doi:10.1038/nature07261

ARTICLES

Identification of *ALK* as a major familial neuroblastoma predisposition gene

Yaël P. Mossé[1], Marci Laudenslager[1], Luca Longo[2], Kristina A. Cole[1], Andrew Wood[1], Edward F. Attiyeh[1], Michael J. Laquaglia[1], Rachel Sennett[1], Jill E. Lynch[1], Patrizia Perri[2,3], Geneviève Laureys[4], Frank Speleman[4], Cecilia Kim[5], Cuiping Hou[1,5], Hakon Hakonarson[5,7], Ali Torkamani[6], Nicholas J. Schork[6], Garrett M. Brodeur[1], Gian P. Tonini[2], Eric Rappaport[1], Marcella Devoto[7,8] & John M. Maris[1,9]

Neuroblastoma is a childhood cancer that can be inherited, but the genetic aetiology is largely unknown. Here we show that germline mutations in the anaplastic lymphoma kinase (*ALK*) gene explain most hereditary neuroblastomas, and that activating mutations can also be somatically acquired. We first identified a significant linkage signal at chromosome bands 2p23–24 using a whole-genome scan in neuroblastoma pedigrees. Resequencing of regional candidate genes identified three separate germline missense mutations in the tyrosine kinase domain of *ALK* that segregated with the disease in eight separate families. Resequencing in 194 high-risk neuroblastoma samples showed somatically acquired mutations in the tyrosine kinase domain in 12.4% of samples. Nine of the ten mutations map to critical regions of the kinase domain and were predicted, with high probability, to be oncogenic drivers. Mutations resulted in constitutive phosphorylation, and targeted knockdown of *ALK* messenger RNA resulted in profound inhibition of growth in all cell lines harbouring mutant or amplified *ALK*, as well as in two out of six wild-type cell lines for *ALK*. Our results demonstrate that heritable mutations of *ALK* are the main cause of familial neuroblastoma, and that germline or acquired activation of this cell-surface kinase is a tractable therapeutic target for this lethal paediatric malignancy.

Neuroblastoma is a cancer of early childhood that arises from the developing autonomic nervous system. It is the most common malignancy diagnosed in the first year of life and shows a wide range of clinical phenotypes, with a few patients having tumours that regress spontaneously, whereas most patients have aggressive metastatic disease[1]. These latter neuroblastoma cases have survival probabilities of less then 40% despite intensive chemoradiotherapy, and the disease continues to account for 15% of childhood cancer mortality[1,2]. Tumours from patients with an aggressive phenotype often show amplification of the *MYCN* oncogene[3], and/or deletions of chromosome arms 1p and 11q (ref. 4). However, because *MYCN* is so aberrantly dysregulated, and no putative tumour suppressor gene at 1p and 11q has been shown to harbour inactivating mutations in more than a small percentage of cases, no tractable molecular target approaches at present exist for this disease.

Like most human cancers, a small subset of neuroblastoma cases is inherited in an autosomal dominant manner[5–7]. A family history of the disease is found in about 1–2% of newly diagnosed cases, with a standardized incidence ratio of 9.7 for siblings of index cases[8]. Neuroblastoma pedigrees show notable heterogeneity in the type of tumours that arise, with both benign and malignant forms occurring in the same family[9]. Familial neuroblastoma patients differ from those with sporadic disease in that they are diagnosed at an earlier age and/or with several primary tumours, clinical characteristics that are hallmarks of cancer predisposition syndromes. Because of the lethality of the condition before reproductive age, previous genetic linkage scans have been underpowered and results have been difficult to replicate[10–12]. Notably, neuroblastoma can occur with a spectrum of disorders related to abnormal development of neural-crest-derived tissues, including central congenital hypoventilation syndrome and Hirschsprung's disease. Missense or nonsense mutations in *PHOX2B*, a homeobox gene that is a master regulator of normal autonomic nervous system development, were recently shown to predispose to this rare field defect of the sympathoadrenal lineage tissues[13–15]. However, *PHOX2B* mutations explain only a small subset of hereditary neuroblastoma, are almost exclusive to cases with associated disorders of neural-crest-derived tissues, and are not somatically acquired in tumours[16,17], leaving the genetic aetiology for most familial neuroblastoma cases unknown.

Identification of germline *ALK* mutations

To identify the location of a hereditary neuroblastoma predisposition gene, we performed a genome-wide scan for linkage at ~6,000 single nucleotide polymorphisms (SNPs) in 20 neuroblastoma families. Because of the rarity of the condition, the genome-wide scan included pedigrees with varying degrees of confidence of actual heritability. Eight families had three or more affected individuals of close relation (high confidence), six families consisted of only two individuals of first-degree relation (moderate confidence), and six other families also only had two affected individuals, but of a more distant relationship (low confidence). We discovered a significant linkage signal at chromosome 2p, with a maximum non-parametric

[1]Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. [2]Translational Pediatric Oncology, National Institute for Cancer Research and Italian Neuroblastoma Foundation, National Institute for Cancer Research, Genoa 16132, Italy. [3]Advanced Biotechnology Center, Genoa 16132, Italy. [4]Center for Medical Genetics, Ghent University Hospital, Gent B-9000, Belgium. [5]The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [6]Scripps Genomic Medicine and the Scripps Research Institute, La Jolla, California 92037, USA. [7]Division of Genetics, The Children's Hospital of Philadelphia, Department of Pediatrics and CCEB, Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. [8]University of Rome "La Sapienza", Department of Experimental Medicine, Rome 00161, Italy. [9]Abramson Family Cancer Research Institute, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA.

likelihood ratio (lod) score of 4.23 at rs1344063 in 18 of the families (two were excluded owing to insufficient DNA). This refined a region previously reported for one of the pedigrees studied here[10]. By mapping informative recombination events, we defined a predisposition locus at chromosome bands 2p23–p24 delimited by SNPs rs1862110 and rs2008535 with 104 genes including the known neuroblastoma oncogene, *MYCN*[3,18], and the *ALK* oncogene located 13.2 megabases (Mb) centromeric. Despite previous work showing that forced overexpression of *MYCN* to the murine neural crest causes neuroblastoma[18], resequencing of the *MYCN* coding region and 18 kilobases (kb) of surrounding genomic DNA in probands from each linked family showed no disease-causal sequence variations.

