yasushi Yatabe[18], Hiroyuki Mano[14,19], Ross A Soo[7,11], Tan Min Chin[7], Wan-Teck Lim[13,20], Yijun Ruan[2,21] & S Tiong Ong[1,3,13,22]

Tyrosine kinase inhibitors (TKIs) elicit high response rates among individuals with kinase-driven malignancies, including chronic myeloid leukemia (CML) and epidermal growth factor receptor–mutated non–small-cell lung cancer (EGFR NSCLC). However, the extent and duration of these responses are heterogeneous, suggesting the existence of genetic modifiers affecting an individual's response to TKIs. Using paired-end DNA sequencing, we discovered a common intronic deletion polymorphism in the gene encoding BCL2-like 11 (BIM). BIM is a pro-apoptotic member of the B-cell CLL/lymphoma 2 (BCL2) family of proteins, and its upregulation is required for TKIs to induce apoptosis in kinase-driven cancers. The polymorphism switched *BIM* splicing from exon 4 to exon 3, which resulted in expression of *BIM* isoforms lacking the pro-apoptotic BCL2-homology domain 3 (BH3). The polymorphism was sufficient to confer intrinsic TKI resistance in CML and EGFR NSCLC cell lines, but this resistance could be overcome with BH3-mimetic drugs. Notably, individuals with CML and EGFR NSCLC harboring the polymorphism experienced significantly inferior responses to TKIs than did individuals without the polymorphism ($P = 0.02$ for CML and $P = 0.027$ for EGFR NSCLC). Our results offer an explanation for the heterogeneity of TKI responses across individuals and suggest the possibility of personalizing therapy with BH3 mimetics to overcome *BIM*-polymorphism–associated TKI resistance.

The use of TKIs has elicited remarkable therapeutic responses in individuals presenting with a broad range of malignancies driven by oncogenic kinases[1]. However, before the use of TKIs, such malignancies were regarded as highly chemoresistant, as exemplified by breakpoint cluster region (BCR)–c-abl oncogene 1, non-receptor tyrosine kinase (ABL1) kinase-driven CML and EGFR NSCLC[2,3]. After the advent of TKIs, treatment responses in both of these cancers typically approached 80% (refs. 4,5). These clinical observations emphasized the importance of classifying tumors according to their molecular drivers and at the same time stimulated the search for biomarkers that could identify the 20% of individuals at risk for primary or intrinsic TKI resistance, as well as guide therapy to overcome this resistance.

[1]Cancer & Stem Cell Biology Signature Research Programme, Duke–National University of Singapore (NUS) Graduate Medical School, Singapore. [2]Genome Institute of Singapore, Singapore. [3]Department of Haematology, Singapore General Hospital, Singapore. [4]Department of Hematology, Nephrology and Rheumatology, Akita University Graduate School of Medicine, Akita, Japan. [5]Department of Epidemiology and Public Health, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [6]Centre for Quantitative Medicine, Duke-NUS Graduate Medical School, Singapore. [7]Department of Hematology-Oncology, National University Cancer Institute of Singapore, National University Health System, Singapore. [8]Clinical Pharmacology Laboratory, National Cancer Centre, Singapore. [9]University of Malaya, Kuala Lumpur, Malaysia. [10]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [11]Cancer Science Institute of Singapore, National University of Singapore, Singapore. [12]Department of Pathology, National University Health System, Singapore. [13]Department of Medical Oncology, National Cancer Centre, Singapore. [14]Division of Functional Genomics, Jichi Medical University, Tochigi, Japan. [15]Department of Respiratory Medicine, Toho University Omori Medical Center, Tokyo, Japan. [16]Institute of Human Genetics, University of Bonn, Bonn, Germany. [17]Singapore Eye Research Institute, Singapore National Eye Centre and National University Health System, Singapore. [18]Department of Pathology and Molecular Diagnostics, Aichi Cancer Center, Aichi, Japan. [19]Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. [20]Office of Clinical Sciences, Duke-NUS Graduate Medical School, Singapore. [21]Department of Biochemistry, National University of Singapore, Singapore. [22]Department of Medicine, Division of Medical Oncology, Duke University Medical Center, Durham, North Carolina, USA. [23]These authors contributed equally to this work. Correspondence should be addressed to Y.R. (ruanyj@gis.a-star.edu.sg), W.-T.L. (dmolwt@nccs.com.sg) or S.T.O. (sintiong.ong@duke-nus.edu.sg).

Received 7 September 2011; accepted 21 February 2012; published online 18 March 2012; doi:10.1038/nm.2713

In this respect, we note that although polymorphisms in genes regulating drug metabolism provide useful information to modify the dosing of therapeutic agents[6], few examples exist in the germline that predict response to targeted therapies.

Accordingly, we investigated whether polymorphisms affecting TKI sensitivity might account for the 20% of TKI-treated individuals with poor responses and whether these polymorphisms might be enriched among genes that are crucial in the apoptotic response to TKIs. One such candidate gene is *BCL2L11* (also known as *BIM*), which encodes a BH3-only protein that is a BCL2 family member. The BH3-only proteins activate cell death by either opposing the prosurvival members of the BCL2 family (BCL2, BCL2-like 1 (BCL-XL, also known as BCL2L1), myeloid cell leukemia sequence 1 (MCL1) and BCL2-related protein A1 (BCL2A1)) or by binding to the pro-apoptotic BCL2 family members (BCL2-associated X protein (BAX) and BCL2-antagonist/killer 1 (BAK1)) and directly activating their pro-apoptotic functions[7]. Others have previously shown that several kinase-driven cancers, including CML and EGFR NSCLC, maintain a survival advantage by suppressing *BIM* transcription and by targeting BIM protein for proteasomal degradation through mitogen-activated protein kinase 1 (MAPK1)-dependent phosphorylation[8–13]. Furthermore, in all of these malignancies, BIM upregulation is required for TKIs to induce apoptosis, and suppression of BIM expression is sufficient to confer *in vitro* TKI resistance[8–13].

Here we describe the discovery of a common deletion polymorphism in the *BIM* gene that results in the generation of alternatively spliced isoforms of BIM that lack the crucial BH3 domain. This polymorphism has a profound effect on the TKI sensitivity of CML and EGFR NSCLC cells, such that one copy of the deleted allele is sufficient to render cells intrinsically TKI resistant. We show that individuals with the polymorphism have markedly inferior responses to TKI than do individuals without the polymorphism. Specifically, the polymorphism correlated with a lesser degree of response to imatinib, a TKI, in CML as well as a shorter progression-free survival (PFS) with EGFR TKI therapy in EGFR NSCLC.

RESULTS

A new *BIM* deletion polymorphism in resistant CML samples

To identify new TKI-resistance mechanisms in CML, we used massively parallel DNA sequencing of paired-end ditags[14,15] to interrogate the genomes of five CML samples obtained from subjects who were either sensitive to or resistant to treatment with TKIs (**Supplementary Tables 1** and **2**). We identified the BCR-ABL1 translocation in all CML samples, but not in control samples from patients in complete remission, and we also identified several CML-specific structural variations (**Supplementary Fig. 1** and **Supplementary Tables 3–6**).

Among the structural variations that were common to all the TKI-resistant samples, one in particular attracted our attention because it occurred in intron 2 of the *BIM* gene (**Fig. 1a**). This structural variation comprised an identical 2,903-bp genomic deletion that was common to all three resistant samples (**Fig. 1a–c**), suggesting that it was germline and polymorphic. After screening 2,597 healthy individuals, we found the deletion polymorphism to occur commonly in East Asian individuals (12.3% carrier frequency), but it was absent in individuals from African and European populations (0%) (**Supplementary Table 7**).

Functional effects of the *BIM* deletion polymorphism

Inspection of *BIM* gene structure suggested that the splicing of exon 3 and the splicing of exon 4 occur in a mutually exclusive manner because of the presence of a stop codon and a polyadenylation signal within exon 3 (**Fig. 2a** and **Supplementary Fig. 2a**)[16,17]. Indeed, sequencing of all identifiable *BIM* transcripts in CML cells confirmed that exons 3 and 4 never occurred in the same transcript (**Supplementary Fig. 2b**), consistent with prior reports[17]. Because of its close proximity (107 bp) to the intron-exon boundary at the 5′ end of exon 3, we hypothesized that the deletion polymorphism would result in preferential splicing of exon 3 over exon 4 (**Fig. 2a**)[18]. To determine whether this was the case, we constructed a minigene to assess whether the deletion leads to the preferential inclusion of exon 3 over exon 4 (**Fig. 2b**)[19] and found that the presence of the deletion favored splicing to exon 3 over exon 4 by at least fivefold (**Fig. 2c**). Notably, primary cells from individuals with CML showed the same phenomenon, as evidenced by the fact that polymorphism-containing

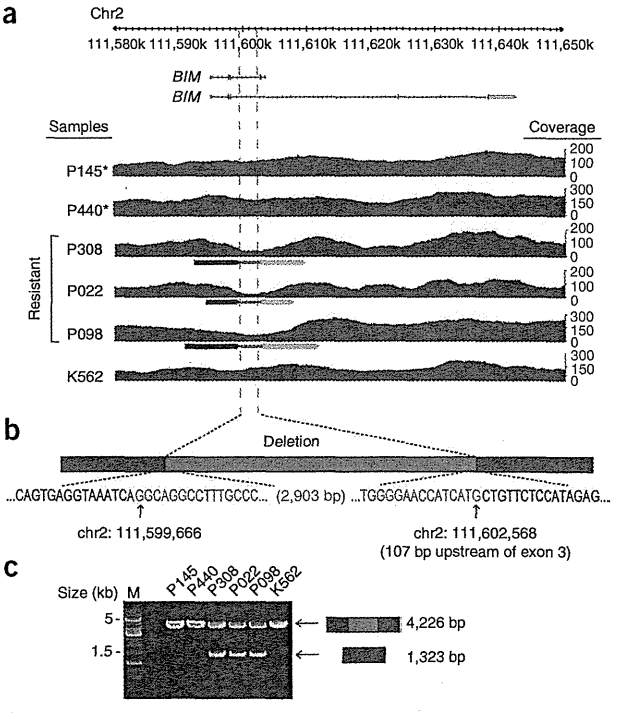

Figure 1 A 2,903-bp deletion polymorphism in intron 2 of *BIM* is present in TKI-resistant CML samples. (**a**) A Genome Browser view of the DNA–paired-end tag (PET) data encompassing chromosome 2 111,580,000–111,650,000 bp from the five clinical CML samples and K562 cells. Detection of the *BIM* deletion polymorphism by DNA-PET analysis in three of three samples from individuals with resistance to imatinib (P308, P022 and P098) but not in samples from subjects or cell lines that are sensitive to imatinib (P145, P440 and K562). The asterisks indicate that the samples (P145 and P440) were obtained from the same individual at presentation in chronic phase CML and when in major molecular remission, respectively. The red tracks represent the number of the sequenced concordant PETs that map to the region (coverage). The burgundy and pink horizontal arrowheads connected by green lines represent mapping regions of discordant PETs and indicate the presence of a deletion. The vertical dashed lines depict the deleted region. (**b**) Schematic depicting the intronic *BIM* deletion polymorphism and its flanking sequences. The breakpoints were identified by Sanger sequencing of PCR products. Deleted sequences are highlighted in blue. The human reference sequence coordinates are based on NCBI Build 36. (**c**) Agarose gel showing the PCR products from the five subject samples and K562 cells using primers that flanked the deletion. PCR products with a size of 4,226 bp and 1,323 bp correspond to the alleles without and with the deletion, respectively. The presence of both the 4,226-bp and 1,323-bp products indicates that the individual is heterozygous for the deletion polymorphism.

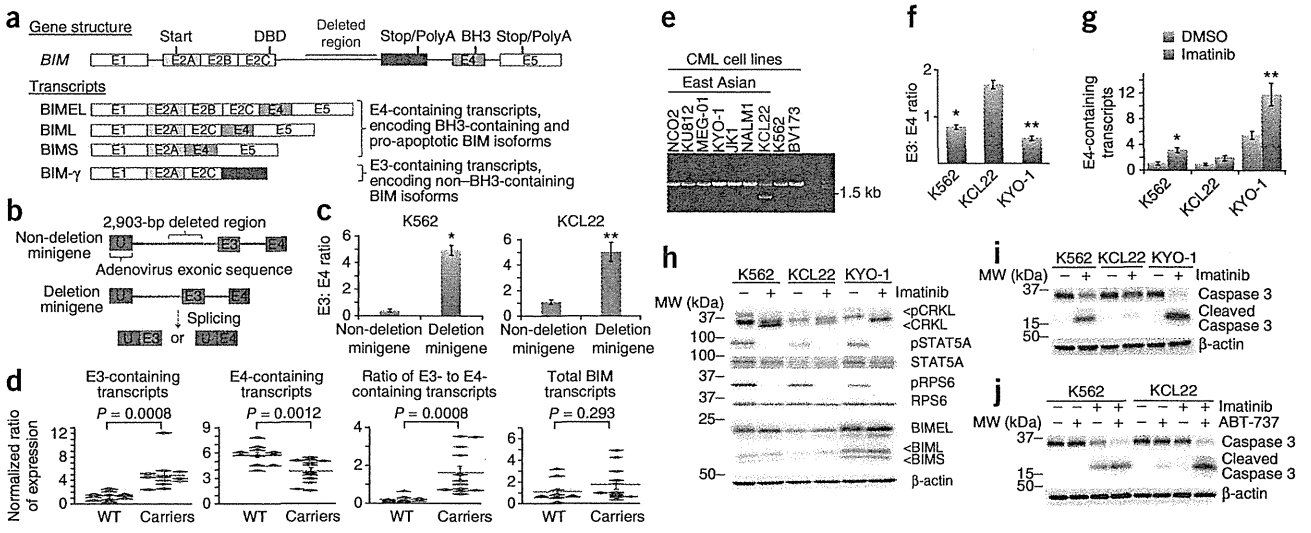

Figure 2 Effects of the deletion polymorphism on *BIM* gene function. (**a**) Genomic organization of *BIM* (top) showing exons for the major *BIM* transcript splice isoforms (bottom), including BIMEL, BIML and BIMS, as well as BIM-γ, which lacks the BH3 domain[17]. The deletion polymorphism between exons 2 and 3 is highlighted with a red line. The exons containing the start codon (start), the dynein-binding domain (DBD), the BH3 domain (BH3) and the stop codon and polyadenylation signal sequences (Stop/PolyA) are also highlighted. Exon 4 encodes for the BH3 domain that is required for BIM apoptotic function, whereas exon 3 lacks this domain. Because exon 3 and exon 4 undergo mutually exclusive splicing, exon-3–containing transcripts will not contain a BH3 domain. The diagram is not drawn to scale. E, exon. (**b**) Schematic of the two minigene constructs used for measuring splicing to exons 3 and 4. (**c**) The increased ratio of exon 3 to exon 4 transcripts in the non-deletion minigene construct compared to the deletion minigene construct in K562 cells (left) and in KCL22 cells (right). Data are mean ± s.e.m. $*P = 0.0002$, $**P = 0.012$ (Student's *t* test). (**d**) Expression of exon-specific transcripts of *BIM* in 23 samples from subjects with CML. $n = 11$ subjects without the deletion (WT), and $n = 12$ subjects with the deletion (carriers). The amounts of the various transcripts containing exons 2A, 3 or 4 are expressed as normalized ratios relative to exon 2A (for exons 3 and 4) or β-actin (ACTB, for exon 2A (total *BIM* transcripts). We measured exon 2A transcripts as a readout for all *BIM* transcripts, as exon 2A contains the start site and is present in all transcripts. The mean and s.e.m. are represented by the red lines and bars. The expressions for the one homozygous carrier are highlighted in green. Statistical significance was determined using the Wilcoxon rank-sum test. (**e**) Agarose gel of the PCR products, using the method described in **Figure 1c**, to detect the polymorphism in a collection of East Asian and non–East Asian CML cell lines. The KCL22 line carries the deletion polymorphism and is highlighted in red. (**f**) Ratio of exon-3– to exon-4–containing transcripts in CML cell lines with (KCL22) and without (K562 and KYO-1) the deletion polymorphism. Data are mean ± s.e.m. $*P = 0.016$, $**P = 0.011$ (Student's *t* test). (**g**) The expression of exon-4–specific transcripts of *BIM* (normalized to β-actin), as measured by quantitative PCR (qPCR) in cell lines with and without the deletion polymorphism treated with DMSO or imatinib. Data are mean ± s.e.m. $*P = 0.01$, $**P = 0.004$ (Student's *t* test) with respect to imatinib-treated KCL22 cells. (**h**) Western blot showing upregulation of BIM and the inhibition of signaling pathways downstream of BCR-ABL1 kinase in CML cell lines as a result of imatinib treatment. CRKL, v-crk sarcoma virus CT10 oncogene homolog (avian)-like; pCRKL, phosphorylated CRKL; STAT5A, signal transducer and activator of transcription 5A; pSTAT5A, phosphorylated STAT5A; RPS6, ribosomal protein S6; pRPS6, phosphorylated RPS6; MW, molecular weight. (**i**) Western blot showing caspase 3 cleavage in cells treated as in **h**. (**j**) Western blot showing caspase 3 cleavage in cell lines treated with imatinib and with or without the BH3-mimetic drug ABT-737.

samples had higher expression of exon-3– compared to exon-4–containing transcripts, whereas general *BIM* transcription was unaffected by the polymorphism (**Fig. 2d**). We observed similar results in lymphoblastoid cell lines obtained from normal healthy HapMap individuals, indicating that the polymorphism has a cell-lineage–independent effect (**Supplementary Fig. 2c**). Taken together, these results suggest that the 2.9-kb deleted region contains *cis* elements that suppress the splicing of *BIM* exon 3, which, in cells harboring the deletion, results in preferential splicing of exon 3 over exon 4.

Because the pro-apoptotic BH3 domain is encoded exclusively by exon 4 of *BIM* (**Fig. 2a**)[17] and is required for BIM's apoptotic function[20,21], our observations suggest a previously unidentified mechanism for TKI resistance. In this model, after TKI exposure, polymorphism-containing CML cells would favor the expression of exon-3– over exon-4–containing *BIM* transcripts, resulting in decreased expression of BH3-containing BIM isoforms and, consequently, impaired BH3-domain–dependent apoptosis. To facilitate the study of this issue, we identified a Japanese CML cell line, KCL22 (ref. 22), that contained the deletion (**Fig. 2e**) and confirmed that cells from the line expressed an increased ratio of exon 3 to exon 4 transcripts compared to cells without the deletion (**Fig. 2f**). KCL22 cells also showed a decreased induction of exon-4–containing transcripts after TKI exposure (**Fig. 2g**), as well as decreased concentrations of BIMEL protein, a major BH3-containing BIM isoform (**Fig. 2h**)[17]. Consistent with previous reports[22–24], KCL22 cells were resistant to imatinib-induced apoptosis (**Supplementary Fig. 2d**) and showed impaired apoptotic signaling after imatinib exposure despite effective BCR-ABL1 inhibition, as confirmed by a decrease in BCR-ABL1–dependent signaling (**Fig. 2h,i**)[25,26]. KCL22 cells were also exquisitely sensitive to induction of apoptosis after increased expression of exon-4–containing and, therefore, BH3-encoding (but not exon-3–containing) *BIM* isoforms (**Supplementary Fig. 2e**). This observation suggested that the impaired imatinib-induced apoptosis in KCL22 cells could be restored by the addition of BH3-mimetic drugs, which functionally mimic BH3-only proteins by binding and inhibiting pro-survival BCL2 family members[27]. As shown in **Figure 2j**, we found that this was indeed the case. In addition, we confirmed that siRNA-mediated knockdown of exon-3–containing transcripts did not sensitize KCL22 cells to imatinib, indicating that exon-3–containing isoforms probably do not have a role in TKI resistance (**Supplementary Fig. 2f–h**).

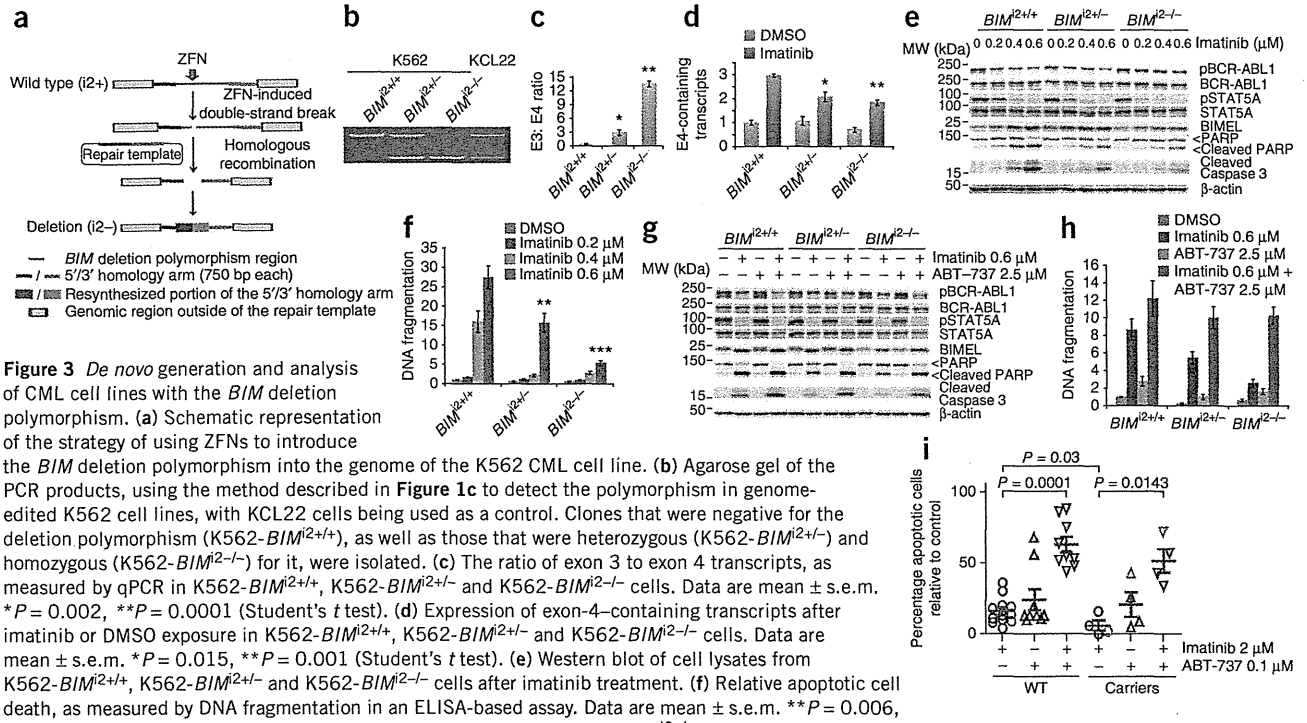

Figure 3 *De novo* generation and analysis of CML cell lines with the *BIM* deletion polymorphism. (**a**) Schematic representation of the strategy of using ZFNs to introduce the *BIM* deletion polymorphism into the genome of the K562 CML cell line. (**b**) Agarose gel of the PCR products, using the method described in **Figure 1c** to detect the polymorphism in genome-edited K562 cell lines, with KCL22 cells being used as a control. Clones that were negative for the deletion polymorphism (K562-*BIM*$^{i2+/+}$), as well as those that were heterozygous (K562-*BIM*$^{i2+/-}$) and homozygous (K562-*BIM*$^{i2-/-}$) for it, were isolated. (**c**) The ratio of exon 3 to exon 4 transcripts, as measured by qPCR in K562-*BIM*$^{i2+/+}$, K562-*BIM*$^{i2+/-}$ and K562-*BIM*$^{i2-/-}$ cells. Data are mean ± s.e.m. $*P = 0.002$, $**P = 0.0001$ (Student's *t* test). (**d**) Expression of exon-4–containing transcripts after imatinib or DMSO exposure in K562-*BIM*$^{i2+/+}$, K562-*BIM*$^{i2+/-}$ and K562-*BIM*$^{i2-/-}$ cells. Data are mean ± s.e.m. $*P = 0.015$, $**P = 0.001$ (Student's *t* test). (**e**) Western blot of cell lysates from K562-*BIM*$^{i2+/+}$, K562-*BIM*$^{i2+/-}$ and K562-*BIM*$^{i2-/-}$ cells after imatinib treatment. (**f**) Relative apoptotic cell death, as measured by DNA fragmentation in an ELISA-based assay. Data are mean ± s.e.m. $**P = 0.006$, $***P = 0.00021$ (Student's *t* test). Results were normalized to the K562-*BIM*$^{i2+/+}$ DMSO-treated control. (**g**) Western blot showing the effect of adding the BH3-mimetic drug ABT-737 to imatinib with respect to apoptotic signaling in K562-*BIM*$^{i2+/+}$, K562-*BIM*$^{i2+/-}$ and K562-*BIM*$^{i2-/-}$ cells. (**h**) Relative apoptotic cell death for cells treated as in **g** using the method described in **f**. Data are mean ± s.e.m. (**i**) Apoptotic cell death as measured by Annexin V staining. Primary CML cells were obtained from individuals without and with the deletion polymorphism. The mean and s.e.m. are represented by the red lines and bars. Results for the one homozygous carrier are highlighted in green. Statistical significance was determined using the Wilcoxon rank-sum test.