We next focused on the *ALK* gene because our group and others had previously identified it as a potential oncogene in neuroblastoma through somatically acquired amplification of the genomic locus[19,20]. In addition, oncogenic fusion proteins leading to constitutive activation of the *ALK* kinase domain occur in many human cancers including anaplastic large cell lymphoma[21], inflammatory myofibroblastic tumours[22], squamous cell carcinomas[23] and non-small-cell lung cancers[24,25]. Resequencing of the 29 *ALK* coding exons identified three separate single base substitutions in the *ALK* tyrosine kinase domain in eight of the probands screened (Fig. 1 and Table 1). These DNA sequence alterations were not present in either single nucleotide polymorphism (dbSNP; http://www.ncbi.nlm.nih.gov/projects/SNP/) or somatic mutation (COSMIC; http://www.sanger.ac.uk/genetics/CGP/cosmic/) databases, and were not detected in direct sequencing of the *ALK* tyrosine kinase domain in 218 normal control alleles. Each substitution was subsequently shown to segregate with the disease in each family (Fig. 1). The sequence variation in family FNB12 (R1275Q) seems to have been acquired *de novo* in the affected father, and non-paternity was excluded by analysis of inheritance of genotypes in this pedigree. There are several asymptomatic obligate carriers identified (FNB2, FNB13, FNB32, FNB52 and FNB56), suggesting that the incomplete penetrance of this disease may be due to lack of the acquisition of a second genetic event such as amplification of the mutant allele, or alternatively, spontaneous regression following malignant transformation in at least a subset of cases. Notable is the very large multiplex family (FNB52), with discordance in twins and several unaffected carriers, which segregates a unique germline mutation (G1128A).

ALK sequence variations occurred only in the families with high or moderate degrees of confidence for harbouring a predisposing allele. Six of the eight families with three or more affected individuals had *ALK* missense alterations. The two families that did not have *ALK* sequence alterations identified were each shown to harbour mutations in the sympathoadrenal lineage-specific *PHOX2B* neurodevelopmental gene[14,16]. Two of the six families consisting of only two affected individuals, but of first-degree relation, had *ALK* sequence variations. Each of these families carried the R1275Q alteration, and in FNB12 we showed that the mutation arose *de novo* in the affected father, whereas in FNB56 the alteration was inherited from an unaffected father (Fig. 1). None of the six families with two distant relations affected with neuroblastoma showed *ALK* alterations, suggesting that the occurrence of a further case of this relatively rare disease in an extended family member was probably a chance occurrence. Because there are several families who share identical mutations, we looked to see if these families shared a common haplotype around the *ALK* gene and showed that the affected individuals with the same mutations did not share haplotypes, arguing against a founder effect.

Because *ALK* functions as an oncogene in other human cancers, we predicted that the sequence variations discovered in the neuroblastoma pedigrees would result in constitutive activation. We therefore used a support vector machine-based statistical classifier to map the putative mutations and determine the probability that they would act as drivers of an oncogenic process[26,27]. Each of the germline alterations occurred at regions of the *ALK* kinase domain that have been shown to be principal targets for cancer driver mutations in other

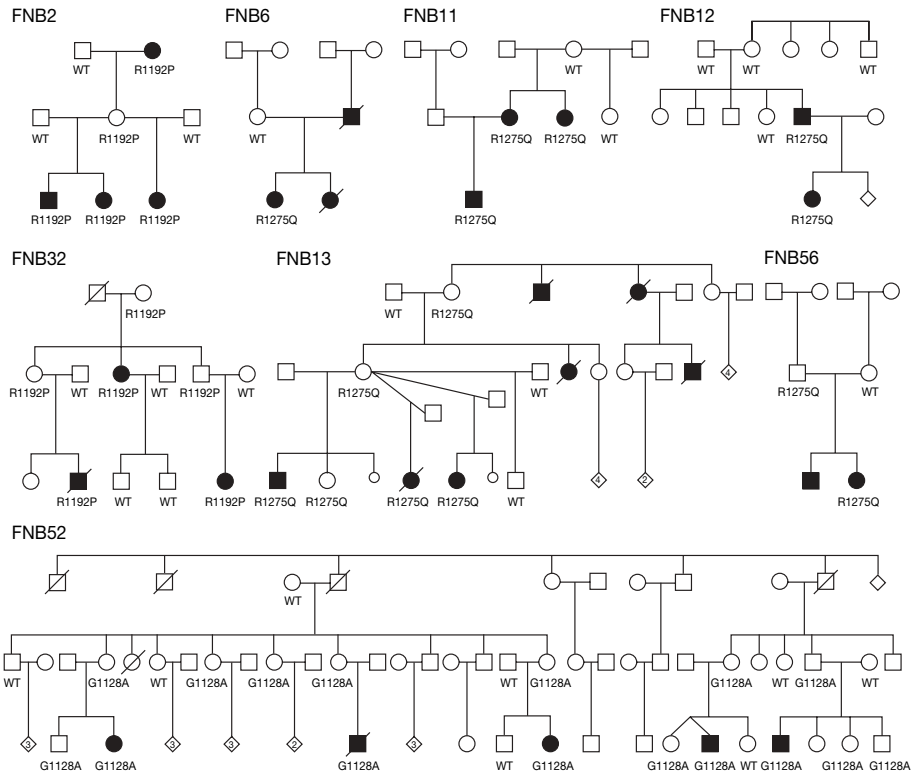

Figure 1 | Eight neuroblastoma pedigrees with *ALK* mutations. All family members with DNA available for genotyping are indicated by either wild-type (WT) for *ALK*, or by mutation in the *ALK* tyrosine kinase domain (R1192P, R1275Q, G1128A). Individuals affected by neuroblastoma are indicated by a filled symbol. The numbers inside the small diamonds indicate the number of other children, the lines through the symbols indicate that the person is deceased, and the smaller circles represent a miscarriage.

Table 1 | *ALK* mutations in neuroblastoma

Mutation	cDNA variation	Type/frequency	Region*	Probability activating mutation†
G1128A	c.3383G > C	Germ line (1/8)	P loop	0.95
R1192P	c.3575G > C	Germ line (2/8)	β4 strand	0.96
R1275Q	c.3824G > A	Germ line (5/8)	Activation loop	0.91
		Somatic (8/24)		
D1091N	c.3271G > A	Somatic (1/24)	N terminus	0.29
M1166R	c.3497T > G	Somatic (1/24)	C helix	0.79
I1171N	c.3512T > A	Somatic (2/24)	C helix	0.85
F1174I	c.3520T > A	Somatic (1/24)	End of C helix	0.92
F1174L	c.3522C > A	Somatic (8/24)	End of C helix	0.96
F1245C	c.3734T > G	Somatic (1/24)	Catalytic loop	0.94
F1245V	c.3733T > G	Somatic (1/24)	Catalytic loop	0.91
I1250T	c.3749T > C	Somatic (1/24)	Catalytic loop	0.87