The *BIM* deletion and intrinsic TKI resistance in CML cells

We next used gene targeting facilitated by zinc finger nuclease (ZFN) to precisely recreate the deletion polymorphism in the *BIM* gene of imatinib-sensitive K562 CML cells (**Fig. 3a**). We then analyzed these cells for changes in *BIM* splicing and expression, as well as for TKI-induced apoptosis. We generated subclones that were heterozygous (K562-*BIM*$^{i2+/-}$) or homozygous (K562-*BIM*$^{i2-/-}$) for the deletion polymorphism (**Fig. 3b**). We confirmed an increased ratio of exon 3 to exon 4 transcripts (**Fig. 3c**), as well as a small but reproducible increase in BIM-γ protein expression (**Supplementary Fig. 3a**), in cells from both subclones in a polymorphism-dosage–dependent manner. We attribute the low expression of BIM-γ protein, even in the cells homozygous for the deletion polymorphism, to the relatively short half-life of BIM-γ (<1 h) (**Supplementary Fig. 3b**). Cells containing the deletion polymorphism also showed decreased induction of exon-4–containing transcripts after imatinib exposure (**Fig. 3d**), as well as impaired upregulation of BIMEL protein, diminished apoptotic signaling and decreased apoptotic cell death, as measured by DNA fragmentation in an ELISA-based assay (**Fig. 3e**,**f** and **Supplementary Fig. 3c**). As in KCL22 cells, the combination of the BH3 mimetic ABT-737 with imatinib enhanced the ability of the latter to activate apoptosis in polymorphism-containing cells (**Fig. 3g**,**h**). In parallel experiments, we re-expressed the most abundant BIM isoform, BIMEL, in polymorphism-containing cells treated with or without imatinib. Analogous to the effects seen with ABT-737 treatment, the forced expression of BIMEL enhanced the ability of imatinib to activate apoptosis in deletion-containing K562 cells (**Supplementary Fig. 3d**). We also found that primary CML cells obtained from subjects with the deletion polymorphism were less sensitive to imatinib-induced death compared to cells from individuals without the deletion and that the relative TKI resistance of the cells with the deletion could be overcome with the addition of ABT-737 (**Fig. 3i**). Taken together, our studies establish that the *BIM* deletion polymorphism impairs the apoptotic response to imatinib by biasing splicing away from BH3-containing *BIM* isoforms and that this bias is sufficient to render CML cells intrinsically resistant to imatinib. We also show that the apoptotic response to imatinib can be restored in polymorphism-containing cells by treatment with BH3-mimetic drugs.

The *BIM* deletion as a biomarker for TKI responses in CML

Next, we performed a retrospective analysis on the influence of the deletion polymorphism on TKI responses in East Asian subjects with CML. Using a group of newly diagnosed persons with chronic phase CML from two independent East Asian (Singapore and Malaysia or Japan) cohorts ($n = 203$), we compared the clinical responses to first-line therapy with a standard dose of imatinib (400 mg per day) in individuals with and without the deletion polymorphism. We classified the clinical responses according to the European LeukemiaNet (ELN) criteria (**Supplementary Table 8**)[5] and defined resistant individuals as 'suboptimal responders' or 'failures' per ELN criteria (which includes subjects who never achieve either a complete cytogenetic response or a 3-log decrease in BCR-ABL1 transcript levels), whereas sensitive individuals corresponded to ELN-defined 'optimal responders'. In both geographic cohorts, subjects with the deletion polymorphism were more likely to have resistant disease than sensitive disease compared to controls (**Table 1**). When analyzed together, the overall odds ratio for resistant disease among subjects with the deletion polymorphism compared to those without it was 2.94 ($P = 0.02$, 95% CI 1.17–7.43).

Table 1 Association of the *BIM* deletion polymorphism with clinical resistance to imatinib in subjects with CML

	No *BIM* deletion polymorphism % (*n*)	*BIM* deletion polymorphism % (*n*)	
Singaporean and Malaysian cohort (*n* = 138)			
Sensitive	51 (64)	33 (5)	OR = 2.73 (95% CI 0.87–8.57)
Resistant	49 (59)	67 (10)	*P* = 0.09
Japanese cohort (*n* = 65)			
Sensitive	43 (23)	17 (2)	OR = 3.52 (95% CI 0.69–18.00)
Resistant	57 (30)	83 (10)	*P* = 0.13
Combined cohorts OR (*n* = 203)			**OR = 2.94 (95% CI 1.17–7.43) *P* = 0.02**

Subjects with newly diagnosed chronic phase CML were analyzed according to their cohorts of origin (Singaporean and Malaysian or Japanese) and divided into those with and those without the *BIM* deletion polymorphism. Individuals were then classified as resistant ('suboptimal response' or 'failure' per ELN criteria) or sensitive ('optimal response' per ELN criteria) to imatinib. Statistical analysis testing for the association between the *BIM* deletion polymorphism and clinical resistance to imatinib was carried out using logistic regression on the individual cohort tables adjusting for any effects of age differences between groups with and without the *BIM* deletion polymorphism (**Supplementary Table 9**). The unadjusted odds ratio (OR) was 2.85 (*P* = 0.02, 95% CI 1.15–7.08). The statistics for the combined cohorts are shown in bold for visualization purposes and to distinguish these results from those of each individual cohort.

By comparison, we found no significant differences between the two groups with respect to other potential prognostic or confounding factors, including median time from diagnosis to initiation of imatinib treatment, Sokal score at diagnosis or prior treatment with interferon (**Supplementary Table 9**). We also noted that the majority of resistant subjects with the polymorphism subsequently did not respond to second-generation TKI therapy (**Supplementary Table 10**), a finding that is in line with the intrinsic resistance we observed in the cell lines.

TKI resistance in CML is most commonly associated with the acquisition of somatic mutations in the *BCR-ABL1* kinase domain, which can be found in up to 50% of resistant individuals in the chronic phase of disease[28]. However, because the deletion polymorphism is germline and is sufficient to cause intrinsic TKI resistance *in vitro* (**Fig. 3**), we predicted that such individuals would be resistant even in the absence of a kinase-domain mutation. Accordingly, we divided the subjects into the following three clinical groups: resistant without a *BCR-ABL1* mutation (group 1), resistant with a *BCR-ABL1* mutation (group 2) or sensitive (group 3). We found that individuals with the polymorphism, compared to those without, were more likely to be in group 1 than in groups 2 and 3 combined (odds ratio = 1.90, 95% CI 1.08–4.35) (**Supplementary Table 11**). These data provide a second clinical validation of our hypothesis.

The *BIM* deletion as a biomarker in EGFR NSCLC

We next validated the role of the *BIM* polymorphism in another kinase-driven cancer, EGFR NSCLC, in which sensitizing mutations in EGFR predict high response rates in individuals treated with EGFR inhibitors[29,30] and in which BIM expression is required for TKI sensitivity[11–13]. An additional and relevant aspect of this cancer is that it is particularly common in East Asian countries, where activating *EGFR* mutations can be found in up to 50% of NSCLCs (compared to 15% in the western countries[4]) and are enriched for among female East Asian nonsmokers[31–33].

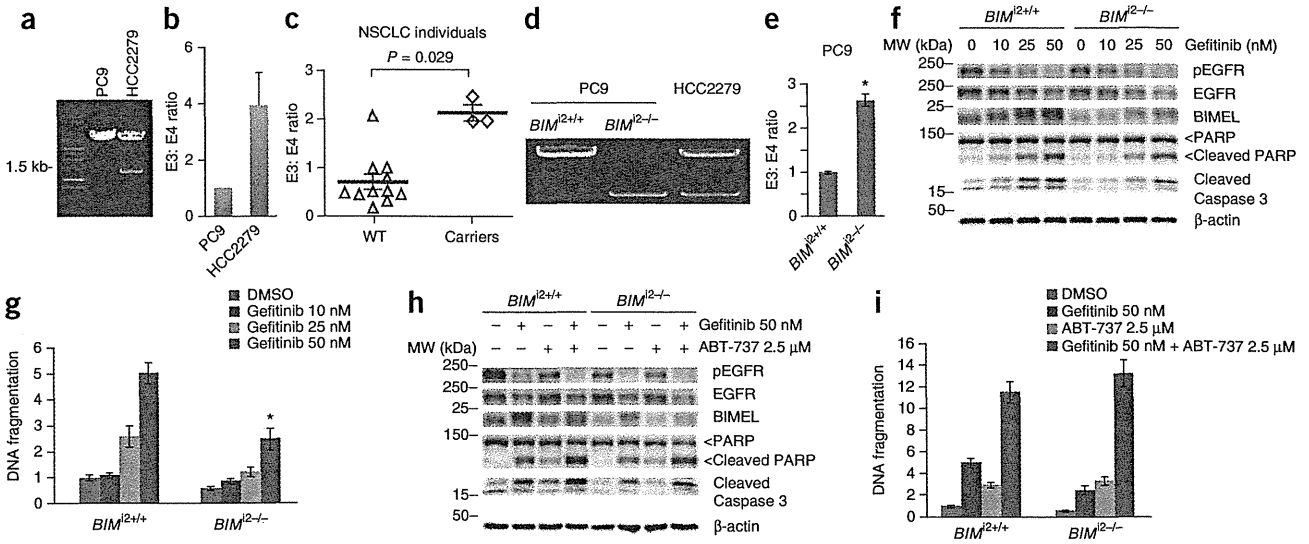

Figure 4 The *BIM* deletion polymorphism is sufficient to cause intrinsic TKI resistance in EGFR NSCLC cell lines. (**a**) An agarose gel of the PCR products, using the method described in **Figure 1c** to detect the polymorphism in HCC2279 cells. (**b**) Ratio of exon-3– to exon-4–containing transcripts in NSCLC cell lines with (HCC2279) and without (PC9) the deletion polymorphism. Data are mean ± s.e.m. (**c**) Peripheral blood mononuclear cells were obtained from subjects with EGFR NSCLC with and without the deletion and were analyzed for the ratio of exon 3 to exon 4 transcripts using qPCR, as described in **Figure 2d**. (**d**) An agarose gel of the PCR products, using the method described in **Figure 1c** to detect the polymorphism in genome-edited PC9 cell lines, with HCC2279 cells being used as a control. Clones that were negative for the deletion polymorphism (PC9-*BIM*$^{i2+/+}$), as well as those homozygous (PC9-*BIM*$^{i2-/-}$) for it, were isolated. (**e**) The ratio of exon 3 to exon 4 transcripts as measured by qPCR in PC9-*BIM*$^{i2+/+}$ and PC9-*BIM*$^{i2-/-}$cells. Data are mean ± s.e.m. *P = 0.0011 (Student's t test). (**f**) Western blots of cell lysates from PC9-*BIM*$^{i2+/+}$ and PC9-*BIM*$^{i2-/-}$ cells after treatment with increasing concentrations of gefitinib. (**g**) Relative apoptotic cell death, using the method described in **Figure 3f**, of PC9-*BIM*$^{i2+/+}$ and PC9-*BIM*$^{i2-/-}$ cells treated using DMSO or different concentrations of gefitinib, as indicated. Data are mean ± s.e.m. *P = 0.0009 (Student's t test). (**h**) Western blot of cell lysates from PC9-*BIM*$^{i2+/+}$ and PC9-*BIM*$^{i2-/-}$ cells treated with gefitinib, ABT-737 or both. (**i**) Relative apoptotic cell death, using the method described in **Figure 3f**, for cells treated as in **h**. Data are mean ± s.e.m.