* The region in which the codon alteration occurs is indicated. Note that the D1091N mutation is immediately adjacent to the tyrosine kinase domain.
† The probability that the amino acid alteration results in oncogenic activation on the basis of previously described methods[26].

oncogenic kinases (Table 1 and Fig. 2). The R1275Q mutation was present in the germline DNA of affected individuals from five pedigrees (Fig. 1), and falls within the kinase activation loop in a region strongly associated with activating mutations in many different protein kinases, such as *BRAF*[28]. This amino acid substitution results in an electropositive residue being replaced by a more electronegative one, possibly mimicking activating phosphorylation events. The R1192P mutation occurred at the beginning of the β4 strand of the kinase domain, and although it is predicted to be a driver mutation with high confidence (Table 1), the mechanism for activation is not yet clear[27]. The G1128A was seen only in the large pedigree with affected individuals in a single generation. The variation falls at the third glycine of the glycine loop, and identical mutations of this glycine residue to alanine in *BRAF* have been shown to increase kinase activity[29].

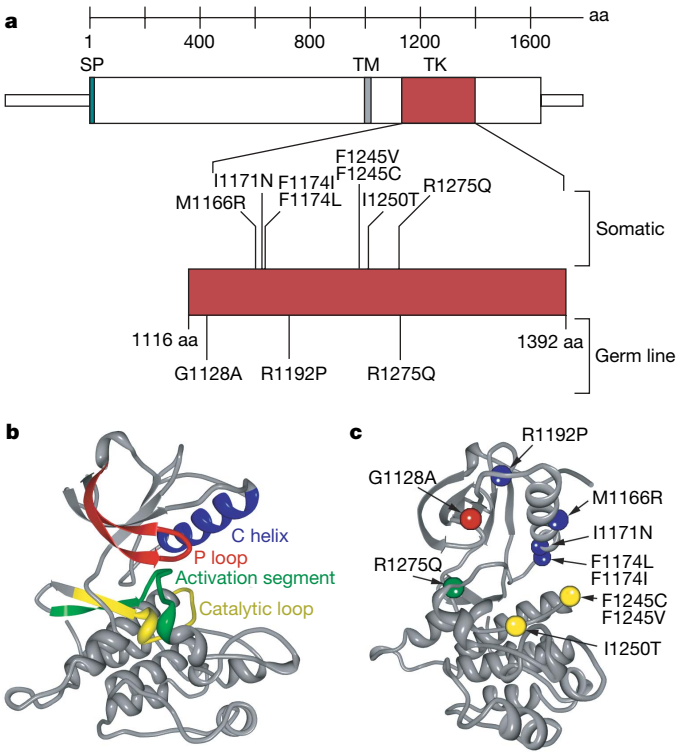

Figure 2 | Germline and somatic *ALK* mutations. a, Schematic diagram indicating protein structure of *ALK*, with mutations discovered in constitutional DNAs of familial cases (germ line) and primary tumours from sporadic cases (somatic) indicated. All but one sequence alteration mapped to the tyrosine kinase (TK) domain (D1091N was just amino-terminal and is not indicated here). Of the three germline mutations discovered, only the R1275Q was found in the tumour DNA samples. Conversely, the I1250T mutation discovered in the tumour set was also present in the matched germline DNA of that patient, whereas all of the other mutations studied here were somatically acquired. aa, amino acids; SP, signal peptide; TM, transmembrane. **b**, Homology model of wild-type *ALK* with each main subdomain indicated[26,27]. **c**, *ALK* mutations mapped onto homology model (different orientation from **b** to show all mutations), with colours indicating the subdomain (identified in **b**) in which the mutation resides (for example, the R1275Q mutation falls within the activation segment, indicated in green).

Identification of somatic *ALK* mutations

Having shown that heritable mutations in the *ALK* tyrosine kinase domain are associated with a highly penetrant predisposition to develop neuroblastoma, we next sought to determine whether *ALK* activation might also be somatically acquired. We examined a representative set of 491 sporadically occurring primary neuroblastoma samples acquired from children at the time of diagnosis on a 550K SNP-based microarray to assess for genome-wide copy number alterations. A total of 112 cases (22.8%) showed unbalanced gain of a large genomic region at 2p including the *ALK* locus (partial trisomy), and an additional 16 cases (3.3%) showed high-level focal amplification of *ALK* (Fig. 3). Each of the high-level amplifications co-occurred with *MYCN* amplification and/or other regions at 2p, except one case with an *ALK* amplicon only. The presence of aberrant *ALK* copy number status (gain or amplification) was highly associated with an aggressive clinical phenotype such as metastasis at diagnosis ($P < 0.0001$) and death from disease ($P = 0.0003$).

Because of the association of *ALK* gain and amplification with high-risk disease features, we next examined a subset of 167 tumour samples from high-risk patients, and 27 human neuroblastoma-derived cell lines (all from high-risk patients), for sequence alterations in the *ALK* tyrosine kinase domain. A total of 14 of the 167 tumour (8.4%) and 10 of the 27 cell line (35.7%) samples showed single base substitutions consistent with activating mutations (Fig. 2). Eight separate single base substitutions were identified, with the R1275Q mutation being the only mutation also seen in the germline DNA of the families studied. Again, none of the sequence variations discovered here was in SNP databases or was identified in our resequencing of the *ALK* tyrosine kinase domain in 109 control subjects (218 alleles). Mutations were equally distributed between cases with and without *MYCN* amplification. Only one case had a co-occurrence of an *ALK* mutation (F1174L) and genomic amplification of the *ALK* locus, and in this case the mutated allele was amplified (data not shown). Germline DNA was available for 9 out of 14 patients with *ALK* mutations, and in one of these cases the sequence alteration (I1250T) was also present in the germ line, suggesting a hereditary predisposition that may or may not be *de novo* in this case.

Using the same statistical classifier used for the germline mutations, we showed that all but one of the sequence variations discovered in the tumour tissues were predicted to be activating mutations (Table 1), and the one that showed a low probability (D1091N) was outside the core kinase domain. Most of the somatically acquired mutations fell into either the catalytic loop or the C-helix kinase domains, both frequent sites for oncogenic activating mutations

(Fig. 2). Catalytic loop mutants, especially I1250T, may promote oncogenesis by altering substrate binding or, most probably, by altering packing of the HRD and DFG motifs towards an activated conformation[30]. The mutations observed in the *ALK* C-helix domain occurred at positions within the kinase domain previously found to be mutated in other tumours. I1171N falls at an equivalent weakly oncogenic position in *MET* (M1149T)[31], and the M1166R, F1174I and F1174L mutants fall at equivalent positions mutated in *ERBB2* (D769, V777) and *EGFR* (D761, V769)[32–34].

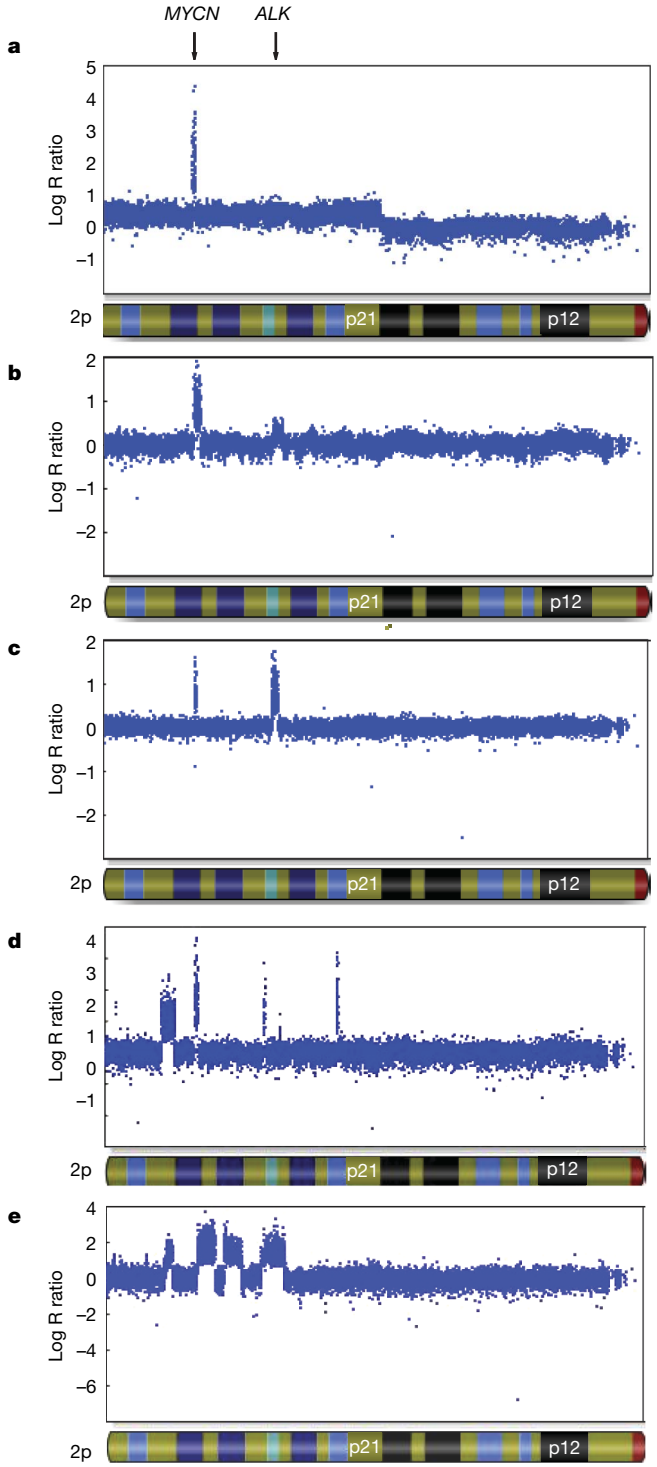

Figure 3 | Representative *ALK* copy number alterations in five neuroblastoma primary tumours. Hybridization intensity reflecting copy number for all SNPs along chromosome 2p in five primary tumours from patients with sporadically occurring disease is shown, and represented on a logarithmic scale. *MYCN* amplification is present in all tumours. **a**, Regional gain (trisomy) of chromosome 2p, including the *ALK* locus. **b**, Focal gain of the *ALK* locus. **c**, Focal amplification of the *ALK* locus. **d**, **e**, Complex rearrangement of the 2p locus, showing various focal amplicons, including *MYCN* and *ALK*.

Functional consequences of *ALK* mutations

We have previously shown that *ALK* is differentially expressed in human primary neuroblastomas with higher expression generally seen in the most aggressive subset of tumours[35]. Using quantitative polymerase chain reaction with reverse transcription (RT–PCR) we confirmed that *ALK* is highly expressed in all but one of the 20 human neuroblastoma-derived cell lines. *ALK* expression was higher in neuroblastoma cells compared to developing fetal brain, and the cell lines harbouring *ALK* mutations ($n = 6$) expressed the mRNA at a significantly higher copy number than the *ALK* wild-type cell lines ($n = 14$, Fig. 4a). Analysis of protein lysates from a panel of neuroblastoma cell lines showed constitutive phosphorylation of the tyrosine residue at codon 1604 in each of the cell lines harbouring mutations, with weak phosphostaining in two wild-type cell lines (Fig. 4b).

To determine whether *ALK* activation by mutation and/or amplification is functionally relevant in models of high-risk neuroblastoma—and thus might offer a tractable therapeutic target—we examined the consequences of disrupting *ALK* signalling by mRNA knockdown. We transiently transfected short interfering RNAs (siRNAs) directed against *ALK* (Dharmacon) into ten neuroblastoma cell lines and screened for inhibition of substrate adherent growth. We demonstrated knockdown of the mRNA and protein in all cell lines studied, but showed differential effects on cellular proliferation (Fig. 5a–l). Each of the cells harbouring *ALK* mutation or amplification showed profound inhibition of proliferation to *ALK* knockdown. In addition, two of the six *ALK* wild-type cell lines showed significant inhibition of growth with *ALK* knockdown and each of these had shown weak evidence for phosphorylation at

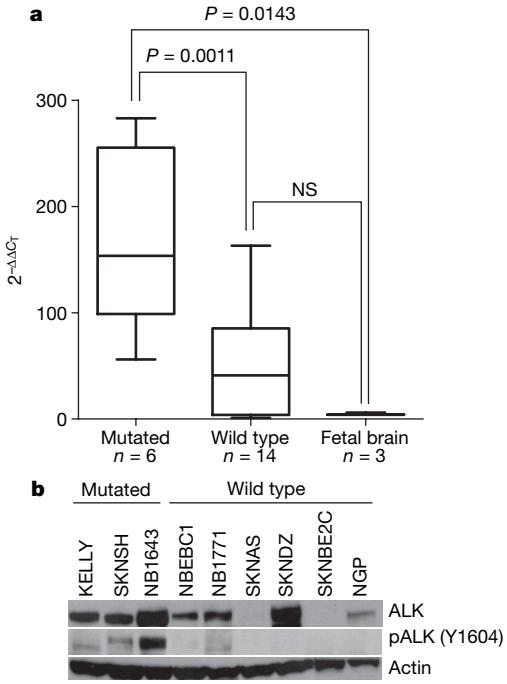

Figure 4 | *ALK* is highly expressed and the kinase is phosphorylated in neuroblastoma cell lines harbouring activating mutations. **a**, Relative *ALK* expression of neuroblastoma cell lines and fetal brain is shown and was determined using the $2^{-\Delta\Delta C_T}$ method[50]. Statistical significance was determined by an unpaired *t*-test. **b**, Immunoblots showing differential ALK expression in neuroblastoma cell lines with phosphorylation of the tyrosine 1604 codon restricted to cell lines with mutations (the wild-type lines NBEBC1 and NB1771 show faint phosphostaining). NS, not significant.