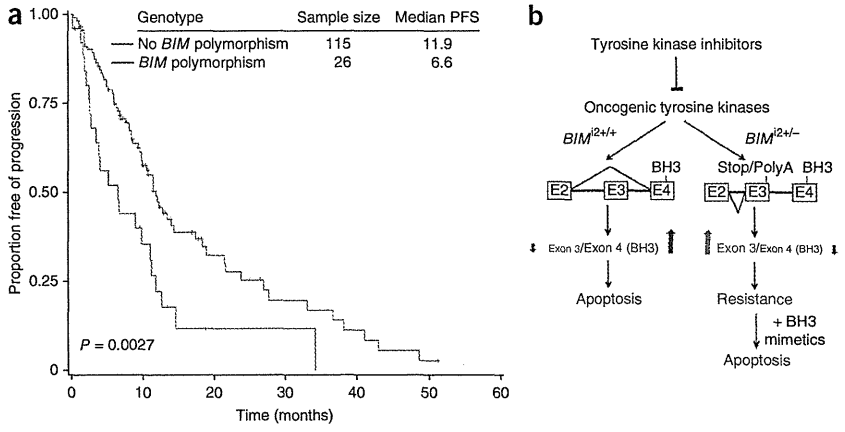

Figure 5 The *BIM* deletion polymorphism predicts shorter PFS in individuals with *EGFR*-mutant NSCLC treated with EGFR TKI therapy. (**a**) The presence or absence of the *BIM* deletion polymorphism was determined in 141 subjects with NSCLC from Singapore and Japan who were known to have activating mutations in EGFR and who received TKI therapy. The PFS for each group was estimated using the Kaplan-Meier method. (**b**) Schematic depicting the mechanism by which the *BIM* deletion polymorphism causes TKI resistance. After TKI exposure, wild-type cells that do not contain the deletion ($BIM^{i2+/+}$) preferentially upregulate expression of exon-4–containing (and, therefore, BH3-encoding) *BIM* transcripts that are capable of activating apoptosis (the red line corresponds to the 2,903-kb deleted region). In contrast, cells that harbor the deletion ($BIM^{i2+/-}$) favor the splicing and expression of exon-3–containing transcripts that do not encode the BH3 domain. The generation of exon-3–containing isoforms occurs at the expense of exon-4–containing isoforms, and as a result, the decreased concentrations of BH3-containing BIM protein isoforms render cells relatively TKI resistant. In these cells, restoration of TKI sensitivity can be brought about by the addition of BH3-mimetic drugs.

First we searched for NSCLC cell lines that harbored TKI-sensitizing *EGFR* mutations but were inexplicably TKI resistant (defined as lacking any of the known secondary-resistance–conferring mutations). We identified one such line, HCC2279, which notably fails to activate apoptosis despite effective EGFR inhibition[34,35]. We confirmed the presence of the *BIM* deletion polymorphism in the HCC2779 cells (**Fig. 4a**) and determined the effects of the deletion polymorphism on BIM function. The deletion resulted in greater expression of exon-3–containing compared to exon-4–containing (and, hence, BH3-containing) BIM isoforms compared to cells without the polymorphism (**Fig. 4b**). Notably, primary peripheral blood mononuclear cells from subjects with EGFR NSCLC, and with or without the deletion polymorphism, showed this same phenomenon (**Fig. 4c**). HCC2279 cells also had decreased induction of exon-4–containing transcripts and BIMEL protein after TKI exposure, as well as impaired activation of apoptotic signaling, as measured by poly (ADP-ribose) polymerase (PARP) cleavage (**Supplementary Fig. 4a,b**). Consistent with the notion that TKI resistance is a result of decreased concentrations of BH3-containing BIM protein, the addition of the BH3-mimetic drug ABT-737 enhanced TKI-induced apoptotic signaling and cell death (**Supplementary Fig. 4c,d**). To confirm that the polymorphism was sufficient to cause TKI resistance in EGFR NSCLC, we introduced it into TKI-sensitive PC9 cells (**Fig. 4d**). Analogous to our findings in K562-$BIM^{i2-/-}$ cells (**Fig. 3**), we found that, compared to PC9-$BIM^{i2+/+}$ cells, PC9-$BIM^{i2-/-}$ cells had decreased expression of exon-4–containing and BH3-containing BIM transcripts and protein, respectively, were intrinsically TKI resistant and were re-sensitized to TKIs by ABT-737 (**Fig. 4e–i**).

Next, we asked whether the deletion correlated with the duration of response to EGFR TKIs in subjects with NSCLC with activating *EGFR* mutations. Individuals with or without the deletion polymorphism did not differ with respect to known prognostic factors, including stage (as more than 85% of the subjects were stage IV) (**Supplementary Table 12**). Nevertheless, the presence of the polymorphism was predictive of a significantly shorter PFS, with a median PFS of 6.6 months in individuals with the polymorphism compared to 11.9 months for those without it ($n = 141$, $P = 0.0027$) (**Fig. 5a**). In multivariate analyses using the Cox regression model, only the deletion polymorphism (hazard ratio = 2.08, 95% CI 1.29–3.38, $P = 0.0028$) and the presence of the TKI-resistant exon 20 mutation (hazard ratio = 5.11, 95% CI 1.43–18.31, $P = 0.012$)[36,37] emerged as independent prognostic factors for shorter PFS.

DISCUSSION

Our findings demonstrate the principle that, although cancers should be classified according to their somatically acquired driver mutations, germline polymorphisms can directly modulate the responses of such cancers to targeted therapies and can strongly influence clinical outcomes. Notably, we show how a common *BIM* deletion polymorphism contributes to the heterogeneity of responses seen among molecularly defined patients with cancer who are treated with targeted therapies. Our data also highlight how a single germline polymorphism can strongly affect clinical outcomes in different cancers that share a common biology and probably reflect the central role of BIM in mediating TKI sensitivity in these diseases[8,11,13]. We anticipate that the list of cancers in which the *BIM* polymorphism influences TKI responses will expand to include others that also depend on BIM expression for TKI sensitivity[38–40].

The *BIM* polymorphism is found only in individuals of East Asian descent. It is therefore interesting to note that in CML, a higher rate of incomplete cytogenetic responses to imatinib has been reported among individuals in East Asia (~50%) compared to individuals in Europe and North America (26%)[41]. To assess the relative contribution of the deletion polymorphism to these ethnic differences, we estimated that the polymorphism underlies resistance in ~21% of East Asian patients (for the population attributable fraction, see the Online Methods). This might explain, in part, the differences in complete cytogenetic response rates observed between these two world populations.

As a germline biomarker for TKI resistance, the *BIM* polymorphism also offers several advantages over biomarkers comprising acquired mutations. First, the *BIM* polymorphism can be used at the time of initial presentation to predict which individuals are at an increased risk of developing TKI resistance, and second, the assessment of an individual's polymorphism status does not require an analysis of tumor-specific DNA. The former characteristic offers the potential for preventing the emergence of TKI resistance by therapeutic means (for example, treatment with a BH3-mimetic drug at the time of initial presentation or at the first sign of resistance), whereas the latter characteristic is particularly advantageous in solid tumor situations, as in EGFR NSCLC, when a second biopsy for tumor-specific tissue usually necessitates an invasive procedure. Although recent work has highlighted the value of *BIM* RNA levels in tumors before treatment in predicting TKI responsiveness[42], our discovery emphasizes the

importance of biomarkers that can also predict the induction of functional isoforms of BIM after TKI exposure.

By elucidating the effects of the deletion polymorphism on BIM function, we also were able to describe a previously unknown splicing mechanism by which the polymorphism contributes to drug resistance in CML and EGFR NSCLC (**Fig. 5b**). Thus, in showing that resistance is caused by impaired expression of BH3-containing BIM isoforms, we confirmed that pharmacologic restoration of BIM function could overcome this particular form of TKI resistance in both cancers. Our findings also support the increasingly recognized role of alterations in the splicing pattern of genes in human disease[43,44] and provide a new example of an inherited mutation that contributes to resistance against targeted cancer therapies. However, we note that although the presence of the deletion polymorphism is strongly associated with clinical TKI resistance and shorter PFS, other genetic factors, both acquired and inherited, will probably dictate the final response to TKI therapy in any individual patient. Indeed, several other mechanisms of EGFR-independent resistance have been described, including upregulated hepatocyte growth factor–dependent signaling[45], nuclear factor κ-light-chain-enhancer of activated B cells (NF-κB)-dependent signaling[46] and v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS) mutations[47]. It will therefore be crucial to determine how these factors interact with each other to contribute to TKI resistance, which, given the relatively low incidence of each individual contributor, will require larger prospective studies.

Clinical resistance to TKIs has been commonly classified as being primary or secondary, with the latter defined as occurring in individuals who experienced an initial response to TKI therapy and then later developed resistance. It is also assumed that secondary resistance is mediated by acquired somatic mutation(s) that emerge under the selective pressure of TKI therapy, whereas intrinsic mechanisms of resistance (including germline polymorphisms) are more likely to present with primary resistance and a lack of any upfront response. This line of reasoning is based on the assumption that resistance-conferring germline polymorphisms result in absolute as opposed to relative TKI resistance. However, by creating both CML and EGFR NSCLC cells with the deletion, we show that the *BIM* polymorphism results in relative TKI resistance. This finding is consistent with cancer cells being sensitive to small changes in BIM protein concentrations[8,48] and with BIM protein concentrations exerting a dose-dependent effect on apoptosis and on the degree of TKI resistance[8]. Accordingly, we expected to see some degree of response in TKI-treated subjects harboring the polymorphism, which we indeed confirmed in the setting of both CML and EGFR NSCLC.

Although our data focus on the effect of polymorphisms on therapeutic responses, it is possible that human polymorphisms also account for heterogeneity among other aspects of cancer biology. Unlike the *BIM* deletion, these other polymorphisms could conceivably result in enhanced therapeutic responses or could even cooperate with driver mutations to accelerate or delay cancer progression. As we have shown, a mechanistic understanding of how such polymorphisms affect gene function may lead to improved management of patients with cancer with respect to prognostication and therapy. In the case of TKI resistance in individuals with the *BIM* polymorphism, the addition of BH3 mimetics to the standard TKI therapy may allow for personalized treatment to overcome resistance or even to prevent its emergence. Finally, although the ethnic segregation of the polymorphism is in itself interesting, the greater importance of our findings may be that it is prototypic of other polymorphisms, yet to be discovered, that account for intrinsic drug resistance in different world populations.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemedicine/.

Accession codes. The sequencing data have been submitted to the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE28303 (clinical samples) and GSE26954 (K562), and were analyzed as described in the Online Methods.

Note: Supplementary information is available on the Nature Medicine website.

ACKNOWLEDGMENTS

This study was supported by grants from the National Medical Research Council of Singapore and the Biomedical Research Council (BMRC) of the Agency for Science, Technology and Research (A*STAR), Singapore. Additional support was also provided by the Genome Institute of Singapore internal research funds from the BMRC and the Department of Clinical Research, Singapore General Hospital. We are grateful for insightful conversations regarding this study with G. Bourque, M. Garcia-Blanco, E. Liu, X. Roca, S. Rosen, S. Shenolikar, D. Virshup and M. Voorhoeve. We thank C.-L. Wei and H. Thoreau for management of the sequencing platform, S.T. Leong, S.C. Neo and P.S. Choi for sequencing, J. Chen and C.S. Chan for help in data processing, H.P. Lim, Y.Y. Sia and Y.H. Choy for PCR validation and A. Lim and T.H. Lim for assistance in the fluorescence *in situ* hybridization (FISH) analysis. We also thank M. Garcia-Blanco (Duke University), K. Itahana (Duke-NUS), A. Vazquez (Institut National de la Santé et de la Recherche Médicale U.1014, Villejuif, France and Université Paris-Sud, Paris, France) and P. Koeffler (Cedars-Sinai Medical Center, Los Angeles, California, USA and Cancer Science Institute of Singapore, Singapore) for the kind gifts of the pl-12 vector, pcDNA3-FLAG3 plasmid, BIM expression vectors and NSCLC cell lines, respectively. Finally, we are grateful to the patients and physicians at the Department of Haematology, Singapore General Hospital, the Department of Hematology-Oncology, Akita University Hospital, Japan, the Toho University Omori Medical Center, Japan, the Aichi Cancer Center, Japan, the National University Cancer Institute, National University Health System, Singapore, National Cancer Centre, Singapore and the University of Malaya Medical Centre, Kuala Lumpur, Malaysia who contributed patient material.