tyrosine 1604 (Fig. 4b), indicating that an alternative mechanism may have resulted in *ALK* kinase activation in these two cell lines.

Discussion

Knudson and Strong predicted that neuroblastoma, like the analogous embryonal cancer retinoblastoma, would follow a two-hit model explaining hereditary and sporadic cases[5]. This model has proved to be correct for most childhood and adult hereditary cancers, and the susceptibility genes are typically tumour suppressors in which the two hits are the sequential inactivation of both alleles. The discovery of heritable mutations in oncogenes as the aetiology of several endocrine neoplasia 1 cancers (*RET*), papillary renal carcinoma (*MET*) and gastrointestinal stromal tumours (*KIT*) challenged this model, but it is now clear that somatically acquired duplication or amplification of the mutant allele provides the second hit[36]. We have shown that heritable mutations in *ALK* are the cause of most hereditary neuroblastoma cases, providing, to our knowledge, the first example of a paediatric cancer arising because of mutations in an oncogene. Together with our recent report that common variations at chromosome band 6p22 predispose to the development of sporadic neuroblastoma[37], the genetic aetiology of this disease is now being defined. The discovery of highly penetrant, heritable *ALK* mutations as the cause of hereditary neuroblastoma is of immediate relevance to patients with a family history; screening with non-invasive techniques such as ultrasonography and measurement of urinary catecholamine metabolites should probably be implemented for unaffected children carrying an *ALK* mutation. In addition, continuing efforts to characterize the full spectrum of mutations in the entire *ALK* coding sequence, as well as determining the frequency of mutations across all neuroblastoma disease subsets, will be required to formulate genetic screening recommendations for newly diagnosed neuroblastoma patients with or without a family history.

ALK is an orphan tyrosine kinase transmembrane receptor with homology to neurotrophin receptors and the *MET* oncogene. Its expression is restricted to the developing nervous system with a postulated role in participating in the regulation of neuronal differentiation[38]. It is now clear that many human cancers activate *ALK* signalling by creating unique oncogenic fusions of *ALK* with a variety of partners through chromosomal translocation events[39]. Previous work had shown that a substantial percentage of human-derived neuroblastoma cell lines express *ALK* transcripts and ALK protein[40], but no definitive role for this oncogene had been proven[19,41–43]. *ALK* was recently identified as a molecular target in neuroblastoma by a screen of human cancer cell lines with pharmacological antagonists of the *ALK* kinase domain[44]. Our current report provides evidence for oncogenic activation of *ALK* by mutation of the kinase domain, and these data provide the genetic basis for the observation of sensitization to *ALK* kinase inhibition. In addition, our discoveries in neuroblastoma may lead to future resequencing efforts in other malignancies, especially those in which oncogenic fusion proteins have recently been discovered. The data presented here clearly establish *ALK* as a critical neuroblastoma oncogene and should increase efforts to identify the ligand for this receptor and to determine whether *ALK*-mediated signalling can be activated by mechanisms other than direct mutation and/or amplification of *ALK* alleles. Furthermore, receptor tyrosine kinases provide tractable targets for pharmacological inhibition, and this work should provide the impetus for developing therapeutic strategies aimed at inhibiting *ALK*-mediated signalling.

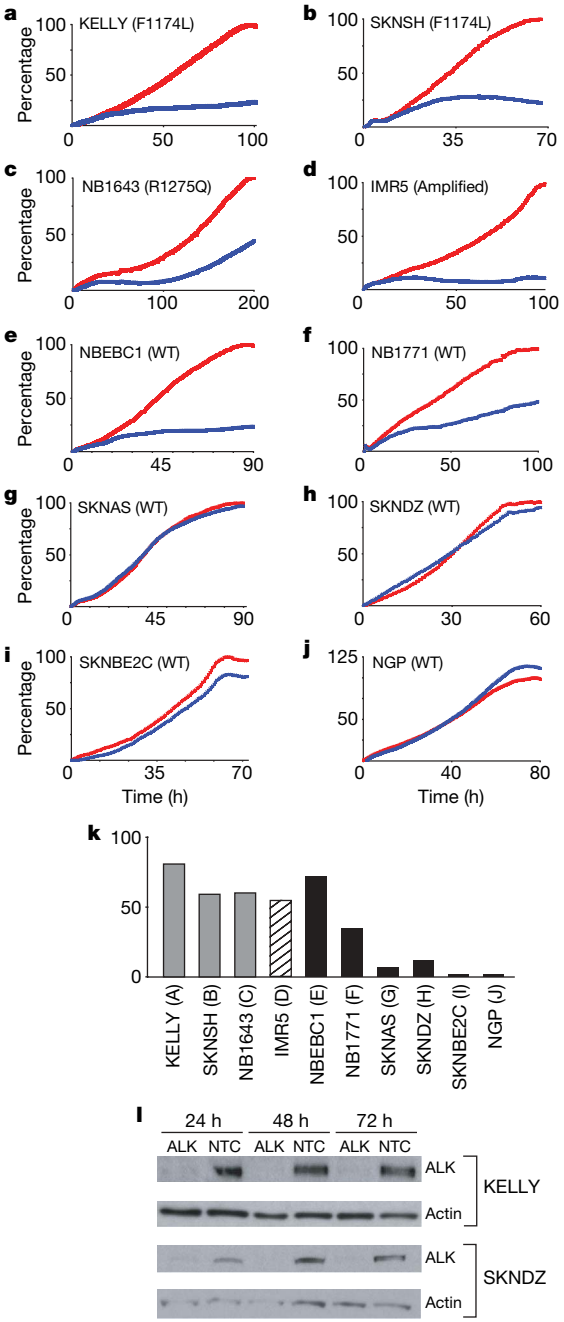

Figure 5 | *ALK* knockdown results in growth inhibition of *ALK* mutated or amplified neuroblastoma cell lines. a–j, Cellular growth for ten neuroblastoma cell lines that were transfected with siRNAs against *ALK* (blue) or *GAPDH* (red; two negative controls and one positive control not shown for clarity). The *x* axis is time in hours after transfection; the *y* axis is the percentage growth normalized to the siRNA against *GAPDH*. **k**, Summary of the percentage growth inhibition with *ALK* siRNA knockdown by *ALK* mutational and allelic status. **l**, Immunoblot showing a time course of ALK protein knockdown in the cell lines KELLY and SKNDZ. NTC, non-targeting control.