AUTHOR CONTRIBUTIONS

K.P.N. and A.M.H. performed data analyses, generated the list of structural variations, validated the paired-end ditag data and wrote the first draft of the manuscript. C.T.H.C. provided CML clinical input and generated and analyzed the clinical data in **Table 1**. W.C.J. and T.K.K. devised and performed the experiments in **Figures 2–4**. C.-T.C. performed the experiments in **Figures 3** and **4**. J.W.J.H. performed FISH and PCR analysis on patient and normal control samples. A.S.M.T. and Y.F. constructed DNA-PET libraries for high-throughput sequencing. P.N.A., W.H.L. and W.-K.S. developed the bioinformatics pipeline for the DNA-PET analysis, N.N. contributed to the pipeline development, and X.Y.W. developed the copy number analysis. W.T.P. ran the bioinformatics pipeline. V.K. and A.T. performed *BIM* deletion screening in the HapMap samples, and A.T. performed the population-level genetic statistical analysis. X.R. managed the high-throughput sequencing, and A.S. managed the bioinformatics infrastructure. C.T.H.C., N.T., K.S., A.L.A., H.T.M., G.F.H., L.Y.Y., L.P.K., B.C., V.S.N., W.J.C., H.T., L.C.L. and Y.T.G. provided samples from patients with CML, as well as clinical data from the same patients. M.M.N. and T.Y.W. provided samples from normal individuals. K.P.N., J.W.J.H. and W.C.J. analyzed CML samples for the *BIM* deletion polymorphism. J.C.A. Jr. performed the statistical analysis of the CML clinical data. V.C.-R. performed and interpreted FISH data and provided scientific advice. S.S. compiled the clinical data and, together with J.C.A. Jr., performed the statistical analyses for **Figure 5a**. K.P.N., J.W.J.H., S.Z., D.P., P.T. and M.S. analyzed samples for *EGFR* mutations and the *BIM* deletion polymorphism. J.-E.S., M.-K.A., N.-M.C., Q.-S.N., D.S.W.T., K.I., Y.Y., H.M., E.H.T., R.A.S., T.M.C. and W.-T.L. provided samples from subjects with EGFR NSCLC, as well as the accompanying clinical data. Y.R. and S.T.O. designed and directed the study and analyzed data. S.T.O. wrote the final draft of the manuscript, which was reviewed by K.P.N., A.M.H., C.T.H.C., W.C.J., T.K.K., W.-T.L. and Y.R.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturemedicine/.

Published online at http://www.nature.com/naturemedicine/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Jänne, P.A., Gray, N. & Settleman, J. Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. *Nat. Rev. Drug Discov.* **8**, 709–723 (2009).
2. Carella, A.M. *et al.* New insights in biology and current therapeutic options for patients with chronic myelogenous leukemia. *Haematologica* **82**, 478–495 (1997).
3. Schiller, J.H. *et al.* Comparison of four chemotherapy regimens for advanced non–small-cell lung cancer. *N. Engl. J. Med.* **346**, 92–98 (2002).
4. Keedy, V.L. *et al.* American Society of Clinical Oncology provisional clinical opinion: epidermal growth factor receptor (EGFR) mutation testing for patients with advanced non–small-cell lung cancer considering first-line EGFR tyrosine kinase inhibitor therapy. *J. Clin. Oncol.* **29**, 2121–2127 (2011).
5. Baccarani, M. *et al.* Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *J. Clin. Oncol.* **27**, 6041–6051 (2009).
6. Wang, L., McLeod, H.L. & Weinshilboum, R.M. Genomics and drug response. *N. Engl. J. Med.* **364**, 1144–1153 (2011).
7. Youle, R.J. & Strasser, A. The BCL-2 protein family: opposing activities that mediate cell death. *Nat. Rev. Mol. Cell Biol.* **9**, 47–59 (2008).
8. Kuroda, J. *et al.* Bim and Bad mediate imatinib-induced killing of Bcr/Abl$^+$ leukemic cells, and resistance due to their loss is overcome by a BH3 mimetic. *Proc. Natl. Acad. Sci. USA* **103**, 14907–14912 (2006).
9. Aichberger, K.J. *et al.* Low-level expression of proapoptotic Bcl-2-interacting mediator in leukemic cells in patients with chronic myeloid leukemia: role of BCR/ABL, characterization of underlying signaling pathways, and reexpression by novel pharmacologic compounds. *Cancer Res.* **65**, 9436–9444 (2005).
10. Kuribara, R. *et al.* Roles of Bim in apoptosis of normal and Bcr-Abl–expressing hematopoietic progenitors. *Mol. Cell. Biol.* **24**, 6172–6183 (2004).
11. Cragg, M.S., Kuroda, J., Puthalakath, H., Huang, D.C. & Strasser, A. Gefitinib-induced killing of NSCLC cell lines expressing mutant EGFR requires BIM and can be enhanced by BH3 mimetics. *PLoS Med.* **4**, 1681–1689 (2007).
12. Gong, Y. *et al.* Induction of BIM is essential for apoptosis triggered by EGFR kinase inhibitors in mutant EGFR-dependent lung adenocarcinomas. *PLoS Med.* **4**, e294 (2007).
13. Costa, D.B. *et al.* BIM mediates EGFR tyrosine kinase inhibitor-induced apoptosis in lung cancers with oncogenic EGFR mutations. *PLoS Med* **4**, 1669–1679 (2007).
14. Fullwood, M.J., Wei, C.L., Liu, E.T. & Ruan, Y. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* **19**, 521–532 (2009).
15. Hillmer, A.M. *et al.* Comprehensive long-span paired-end-tag mapping reveals characteristic patterns of structural variations in epithelial cancer genomes. *Genome Res.* **21**, 665–675 (2011).
16. Liu, J.W., Chandra, D., Tang, S.H., Chopra, D. & Tang, D.G. Identification and characterization of Bimgamma, a novel proapoptotic BH3-only splice variant of Bim. *Cancer Res.* **62**, 2976–2981 (2002).
17. Adachi, M., Zhao, X. & Imai, K. Nomenclature of dynein light chain-linked BH3-only protein Bim isoforms. *Cell Death Differ.* **12**, 192–193 (2005).
18. Ladd, A.N. & Cooper, T.A. Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol.* **3**, reviews0008 (2002).
19. Carstens, R.P., McKeehan, W.L. & Garcia-Blanco, M.A. An intronic sequence element mediates both activation and repression of rat fibroblast growth factor receptor 2 pre-mRNA splicing. *Mol. Cell. Biol.* **18**, 2205–2217 (1998).
20. Cheng, E.H. *et al.* BCL-2, BCL-X(L) sequester BH3 domain-only molecules preventing BAX- and BAK-mediated mitochondrial apoptosis. *Mol. Cell* **8**, 705–711 (2001).
21. Huang, D.C. & Strasser, A. BH3-only proteins-essential initiators of apoptotic cell death. *Cell* **103**, 839–842 (2000).
22. Kubonishi, I. & Miyoshi, I. Establishment of a Ph1 chromosome-positive cell line from chronic myelogenous leukemia in blast crisis. *Int. J. Cell Cloning* **1**, 105–117 (1983).
23. Mahon, F.X. *et al.* Selection and characterization of BCR-ABL positive cell lines with differential sensitivity to the tyrosine kinase inhibitor STI571: diverse mechanisms of resistance. *Blood* **96**, 1070–1079 (2000).
24. Deininger, M.W., Goldman, J.M., Lydon, N. & Melo, J.V. The tyrosine kinase inhibitor CGP57148B selectively inhibits the growth of BCR-ABL–positive cells. *Blood* **90**, 3691–3698 (1997).
25. Shah, N.P. *et al.* Transient potent BCR-ABL inhibition is sufficient to commit chronic myeloid leukemia cells irreversibly to apoptosis. *Cancer Cell* **14**, 485–493 (2008).
26. Ly, C., Arechiga, A.F., Melo, J.V., Walsh, C.M. & Ong, S.T. Bcr-Abl kinase modulates the translation regulators ribosomal protein S6 and 4E–BP1 in chronic myelogenous leukemia cells via the mammalian target of rapamycin. *Cancer Res.* **63**, 5716–5722 (2003).
27. Cragg, M.S., Harris, C., Strasser, A. & Scott, C.L. Unleashing the power of inhibitors of oncogenic kinases through BH3 mimetics. *Nat. Rev. Cancer* **9**, 321–326 (2009).
28. La Rosée, P. & Hochhaus, A. Resistance to imatinib in chronic myelogenous leukemia: mechanisms and clinical implications. *Curr. Hematol. Malig. Rep.* **3**, 72–79 (2008).
29. Paez, J.G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
30. Lynch, T.J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139 (2004).
31. Shepherd, F.A. *et al.* Erlotinib in previously treated non–small-cell lung cancer. *N. Engl. J. Med.* **353**, 123–132 (2005).
32. Kim, E.S. *et al.* Gefitinib versus docetaxel in previously treated non–small-cell lung cancer (INTEREST): a randomised phase III trial. *Lancet* **372**, 1809–1818 (2008).
33. Park, K. & Goto, K. A review of the benefit-risk profile of gefitinib in Asian patients with advanced non-small-cell lung cancer. *Curr. Med. Res. Opin.* **22**, 561–573 (2006).
34. Lu, Y., Liang, K., Li, X. & Fan, Z. Responses of cancer cells with wild-type or tyrosine kinase domain–mutated epidermal growth factor receptor (EGFR) to EGFR-targeted therapy are linked to downregulation of hypoxia-inducible factor-1α. *Mol. Cancer* **6**, 63 (2007).
35. Machida, K. *et al.* Characterizing tyrosine phosphorylation signaling in lung cancer using SH2 profiling. *PLoS ONE* **5**, e13470 (2010).
36. Wu, J.Y. *et al.* Lung cancer with epidermal growth factor receptor exon 20 mutations is associated with poor gefitinib treatment response. *Clin. Cancer Res.* **14**, 4877–4882 (2008).
37. Sasaki, H. *et al.* EGFR exon 20 insertion mutation in Japanese lung cancer. *Lung Cancer* **58**, 324–328 (2007).
38. Gordon, P.M. & Fisher, D.E. Role for the proapoptotic factor BIM in mediating imatinib-induced apoptosis in a c-KIT-dependent gastrointestinal stromal tumor cell line. *J. Biol. Chem.* **285**, 14109–14114 (2010).
39. Will, B. *et al.* Apoptosis induced by JAK2 inhibition is mediated by Bim and enhanced by the BH3 mimetic ABT-737 in JAK2 mutant human erythroid cells. *Blood* **115**, 2901–2909 (2010).
40. Soda, M. *et al.* Identification of the transforming EML4-ALK fusion gene in non–small-cell lung cancer. *Nature* **448**, 561–566 (2007).
41. Au, W.Y. *et al.* Chronic myeloid leukemia in Asia. *Int. J. Hematol.* **89**, 14–23 (2009).
42. Faber, A. *et al.* BIM expression in treatment naive cancers predicts responsiveness to kinase inhibitors. *Cancer Discov.* **1**, 352–365 (2011).
43. Cartegni, L., Chew, S.L. & Krainer, A.R. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.* **3**, 285–298 (2002).
44. López-Bigas, N., Audit, B., Ouzounis, C., Parra, G. & Guigo, R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* **579**, 1900–1903 (2005).
45. Yano, S. *et al.* Hepatocyte growth factor induces gefitinib resistance of lung adenocarcinoma with epidermal growth factor receptor–activating mutations. *Cancer Res.* **68**, 9479–9487 (2008).
46. Bivona, T.G. *et al.* FAS and NF-κB signalling modulate dependence of lung cancers on mutant EGFR. *Nature* **471**, 523–526 (2011).
47. Takeda, M. *et al. De novo* resistance to epidermal growth factor receptor–tyrosine kinase inhibitors in EGFR mutation-positive patients with non-small cell lung cancer. *J. Thorac. Oncol.* **5**, 399–400 (2010).
48. Egle, A., Harris, A.W., Bouillet, P. & Cory, S. Bim is a suppressor of Myc-induced mouse B cell leukemia. *Proc. Natl. Acad. Sci. USA* **101**, 6164–6169 (2004).

ONLINE METHODS

Ethics committee approval. Clinical CML samples were obtained from patients seen at the Singapore General Hospital, the Akita University Hospital, the University of Malaya Medical Centre and the National University Cancer Institute, Singapore. German control samples were obtained from blood donors at the University Hospital of Bonn. Malay, Chinese and Indian control samples were derived from recent local population studies[49,50]. Clinical NSCLC samples were obtained from patients seen at the National Cancer Centre, Singapore, the Toho University Omori Medical Center, Japan, the Aichi Cancer Center, Japan and the National University Cancer Institute, National University Health System, Singapore. Written informed consent and institutional review board approval at the participating institutions were obtained from all patients and normal individuals who contributed samples to this study.

DNA-PET sequencing and structural variation detection. DNA-PET sequencing and clustering of discordant PETs (dPETs) for structural variation detection has been described in Hillmer *et al.*[51]. DNA-PET libraries with 5-, 7- and 9-kb DNA fragments (**Supplementary Table 2**) were sequenced using the SOLiD platform (Applied Biosystems). The sequencing data from this study have been submitted to NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE28303 for the five patient samples and under accession number GSE26954 for K562. The genomic region that was covered by the 5′ tags of a dPET cluster was defined as the 5′ anchor, and the genomic region that was covered by the 3′ tags of a cluster was defined as the 3′ anchor. dPET clusters with anchor regions <500 bp were excluded from further analyses. The DNA-PET sequencing of K562 has been described earlier[51]. The same K562 dPET clusters used previously were also used in the present study, but the centromeric regions were not excluded here, dPET clusters with anchor regions <500 bp were excluded here (previously clusters with anchor regions <1,000 bp were excluded), and a new exclusion and filtering procedure was applied here (**Supplementary Note** and **Supplementary Table 13**).

Genotyping of the *BIM* polymorphic deletion. *Determination of patient genotype.* We extracted genomic DNA from either patients' peripheral blood (for both CML and NSCLC) or from formalin-fixed paraffin-embedded (FFPE) biopsy slides and blocks (for NSCLC). For DNA extracted from blood samples, we genotyped the deletion in the samples by a single PCR reaction using the primers 5′-AATACCACAGAGGCCCACAG-3′ and 5′-GCCTGAAGGTGCTGAGAAAG-3′ and JumpStart RedAccuTaq LA DNA Polymerase (Sigma) with the following thermo cycling conditions: 96 °C for 30 s, (94 °C for 15 s, 60 °C for 60 s and 68 °C for 10 min) ×29 and 68 °C for 20 min. The resulting PCR products from the deletion (1,323 bp) and the wild-type (4,226 bp) alleles were analyzed on 1% agarose gels.

For DNA recovered from FFPE tissues, we performed two separate PCR reactions to determine the presence of the wild-type and deletion alleles. The wild-type allele was genotyped using the forward primer 5′-CCA CCAATGGAAAAGGTTCA-3′ and the reverse primer 5′-CTGTCATTTC TCCCCACCAC-3′. The deletion allele was genotyped using the forward primer 5′-CCACCAATGGAAAAGGTTCA-3′ and the reverse primer 5′-GGC ACAGCCTCTATGGAGAA-3′. We performed PCR reactions using GoTaq Hot start Polymerase (Promega) with the following thermo cycling conditions: 95 °C for 5 min, (95 °C for 50 s, 58 °C for 50 s and 72 °C for 1 min) ×39 and 72 °C for 10 min. The PCR products for the deletion (284 bp) and the wild-type (362 bp) alleles were analyzed on a 2% agarose gel and were sequenced.

Determination of population frequency. We used Affymetrix Genome-Wide Human SNP Array 6.0 intensity data downloaded from the HapMap[52] homepage (http://snp.cshl.org/) to infer the copy number of the deletion polymorphism in *BIM* for the HapMap samples. Two genotyped single nucleotide positions were located within the deletion: SNP_A-4195083 and CN_173550. The raw intensities of the two markers were used to call the copy number variation event using a Gaussian mixture model similar to the algorithm proposed by Korn and colleagues[53]. Using this procedure, we predicted the copy number and, therefore, the presence or absence of the deletion in unrelated HapMap samples of European (n = 60), Yoruban (n = 60) and Chinese or Japanese (n = 90) origin. We then genotyped the deletion in the Chinese and Japanese samples for which we had available DNA (n = 74) by PCR as described above with the following slight modifications of the thermo cycling conditions: 96 °C for 30 s, (94 °C for 15 s, 64 °C for 30 s and 68 °C for 5 min) ×12, (94 °C for 15 s, 60 °C for 30 s and 68 °C for 5 min) ×18 and 68 °C for 20 min. We used the PCR-based genotypes to refine the single-nucleotide intensity cutoffs for genotype calling in the European and Yoruban samples and used only the PCR-validated genotypes of the East Asian samples for frequency assessment.

To investigate further whether the deletion in the European population is at moderate frequency but has been missed by chance in the HapMap samples and to determine more precisely the deletion frequency in Asia, we genotyped by PCR assay 595 German, 600 Malay, 608 Chinese and 605 Indian samples.

Calculation of attributable fractions for the *BIM* deletion. To calculate the population attributable fraction (PAF) of treatment resistance in East Asian patients, we used PAF = (f(OR − 1))/(f(OR − 1) + 1), where f is the frequency of deletion carriers among patients (f = 0.133), and OR is the odds ratio of the deletion carriers between patients being resistant and patients being sensitive to TKI treatment (OR = 2.94).

FISH. We used Vysis LSI (Locus specific identifier) BCR/ABL1 dual-fusion translocation probes (Abbott Molecular) for detecting *BCR-ABL1*. The LSI BCR probe is labeled with SpectrumGreen, and the LSI ABL1 probe is labeled with SpectrumOrange. We treated cells with 0.75 M KCl for 15 min at 37 °C. After fixation, we dropped the nuclei on slides for FISH according to the manufacturer's instructions with slight modifications. Briefly, we dehydrated the slides in a co-denaturated alcohol series for 3 min at 75 °C, which was followed by an overnight hybridization at 37 °C. We evaluated FISH signals in 200 interphase nuclei using a fluorescence microscope (Olympus BX60) under 1,000× magnification.

Cell lines, culture and chemicals. We purchased CML lines from American Type Culture Collection (ATCC) (MEG-01 and KU812), the Japanese Collection of Research Bioresources (NCO2) and the German Collection of Microorganisms and Cell Cultures (KCL22, K562, KYO-1, JK1, BV173 and NALM1). NSCLC cells (PC9 and HCC2279) were a gift from P. Koeffler. We cultured cells in RPMI-1640 medium supplemented with penicillin/streptomycin, glutamine and 10% FBS and incubated them in a humidified incubator at 37 °C with 5% CO_2. Zinc-finger-nuclease–edited K562 and PC9 cells were generated and maintained in RPMI-1640 medium supplemented with penicillin/streptomycin, glutamine and 20% FBS. Drugs were dissolved in DMSO (50% for imatinib; 100% for gefitinib and ABT-737) and kept at −20 °C. We used 1 μM imatinib and 0.5 μM gefitinib for all experiments, unless otherwise indicated. The treatment time was 12 h (**Figs. 2g** and **3d**), 24 h (**Figs. 2h–j** and **4**) or 48 h (**Fig. 3e–i**).

Real-time PCR analysis of exon-specific *BIM* transcripts. We extracted total cellular RNAs using the RNeasy Mini Kit (Qiagen). RNA was reverse transcribed using Superscript III First-Strand Synthesis System (Invitrogen) and quantitatively assessed using the iQ5 Multicolor Real-Time Detection System (Bio-Rad) with a total reaction volume of 25 μl. Primers were annealed at 59 °C for 20 s, and the amplicon was extended at 72 °C for 30 s. The total number of cycles quantified was 40. Transcript levels of β-actin or exon 2A of *BIM* were used to normalize between samples. The following primers were used: *BIM* exon 2A (forward: 5′-ATGGCAAAGCAACCTTCTGATG-3′; reverse: 5′-GGCTCTGTCTGTAGGGAGGT-3′), *BIM* exon 3 (forward: 5′-CA ATGGTAGTCATCCTAGAGG-3′; reverse: 5′-GACAAAATGCTCAAGGA AGAGG-3′), *BIM* exon 4 (forward: 5′-TTCCATGAGGCAGGCTGAAC-3′; reverse: 5′-CCTCCTTGCATAGTAAGCGTT-3′) and β-actin (forward: 5′-GGAC TTCGAGCAAGAGATGG-3′; reverse: 5′-AGCACTGTGTTGGCGTACAG-3′).

RT-PCR and sequencing of *BIM* transcripts. To assess whether the splicing of *BIM* exons 3 and 4 are indeed mutually exclusive, we performed RT-PCR and sequenced all *BIM* transcripts that were amplified. Total cellular RNA extraction and reverse transcription was performed using the method described above. We used the forward primer 5′-ATGGCAAAGCAACCTTCTGA-3′ and the reverse primer 5′-TCAATGCATTCTCCACACCA-3′ to amplify all transcripts that contained exons 2 and 5. These primers were annealed at 57 °C for 30 s, and the amplicons were extended at 72 °C for 1 min. To amplify transcripts containing

exon 3, we used the forward primer 5′-TGACTCTCGGACTGAGAAACG-3′ and the reverse primer 5′-CCAAAGCACAGTGAAAGATCA-3′. These primers were annealed at 55 °C for 30 s, and the amplicons were extended at 72 °C for 30 s. All PCR products were cloned into pJET1.2/blunt vector (Fermentas) before they were sent for sequencing analysis.

Western blot. We used antibodies to the following to perform western blotting: BCR-ABL1 (#2802), pBCR-ABL1 (#2861), BIM (#2819), CRKL (#3182), pCRKL (#3181), CASPASE 3 (#9662), cleaved CASPASE 3 (#9661), STAT5A (#9310), pSTAT5A (#9359), ribosomal protein S6 (RPS6; #2317), pRPS6 (#2211), PARP (#9542), phospho-EGFR (Y1068,#2234) (all from Cell Signaling Technology), Flag-M2 clone and β-actin (#AC-15, Sigma). The antibody dilutions used were 1 in 1,000, except for pRPS6 (1 in 2,000) and β-actin (1 in 5,000). A BIM-γ-specific antibody was generated by a commercial entity (Open Biosystems). HRP-conjugated secondary antibodies were specific to rabbit (Sigma) or mouse IgG (Santa Cruz biotechnology). The protein bands on the membrane were visualized using the Western Lightning chemiluminescence reagent (PerkinElmer).

Minigene vector construction. We used the pI-12 splicing vector (a gift from M. Garcia-Blanco) to construct the minigene vectors pI-12-MUT and pI-12-WT, which contained and did not contain the deletion polymorphism, respectively. Briefly, *BIM* exon 4, together with a 659-bp sequence upstream of exon 4, was amplified from KCL22 genomic DNA using forward primer 5′-GCC GCTCGAGTCTCTCCATGTGGTGTTTG-3′ and reverse primer 5′-GCC GAAGCTTCCTCCTTGCATAGTAAGCGTT-3′. The PCR product was subcloned into the XhoI and HindIII sites in the pI-12 plasmid to generate an intermediate vector. *BIM* exon 3 and the upstream region with and without the deletion polymorphism were amplified from KCL22 genomic DNA using forward primer 5′-GCCGGATATCATGGAAGGAACTGACCTGGTG-3′ and reverse primer 5′-GCCGATCGATGTAGGAAACTGGGTGAATGGC-3′. The two PCR products (4,500 bp and 1,597 bp) were subcloned into the *EcoR*V and *ClaI* sites in the intermediate vector to obtain the pI-12-WT and pI-12-MUT constructs. The ratios of exon 3 to exon 4 transcripts in the transfected cells were obtained by qPCR using specific primers for the U-E3 and U-E4 transcripts. Transcript levels were normalized to the adenovirus exonic sequence (U). The following primers were used: adenovirus exon (forward: 5′-CGA GCTCACTCTCTTCCGC-3′; reverse: 5′-CTGGTAGGGTACCTCGCA-3′), U-E3 transcript (forward: 5′- CGAGCTCACTCTCTTCCGC-3′; reverse: 5′-CTCTA GGATGACTACTGGTAGGGT-3′) and U-E4 transcript (forward: 5′-CGAGC TCACTCTCTTCCGC-3′; reverse: 5′- CCTCATGGAAGCTGGTAGGGT-3′).

siRNA knockdown of E3-containing *BIM* transcripts. siRNAs against E3-containing *BIM* transcripts (BIM-γ siRNA1: 5′-CCACCAUAGUCAAGAUACA-3′; BIM-γ siRNA2: 5′-CAGAACAACUCAACCACAA-3′) and negative control siRNA (ON-TARGETplus Non-targeting siRNA #1) were purchased from Dharmacon Inc (Lafayette). Nucleofection was performed on KCL22 cells using Nucleofector Solution V (Lonza) in the presence of siRNAs.