METHODS SUMMARY

Twenty probands with neuroblastoma and a family history of the disease were identified. Eight pedigrees had three or more affected individuals; six pedigrees contained only two affected individuals, but of first degree relation; and six pedigrees consisted of only two affected individuals, but of second, third or ≥fourth degree relationship. A total of 176 individuals (49 affected with neuroblastoma) were genotyped genome-wide, and two families were excluded owing to insufficient DNA for genotyping. We simulated marker data under a model of genetic homogeneity and autosomal dominant inheritance, and analysed the data using an affected-only approach comparable to the model-free approach used in the actual linkage analysis. Genotype data were checked for Mendelian inconsistencies using PedStats[45], and analysed for linkage using Merlin[46] and Lamp[47]. Regional candidates were re-sequenced using Sanger methodology. Predictions on the probability that DNA sequence alterations encode a mutant protein were performed using a support vector machine-based statistical classifier[26,27]. Four-hundred-and-ninety-one primary tumour samples and 27 cell lines were used for whole genome SNP-array analyses (550K) to determine copy number alterations[37]. mRNA knockdown of *ALK* and control targets was achieved with siRNAs against each target. The effects of siRNA knockdown on substrate adherent growth was quantified using the RT-CES

microelectronic cell sensor system (ACEA)[48,49]. Whole cell lysates were collected from the cell lines and from cells treated with either *ALK* siRNA or non-targeting control after transfection. Proteins were separated by SDS–PAGE gels and immunoblotted using anti-ALK and anti-phospho-ALK antibodies.

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

Received 21 April; accepted 14 July 2008.
Published online 24 August 2008.

1. Maris, J. M., Hogarty, M. D., Bagatell, R. & Cohn, S. L. Neuroblastoma. *Lancet* **369,** 2106–2120 (2007).
2. Matthay, K. K. *et al.* Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-cis-retinoic acid. Children's Cancer Group. *N. Engl. J. Med.* **341,** 1165–1173 (1999).
3. Schwab, M. *et al.* Chromosome localization in normal human cells and neuroblastomas of a gene related to c-*myc*. *Nature* **308,** 288–291 (1984).
4. Attiyeh, E. F. *et al.* Chromosome 1p and 11q deletions and outcome in neuroblastoma. *N. Engl. J. Med.* **353,** 2243–2253 (2005).
5. Knudson, A. G. J. & Strong, L. C. Mutation and cancer: Neuroblastoma and pheochromocytoma. *Am. J. Hum. Genet.* **24,** 514–522 (1972).
6. Kushner, B. H., Gilbert, F. & Helson, L. Familial neuroblastoma. Case reports, literature review, and etiologic considerations. *Cancer* **57,** 1887–1893 (1986).
7. Maris, J. M. *et al.* Molecular genetic analysis of familial neuroblastoma. *Eur. J. Cancer* **33,** 1923–1928 (1997).
8. Friedman, D. L. *et al.* Increased risk of cancer among siblings of long-term childhood cancer survivors: a report from the childhood cancer survivor study. *Cancer Epidemiol. Biomarkers Prev.* **14,** 1922–1927 (2005).
9. Maris, J. M. & Brodeur, G. M. in *Neuroblastoma* (eds Cheung, N.-K. V. & Cohn, S. L.) 21–26 (Springer, 2005).
10. Longo, L. *et al.* Genetic predisposition to familial neuroblastoma: identification of two novel genomic regions at 2p and 12p. *Hum. Hered.* **63,** 205–211 (2007).
11. Maris, J. M. *et al.* Evidence for a hereditary neuroblastoma predisposition locus at chromosome 16p12–13. *Cancer Res.* **62,** 6651–6658 (2002).
12. Perri, P. *et al.* Weak linkage at 4p16 to predisposition for human neuroblastoma. *Oncogene* **21,** 8356–8360 (2002).
13. Amiel, J. *et al.* Polyalanine expansion and frameshift mutations of the paired-like homeobox gene *PHOX2B* in congenital central hypoventilation syndrome. *Nature Genet.* **33,** 459–461 (2003).
14. Mosse, Y. P. *et al.* Germline *PHOX2B* mutation in hereditary neuroblastoma. *Am. J. Hum. Genet.* **75,** 727–730 (2004).
15. Trochet, D. *et al.* Germline mutations of the paired-like homeobox 2B (*PHOX2B*) gene in neuroblastoma. *Am. J. Hum. Genet.* **74,** 761–764 (2004).
16. Raabe, E. H. *et al.* Prevalence and functional consequence of *PHOX2B* mutations in neuroblastoma. *Oncogene* **27,** 469–476 (2008).
17. van Limpt, V. *et al.* The Phox2B homeobox gene is mutated in sporadic neuroblastomas. *Oncogene* **23,** 9280–9288 (2004).
18. Weiss, W. A., Aldape, K., Mohapatra, G., Feuerstein, B. G. & Bishop, J. M. Targeted expression of *MYCN* causes neuroblastoma in transgenic mice. *EMBO J.* **16,** 2985–2995 (1997).
19. Osajima-Hakomori, Y. *et al.* Biological role of anaplastic lymphoma kinase in neuroblastoma. *Am. J. Pathol.* **167,** 213–222 (2005).
20. George, R. E. *et al.* Genome-wide analysis of neuroblastomas using high-density single nucleotide polymorphism arrays. *PLoS ONE* **2,** e255 (2007).
21. Morris, S. W. *et al.* Fusion of a kinase gene, *ALK*, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* **263,** 1281–1284 (1994).
22. Griffin, C. A. *et al.* Recurrent involvement of 2p23 in inflammatory myofibroblastic tumors. *Cancer Res.* **59,** 2776–2780 (1999).
23. Jazii, F. R. *et al.* Identification of squamous cell carcinoma associated proteins by proteomics and loss of beta tropomyosin expression in esophageal cancer. *World J. Gastroenterol.* **12,** 7104–7112 (2006).
24. Soda, M. *et al.* Identification of the transforming *EML4–ALK* fusion gene in non-small-cell lung cancer. *Nature* **448,** 561–566 (2007).
25. Rikova, K. *et al.* Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* **131,** 1190–1203 (2007).
26. Torkamani, A. & Schork, N. J. Accurate prediction of deleterious protein kinase polymorphisms. *Bioinformatics* **23,** 2918–2925 (2007).
27. Torkamani, A. & Schork, N. J. Prediction of cancer driver mutations in protein kinases. *Cancer Res.* **68,** 1675–1682 (2008).
28. Ikenoue, T. *et al.* Functional analysis of mutations within the kinase activation segment of *B-Raf* in human colorectal tumors. *Cancer Res.* **63,** 8132–8137 (2003).
29. Ikenoue, T. *et al.* Different effects of point mutations within the *B-Raf* glycine-rich loop in colorectal tumors on mitogen-activated protein/extracellular signal-regulated kinase kinase/extracellular signal-regulated kinase and nuclear factor κB pathway and cellular transformation. *Cancer Res.* **64,** 3428–3435 (2004).
30. Kannan, N. & Neuwald, A. F. Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? *J. Mol. Biol.* **351,** 956–972 (2005).
31. Jeffers, M. *et al.* Activating mutations for the met tyrosine kinase receptor in human cancer. *Proc. Natl Acad. Sci. USA* **94,** 11445–11450 (1997).
32. Balak, M. N. *et al.* Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clin. Cancer Res.* **12,** 6494–6501 (2006).
33. Lee, J. W. *et al. ERBB2* kinase domain mutation in the lung squamous cell carcinoma. *Cancer Lett.* **237,** 89–94 (2006).
34. Lee, J. W. *et al.* Somatic mutations of *ERBB2* kinase domain in gastric, colorectal, and breast carcinomas. *Clin. Cancer Res.* **12,** 57–61 (2006).
35. Wang, Q. *et al.* Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. *Cancer Res.* **66,** 6050–6062 (2006).
36. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nature Med.* **10,** 789–799 (2004).
37. Maris, J. M. *et al.* Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. *N. Engl. J. Med.* **358,** 2585–2593 (2008).
38. Iwahara, T. *et al.* Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. *Oncogene* **14,** 439–449 (1997).
39. Chiarle, R., Voena, C., Ambrogio, C., Piva, R. & Inghirami, G. The anaplastic lymphoma kinase in the pathogenesis of cancer. *Nature Rev. Cancer* **8,** 11–23 (2008).
40. Lamant, L. *et al.* Expression of the *ALK* tyrosine kinase gene in neuroblastoma. *Am. J. Pathol.* **156,** 1711–1721 (2000).
41. Motegi, A., Fujimoto, J., Kotani, M., Sakuraba, H. & Yamamoto, T. ALK receptor tyrosine kinase promotes cell growth and neurite outgrowth. *J. Cell Sci.* **117,** 3319–3329 (2004).
42. Miyake, I. *et al.* Activation of anaplastic lymphoma kinase is responsible for hyperphosphorylation of ShcC in neuroblastoma cell lines. *Oncogene* **21,** 5823–5834 (2002).
43. Dirks, W. G. *et al.* Expression and functional analysis of the anaplastic lymphoma kinase (*ALK*) gene in tumor cell lines. *Int. J. Cancer* **100,** 49–56 (2002).
44. McDermott, U. *et al.* Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. *Cancer Res.* **68,** 3389–3395 (2008).
45. Wigginton, J. E. & Abecasis, G. R. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* **21,** 3445–3447 (2005).
46. Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genet.* **30,** 97–101 (2002).
47. Li, M., Boehnke, M. & Abecasis, G. R. Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. *Am. J. Hum. Genet.* **76,** 934–949 (2005).
48. Yu, N. *et al.* Real-time monitoring of morphological changes in living cells by electronic cell sensor arrays: an approach to study G protein-coupled receptors. *Anal. Chem.* **78,** 35–43 (2006).
49. Cole, K. A. *et al.* A functional screen identifies miR-34a as a candidate neuroblastoma tumor suppressor gene. *Mol. Cancer Res.* **6,** 735–742 (2008).
50. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}_{T}$ method. *Methods* **25,** 402–408 (2001).