Determination of protein stability and apoptotic activity of different BIM isoforms. We cloned the complementary DNA of different BIM isoforms (BIMEL, BIML, BIMS and BIM-γ; gifts from A. Vazquez) into the pcDNA3-FLAG3 vector (a gift from K. Itahana). We transfected 5 μg of plasmid into KCL22 (**Supplementary Fig. 2e**) or K562 cells (**Supplementary Fig. 3a,b,d**) by nucleofection. To determine apoptotic activity, we used Annexin V-FITC and 7-AAD staining and flow cytometry. To determine the stability of the BIM-γ and BIML proteins, we treated transfected cells with 50 μg/ml of cycloheximide 44 h after nucleofection. Then we harvested the cells at various time points after treatment (0, 0.5, 1, 3, 5 and 7 h), and we determined the stability of Flag-tagged BIM-γ or BIML by western blot using antibodies to the Flag epitope.

Creation of *BIM* deletion polymorphism by ZFNs. The ZFN was custom made by Sigma-Aldrich CompoZr TM ZFN Technology with the following binding and cleavage sites: 5′-**CCTTCCCTGGAA**-ctggga-**ATAGTGGGTGAGATAGTG**-3′ (with the binding site in bold and the cleavage site not bolded). The cleavage site is located 551 bp downstream of the 5′ end of the *BIM* deletion polymorphism region. The repair template contained only the two flanking homology arms but not the *BIM* deletion polymorphism region. The repair template was constructed using a PCR strategy that used the KCL22 genomic DNA as a template and the forward primer 5′-CATAAATACCACAGAGGCCCACAGC-3′ (corresponding to a site 619 bp upstream from the 5′ end of the *BIM* deletion polymorphism) and reverse primer 5′-CCCTCGAAGACACCTCTATTGGGAGGC-3′ (corresponding to a site 743 bp downstream of the 3′ end of the *BIM* deletion polymorphism). We subcloned the 1,362-bp PCR product into the vector pCR-Blunt II-TOPO (Invitrogen), and we confirmed the correct template by sequencing.

The repair template and ZFN-encoding plasmids were transfected into K562 cells using the protocol mentioned previously[54]. To generate genome-edited cells, PC9 cells were seeded at a density of 2×10^5 cells per well in a six-well plate 1 d before transfection. The cells were transfected with the repair template (6 μg) and ZFN-encoding plasmids (0.6 μg each) using Fugene HD (Promega, USA). One day later, the transfected PC9 cells were arrested at the G2 phase for 18 h with 0.2 μM vinblastine (Sigma, USA). The cells were released from G2 arrest by washing twice in PBS, re-plating in a new tissue culture plate and being allowed to recover for 72 h.

We isolated the genome-edited K562 and PC9 clones by dilution cloning. We diluted the transfected cells to a density of 2.5 cells/ml and seeded 200 μl of the diluted cells into each well of a 96-well plate. Clones that successfully amplified from each well were harvested, and the genomic DNA was isolated using a Qiagen DNEasy kit (Hilden).

We screened for clones having the *BIM* deletion polymorphism by PCR using primers annealing to the *BIM* intronic region outside of the repair template, an approach that ensured the repair template would not be amplified. We used the forward primer 5′- GGCCTTCAACCACTATCTCAGTGCAATGG-3′ (corresponding to a site 1,507 bp upstream from the 5′ end of the *BIM* deletion polymorphism) and the reverse primer 5′-GGTTTCAGAGACAGAGCTGGG ACTCC-3′ (corresponding to a site 767 bp downstream of the 3′ end of the *BIM* deletion polymorphism) for PCR.

ELISA-based DNA fragmentation detection and western blotting on genome-edited K562 and PC9 clones. The presence of mono- and oligo-nucleosomes in the apoptotic cells was detected using the Cell Death Detection ELISA (Roche), following the manufacturer's instructions. Genome-edited K562 cells were seeded at a density of 2×10^5 cells/ml. Five milliliters or 0.5 ml of cells were used for western blot or the apoptotic assay, respectively. The cells were harvested 48 h after treatment. Genome-edited PC9 cells were seeded at a density of 5×10^4 cells/ml. Ten milliliters or 0.5 ml of cells were used for western blot or the apoptotic assay, respectively. The cells were harvested 24 h after treatment.

Apoptosis assay in primary CML samples. We measured apoptosis in primary CML samples using the Annexin V-FITC kit (Beckman Coulter, IN) with propidium iodine, following the manufacturer's instructions. Statistical significance was determined using a one-tailed Wilcoxon rank sum test, as deletion-containing cells are expected to be more resistant than non–deletion-containing cells.

Trypan blue assay. PC9 and HCC2279 cells (5×10^5 and 1.6×10^5 cells, respectively) were seeded in triplicate and treated for 48 h. The cells were trypsinized, and the number of viable cells was determined by trypan blue exclusion.

Mutation analysis for EGFR. FFPE slides of lung tumors were deparaffined by washing the slides in xylene and absolute ethanol. Lung cancer regions from each slide were scraped and transferred into a 1.5-ml tube, and genomic DNA was extracted using a QIAamp FFPE Tissue kit (Qiagen). EGFR exons 18–21 were sequenced. Fifty nanograms of FFPE genomic DNA was amplified by PCR in a 20 μl reaction volume containing 10 μl of GoTaq hot start Taq colorless master mix (M5133, Promega) in the following PCR conditions: 95 °C for 5 min, DNA amplification for 35 cycles at 95 °C for 50 s, 58 °C for 50 s, 72 °C for 60 s and a final extension at 72 °C for 10 min. The PCR primers used were: exon 18 (forward: 5′-TGGCACTGCTTTCCAGCATGG-3′; reverse: 5′-CTCCCCACCAGACCATGAGAGG-3′), exon 19 (forward: 5′-ATC ACTGGGCAGCATGTGGCA-3′; reverse: 5′-CCTGAGGTTCAGAGCCAT GGAC-3′), exon 20 (forward: 5′-CATGCGAAGCCACACTGACGTG-3′; reverse: 5′-GCATGTGAGGATCCTGGCTC-3′) and exon 21 (forward: 5′-GA TCTGTCCCTCACAGCAGG-3′; reverse: 5′-GGTGTCAGGAAAATGCTGG

CTG-3′). PCR products were purified by Exonuclease I (M0293L, New England Biolabs) and Shrimp Alkaline Phosphatase (Promega) treatments. Purified PCR products were sequenced in the forward and reverse directions using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Version 3) on an ABI PRISM 3730 Genetic Analyzer (Applied Biosystems, Foster City, California). Chromatograms were analyzed by SeqScape V2.5 and manual review.

Statistical analysis for PFS of patients with EGFR NSCLC. The primary endpoint in this study was to examine the effect of the *BIM* deletion polymorphism on the PFS of patients with EGFR-NSCLC from East Asian countries who were treated with EGFR TKIs. We calculated the PFS from the initiation of EGFR TKI therapy until either tumor progression or death from any cause. Observations were censored if TKI therapy was stopped because of side effects or if treatment was ongoing at the time of the analysis. We calculated the *P* values of the Kaplan-Meier test comparing survival curves using the Wilcoxon test. We used the *t* test and Fisher's exact test to test for differences between clinical characteristics of the *BIM*-deleted and wild-type populations. Using Cox proportional hazard regression analyses, univariate and multivariate hazard ratios were generated for the following factors: age, gender, histology, smoking history, type of EGFR mutation by exon and specific mutation, stage, first- or second-line TKI therapy, race, country (Japan or Singapore), TKI (gefitinib or erlotinib) and ECOG status. The significance level for entering variables in a stepwise regression was 0.05. We used the SAS System for Windows Version 9.2 LIFETEST, TTEST and PHREG procedures to perform the calculations.

49. Foong, A.W. *et al.* Rationale and methodology for a population-based study of eye diseases in Malay people: the Singapore Malay eye study (SiMES). *Ophthalmic Epidemiol.* **14**, 25–35 (2007).
50. Lavanya, R. *et al.* Methodology of the Singapore Indian Chinese Cohort (SICC) eye study: quantifying ethnic variations in the epidemiology of eye diseases in Asians. *Ophthalmic Epidemiol.* **16**, 325–336 (2009).
51. Hillmer, A.M. *et al.* Comprehensive long-span paired-end-tag mapping reveals characteristic patterns of structural variations in epithelial cancer genomes. *Genome Res.* **21**, 665–675 (2011).
52. Frazer, K.A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
53. Korn, J.M. *et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.* **40**, 1253–1260 (2008).
54. Urnov, F.D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).

Published OnlineFirst May 21, 2012; DOI:10.1158/2159-8290.CD-12-0009

REVIEW

ALKoma: A Cancer Subtype with a Shared Target

Hiroyuki Mano

ABSTRACT Anaplastic lymphoma kinase (ALK) is a receptor-type protein tyrosine kinase that is currently the focus of much attention in oncology. ALK is rendered oncogenic as a result of its fusion to NPM1 in anaplastic large cell lymphoma, to TPM3 or TPM4 in inflammatory myofibroblastic tumor, to EML4 in non-small cell lung carcinoma, and to VCL in renal medullary carcinoma. It is also activated as a result of missense mutations in neuroblastoma and anaplastic thyroid cancer. Whereas these various tumors arise in different organs, they share activated ALK, and a marked clinical efficacy with ALK inhibitors has already been shown for some of the tumors with ALK fusions. One of such compound, crizotinib, is now approved in the United States for the treatment of lung cancer positive for *ALK* rearrangement. I propose that tumors carrying abnormal ALK as an essential growth driver be collectively termed "ALKoma."

Significance: ALK acquires transforming ability through gene fusion or missense mutation in a wide range of human cancers. Some of these cancers, which I propose be collectively referred to as "ALKoma," may all be effectively treated with small compounds or antibodies targeted to activated ALKs. *Cancer Discov; 2(6); 495–502. ©2012 AACR.*

INTRODUCTION

Anaplastic lymphoma kinase (ALK) is a protein tyrosine kinase (PTK) that possesses a single transmembrane domain and consists of 1,620 amino acid residues in humans (Fig. 1; refs. 1, 2). The extracellular region of ALK contains 2 MAM (meprin, A5 protein, and protein phosphatase μ) domains (3) and a putative ligand-binding domain. ALK is relatively isolated in the phylogenetic tree of the PTK superfamily, with its kinase domain sharing the highest sequence identity with that of the leukocyte receptor tyrosine kinase (LTK; 79% identity) and that of the proto-oncoprotein ROS1 (50% identity).

In the mouse, expression of ALK is prominent in the brain and peripheral nervous system of developing embryos, but it decreases rapidly after birth. In the adult human, ALK is expressed at a low level in the central nervous system but is not detected in other tissues (4). Given these expression profiles, ALK has been thought to play an important role in the development or orchestration of the nervous system. Surprisingly, however, homozygous deletion of *Alk* in mice does not give rise to any apparent anomalies in morphology, internal organs, or fertility, suggesting that the gene is not essential for mouse development (5). A more detailed analysis of *Alk* knockout mice revealed an increase in the number of hippocampal progenitor cells in the brain and an enhanced performance in a test of novel object recognition. Another strain of such mice manifested increased ethanol consumption compared with heterozygous controls (6). These observations thus suggest that ALK may contribute to behavioral control in adult mice.

The structure of ALK, together with its expression pattern, suggests that it may function as a cell surface receptor for specific ligands that may regulate the proliferation or differentiation of neural cells. In the fruit fly, jelly belly (jeb) binds and activates an ortholog of ALK (7), but a mammalian ortholog of jeb has not been identified. The growth factors pleiotrophin and midkine have been proposed as candidate ligands for mammalian ALK (8, 9), but this notion remains controversial (10), leaving the biologic relevance of pleiotrophin and midkine unclear.

ALK first emerged in the field of oncology in 1994 as a result of the identification of its fusion to nucleophosmin (NPM1) in anaplastic large cell lymphoma (ALCL; Figs. 1 and 2; refs. 11, 12). The NPM1 portion mediates constitutive dimerization and consequent activation of the NPM1-ALK fusion protein (13). In addition to NPM1-ALK, various other ALK fusion proteins have now been identified in ALCL, including TFG-ALK, ATIC-ALK, and CLTC-ALK. The next disease linked to ALK was inflammatory myofibroblastic

Author's Affiliations: Division of Functional Genomics, Jichi Medical University, Tochigi; Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; and CREST, Japan Science and Technology Agency, Saitama, Japan

Corresponding Author: Hiroyuki Mano, Division of Functional Genomics, Jichi Medical University, 3311-1 Yakushiji, Shimotsukeshi, Tochigi 329-0498, Japan. Phone: 81-285-58-7449; Fax: 81-285-44-7322; E-mail: hmano@jichi.ac.jp

doi: 10.1158/2159-8290.CD-12-0009

Published OnlineFirst May 21, 2012; DOI:10.1158/2159-8290.CD-12-0009

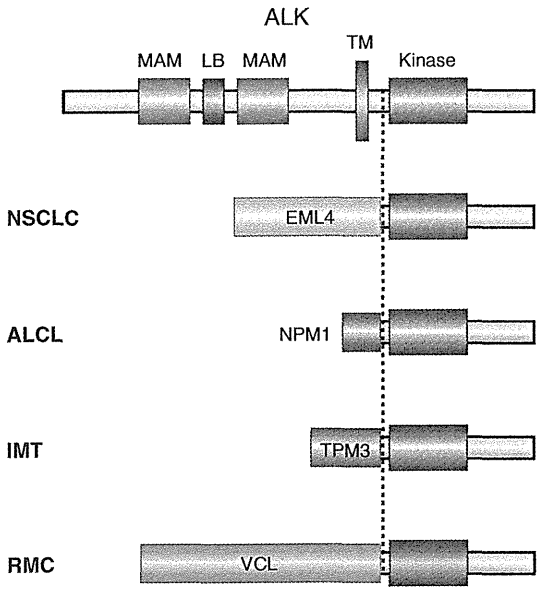

Figure 1. Structural organization of ALK and its fusion proteins. ALK is a receptor-type PTK with a single transmembrane (TM) domain as well as 2 MAM domains and a putative ligand-binding (LB) domain in the extracellular region. The intracellular region contains a kinase domain and is fused to EML4 in NSCLCs, to NPM1 in ALCLs, to TPM3 in IMTs, and to VCL in RMCs.

tumor (IMT), with increased expression of ALK and a rearranged *ALK* locus (chromosome band 2p23) being detected in a subset of IMT cases. A more detailed analysis identified *TPM3-ALK* and *TPM4-ALK* as fusion genes in IMT (Figs. 1 and 2; ref. 14). Further screening of IMT specimens revealed additional *ALK* fusions including *RANBP2-ALK* and *CARS-ALK*. Interestingly, some of these fusion genes, such as *TPM3/4-ALK* and *CLTC-ALK*, have been identified in both ALCLs and IMTs. The presence of *TPM3/4-ALK* was also reported in esophageal squamous cell carcinoma (15) and renal cancer (16).

In 2007, interest in ALK and the therapeutic potential of its specific inhibitors was boosted in response to the discovery of another fusion gene, *EML4-ALK*, this time in non–small cell lung carcinoma (NSCLC; ref. 17). Such interest was further increased the next year by the identification of activating point mutations in *ALK* in cases of neuroblastoma (18–21). In this review, I focus on activating genetic changes of *ALK* relevant to human cancer. Other aspects of *ALK* alterations (such as overexpression) in cancer have been elegantly reviewed elsewhere (22, 23).

EML4-ALK

EML4-ALK in NSCLC

Chronic myeloid leukemia (CML) is characterized by the presence of a fusion-type PTK, BCR-ABL1, that is generated as a result of a balanced chromosomal translocation, t(9;22). The remarkable therapeutic efficacy of the ABL1 inhibitor imatinib in individuals with this condition (24) suggested that targeting of the essential growth drivers in different types of cancer is a promising treatment strategy. To identify such growth drivers in clinical specimens, we developed a sensitive functional screening system based on retroviral cDNA expression libraries. The application of this approach to a lung adenocarcinoma specimen resulted in the discovery of *EML4-ALK* as a fusion-type oncogene (Figs. 1 and 2; ref. 17).

EML4 and *ALK* loci both map to the short arm of human chromosome 2 in opposite orientations and are separated by a distance of approximately 12 Mbp. A small inversion, inv(2)(p21p23), affecting both loci gives rise to the fusion gene. *EML4-ALK* was the first recurrent fusion-type oncogene in NSCLC, and, together with *ETS* fusions in prostate cancer (25), its existence argues against the previous notion that oncogenesis mediated by chromosome rearrangement is relatively specific to hematologic malignancies and sarcomas (rather than epithelial tumors).

EML4-ALK encodes a protein consisting of an amino-terminal portion of EML4 fused to the intracellular portion of ALK. EML4-ALK undergoes constitutive dimerization mediated through the coiled-coil domain of the EML4 portion and thereby acquires transforming ability in a manner

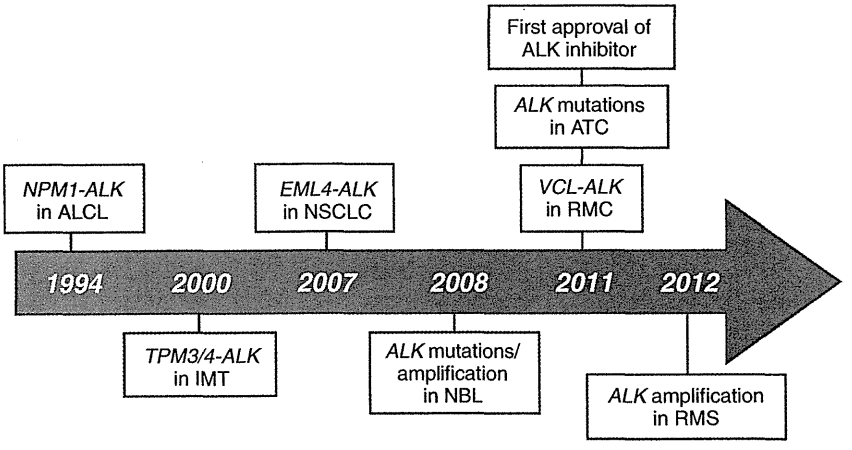

Figure 2. History of ALK in oncology. The first discovery of an oncogenic mutant of *ALK* was that of *NPM1-ALK* associated with ALCLs in 1994, which was followed by the identification of *TPM3/4-ALK* associated with IMTs, *EML4-ALK* associated with NSCLCs, mutated/amplified *ALK* associated with neuroblastoma (NBL), *VCL-ALK* associated with RMCs, mutated *ALK* associated with ATCs, and amplified *ALK* associated with rhabdomyosarcoma (RMS). Whereas several ALK inhibitors are currently in clinical trials, one such compound, crizotinib, was approved as a therapeutic drug for *ALK*-rearranged NSCLCs in 2011.

Published OnlineFirst May 21, 2012; DOI:10.1158/2159-8290.CD-12-0009

dependent on the associated upregulation of PTK activity. Indeed, several cell lines positive for *EML4-ALK* were found to be dependent on ALK catalytic activity for their proliferation (26, 27). Furthermore, transgenic mice that express *EML4-ALK* in lung type II alveolar cells manifest hundreds of adenocarcinoma nodules in both lungs soon after birth, and treatment of the animals with an ALK inhibitor results in the rapid clearance of these nodules (28). Such successful treatment of *EML4-ALK* transgenic mice with an ALK inhibitor was subsequently confirmed independently (29). These data suggested that EML4-ALK is the essential growth driver for NSCLC positive for this fusion kinase, and that targeting of ALK activity may therefore prove therapeutically effective in the clinic.

Clinicopathologic Features and Diagnosis of *EML4-ALK*-Positive Tumors

EML4-ALK is present in 3% to 6% of NSCLC cases (30–32). Of note, the clinical characteristics of patients with NSCLCs positive for *EML-ALK* are similar to those of such individuals who harbor activating mutations in the EGF receptor gene (*EGFR*): Both groups of patients thus tend to manifest an adenocarcinoma histology and to be non- or light smokers. However, the presence of *EML4-ALK* and that of *EGFR* or *KRAS* mutations are mutually exclusive, albeit with rare exceptions (33). Furthermore, whereas NSCLC with *EGFR* mutations is more prevalent in Asian populations than in Caucasians, no such ethnic differences have been reported for NSCLCs with *EML4-ALK*. With regard to pathologic characteristics, EML4-ALK–positive NSCLC has been found to contain signet-ring cells or to manifest a mucinous cribriform pattern (34–36), but other pathologic subtypes of NSCLCs also may harbor EML4-ALK. *EML4-ALK* may be infrequently present in other types of human cancers (16, 37).

The original *EML4-ALK* we identified resulted from ligation of intron 13 of *EML4* to intron 19 of *ALK* (17), but many other variants of the fusion gene have since been described. The vast majority of *ALK* translocations, including *EML4-ALK*, *NPM1-ALK*, and *TPM3-ALK*, involve intron 19 of *ALK*. Theoretically, ligation of *EML4* introns 1, 2, 6, 13, 18, 20, or 21 to intron 19 of *ALK* should generate an in-frame fusion at the mRNA level. Furthermore, given that exon 2 of *EML4* encodes the coiled-coil domain essential for EML4-ALK activation, all such fusions with the exception of that of *EML4* intron 1 would be expected to generate an oncogenic kinase. Large-scale screening of clinical specimens has revealed that the original *EML4-ALK* variant (referred to as E13;A20 or variant 1) together with a variant in which intron 6 of *EML4* is fused to intron 19 of *ALK* (referred to as E6;A20 or variant 3; ref. 26) account for more than 90% of all *EML4-ALK*–positive cases of NSCLCs in Japan (M. Soda, personal communication), but many rare variants have also been identified (17, 26, 27, 30, 31, 34, 37, 38).

Importantly, *EML4* exons theoretically unable to undergo in-frame fusion to exon 20 of *ALK* may be involved in the generation of oncogenic EML4-ALK proteins. For instance, an *EML4-ALK* cDNA in which exon 14 of *EML4* is fused in-frame to the 13th nucleotide of *ALK* exon 20 has been identified (30). In addition, another variant in which the nucleotide located 19 bp upstream of the 3′ end of *EML4* exon 15 is fused to that located 20 bp downstream of the 5′ end of *ALK* exon 20 was described (27). It is therefore important that clinics adopt diagnostic techniques, such as those based on multiplex reverse transcription PCR (RT-PCR), that are able to capture all these variants of *EML4-ALK* reliably.

Currently, no single technique may be suitable for the analysis of all specimen types. Immunohistochemistry and FISH, for example, are applicable to formalin-fixed, paraffin-embedded (FFPE) tissue specimens (32, 34, 35) but may not be suitable for the analysis of sputum, pleural effusion, bronchial lavage fluid, or frozen tissue. Conversely, the latter specimen types are readily examined by RT-PCR (30), whereas the former ones may not be suitable for this technique. I thus propose that diagnostic tools for the detection of EML4-ALK should be selected on the basis of the available specimen types. FISH and immunohistochemistry should be applied to FFPE tissue samples, whereas multiplex RT-PCR is appropriate for the other specimen types.

Crizotinib

The first ALK inhibitor to enter clinical trials was crizotinib, which is also known as PF-02341066 and is actually a dual inhibitor for both ALK and MET kinases (Table 1; ref. 39). At the time of the discovery of *EML4-ALK*, crizotinib had already entered a phase I trial that mainly targeted digestive tract cancers positive for *MET* amplification. After the report of *EML4-ALK*, the phase I trial with crizotinib was expanded to include tumors positive for ALK rearrangement, and the drug soon proved as therapeutically efficacious for such tumors.

Table 1. ALK inhibitors under clinical trials or already approved

Inhibitor	Company	Trial phase	References
Crizotinib	Pfizer	Phase III (approved in United States, South Korea, and Japan)	39, 68
CH5424802	Chugai Pharmaceutical	Phase I/II	62
ASP3026	Astellas Pharma	Phase I	Not available
LDK378	Novartis	Phase I	Not available
AP26113	Ariad	Phase I/II	61