Acknowledgements We acknowledge the families and children that participated in this research study, and the Children's Oncology Group for providing specimens. We thank W. London for providing statistical analyses related to the Children's Oncology Group tumour set, H. Rydbeck for his assistance with the linkage analysis, M. Wasik for technical assistance, and J. Felgenhauer, N. Van Roy and C. McConville for providing neuroblastoma pedigrees. This work was supported in part by National Institutes of Health grants K08-111733 (Y.P.M.), R01-CA78454 (J.M.M.), R01-CA87847 (J.M.M.), an American Society of Clinical Oncology Career Development Award (Y.P.M.), the Foerderer-Murray Fund (Y.P.M.), the Carly Hillman Fund (Y.P.M.), the Alex's Lemonade Stand Foundation (J.M.M.), the Andrew's Army Foundation (J.M.M.), the Giulio D'Angio Endowed Chair (J.M.M.), the Italian Neuroblastoma Foundation (L.L.), the Center for Applied Genomics at the Joseph Stokes Research Institute (H.H.), Scripps Genomic Medicine (A.T., N.J.S.), the Scripps Dickinson Scholarship (A.T.), and the Abramson Family Cancer Research Institute (J.M.M.).

Author Contributions Y.P.M. and J.M.M. designed the experiments and wrote the manuscript. Y.P.M., M.L., J.M.M., G.L., F.S., P.P. and G.P.T. collected the families for the linkage analysis. Y.P.M., M.L., L.L., C.K., C.H., E.R., H.H. and M.D. performed the genome-wide genotyping and linkage analysis. M.L. performed the DNA sequencing and analyses. K.A.C., A.W. and M.J.L. performed the siRNA experiments. E.F.A., H.H. and Y.P.M. performed the tumour SNP genotyping/copy number analyses. J.E.L., K.A.C. and A.W. performed the expression analyses. K.A.C., R.S. and M.L. performed the protein work. G.M.B. initiated the collection of neuroblastoma pedigrees. A.T. and N.J.S. performed the structural analysis of *ALK* coding mutations.

Author Information All sequence variations have been deposited to GenBank under accession numbers EU660517 to EU660527. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.M.M. (maris@chop.edu).

METHODS

Research subjects and samples. Families with a history of neuroblastoma in at least one other relative were eligible to participate. Only germline DNA from the neuroblastoma pedigrees was studied as no tumour tissue was available. Sporadic neuroblastoma tumour samples with matched constitutional DNA were acquired from the Children's Oncology Group neuroblastoma tumour bank. The Children's Hospital of Philadelphia Institutional Review Board approved this research.

Linkage analysis. A genome-wide linkage scan was performed using the Illumina Linkage IVb SNP panel. Genotype data were checked for Mendelian inconsistencies using PedStats[45], and analysed for linkage using Merlin[46] and Lamp[47]. The genome-wide screen for linkage was performed with both maximum-likelihood allele frequency estimates and model-free analyses. Because the pattern of inheritance is complex, we used a model-free approach so as not to assume any specific mode of inheritance. Model-based analyses were performed for all SNPs included in the critical region under a dominant mode of inheritance, assuming four different gene frequencies (0.0001, 0.001, 0.01 and 0.1) and dominant transmission with varying penetrance of the disease across a broad range, from 0.0001 to 0.68. Model-based linkage analysis was also performed using the method implemented in LAMP, assuming a prevalence of the disease of 0.000143 (1 in 7,000), and maximizing the lod scores over all possible disease models (MOD score analysis). In every analysis, the critical interval was defined as the region with associated lod scores that were greater than the maximum lod score minus three. Because linkage disequilibrium among markers is known to inflate the lod scores from linkage analysis in the presence of missing founders, its impact on the lod scores was assessed at the chromosome 2p critical interval. To model linkage disequilibrium, we organized markers into clusters by means of Merlin, which uses population haplotype frequencies derived from the HapMap project (http://www.hapmap.org/).

DNA sequencing. Shotgun resequencing from templates generated by long PCR for an 18 kb region surrounding the *MYCN* locus was performed using a 454 Life Sciences instrument (Branford) after bi-directional sequencing of the three coding exons showed no disease causal sequence variations in the pedigrees. Bi-directional sequencing of the *ALK* coding sequence (primer sequences available on request) was performed in the following distinct sample sets: (1) constitutional DNA from the proband and an unaffected first degree relative from the 20 neuroblastoma pedigrees, with repeat sequencing of amplicons containing any DNA sequence variations, and sequencing of amplicons containing confirmed variations in the remaining family members; (2) 27 human neuroblastoma-derived cell line DNAs maintained at the Children's Hospital of Philadelphia (available on request); (3) tumour DNA from 167 sporadic neuroblastomas from the Children's Oncology Group tumour bank; and (4) 109 normal constitutional DNAs from the SNP500 Cancer Resource panel purchased from the Coriell Institute for Medical Research. To verify neuroblastoma cell line integrity, all lines were routinely genotyped (AmpFLSTR Identifiler kit; Applied Biosystems) and tested for mycoplasma.

Mutation prediction. Cancer mutant predictions and analysis were performed as described previously[27]. In brief, a support vector machine was trained on common SNPs (presumed neutral) and congenital disease-causing SNPs characterized by a variety of sequence, structural and phylogenetic parameters. Training and predictions were performed using somatic mutations occurring in and outside the kinase catalytic core separately. The support vector machine-based method was then applied to the *ALK* mutants, and the probability that each mutant is a driver was computed using the support vector machine. The threshold taken for calling a SNP a driver was taken to be 0.49 for catalytic domain mutations and 0.53 for all other mutations[26]. For comparison to previously observed cancer mutations, *ALK* mutants were mapped to positions of a catalytic core alignment generated with characteristic site motifs, and previously observed cancer mutants mapping to the same positions were noted[27].

Tumour copy number analysis. Tumour samples were assayed on the Illumina Infinium II HumanHap550 BeadChip technology, at the Center for Applied Genomics at the Children's Hospital of Philadelphia. A total of 750 ng of genomic DNA was used as the input in each case; the assay was performed and data were analysed following the manufacturer's recommendations and as previously described[37].

Quantitative mRNA expression. Relative *ALK* expression was determined using the $2^{-\Delta\Delta C}{}_{T}$ method[50], using *GAPDH* as the endogenous control and using the second dCT (change in cycle threshold values) as the lowest expressed cell line, using methods as previously described[49].

***ALK* siRNA knockdown in neuroblastoma cell lines.** A total of $1–5 \times 10^4$ neuroblastoma cells were plated in triplicate overnight in antibiotic-free complete media in the 96-well RT-CES microelectronic cell sensor system (ACEA)[48,49]. Cells were transiently transfected with 200 µl containing 50 nM of pooled siRNAs (four separate siRNAs per transcript targeted) against *ALK* (M-003103-02), *GAPDH* (D-001140-01-20) negative control, non-targeting negative control, or *PLK1* (M-003290-01) positive control (siGENOME SMARTpool siRNA, Dharmacon) using 0.1% (v/v) Dharmafect I, according to the manufacturer's protocol (Dharmacon). The four separate *ALK*-directed siRNA sense-direction sequences are: GGGCCUGUAUACCGGAUAAUU (ALK J-003103-10), GUGCCAUGCUGCCAGUUAAUU (ALK J-003103-11), CCGCUUUGCCGAUAGAAUAUU (ALK J-003103-12) and GGAGCCACCUACGUAUUUAUU (ALK J-003103-13). In brief, 35 µl of 1 µM siRNA and 35 µl of serum-free media were combined with 0.7 µl Dharmafect I in 70 µl of serum-free media, incubated for 20 min at room temperature, and then 560 µl of antibiotic-free complete media was added. The culture media was gently removed from the plated cells and replaced by 200 µl of fresh media containing the siRNA, mock or complete media. Cell growth was monitored continuously and recorded as a cell index every 30 min for a minimum of 96 h. The cell index is derived from the change in electrical impedance as the living cells interact with the biocompatible microelectrode surface in the microplate well effectively measuring cell number, shape and adherence. Forty-eight hours after siRNA transfection, total RNA was extracted from the cells that had been plated in a parallel 96-well plate using the Qiagen mini extraction kit, with DNAase treatment. Two hundred nanograms of total RNA was oligodT primed and reverse transcribed using Superscript II reverse transcriptase (Invitrogen). *ALK*, *HPRT*, *GAPDH* and *PPIB* expression levels were measured by quantitative RT–PCR using Taqman gene expression assays (ABI), quantified on corresponding standard curves and normalized to the geometric mean of the three housekeeping genes. Two independent experiments were performed in triplicate. Growth inhibition of the neuroblastoma cell lines was determined by comparing the growth curve for *ALK* siRNA to the curve for *GAPDH* siRNA at the time of maximum cell index (CI_{max}): % growth inhibition $= (1 - CI_{ALK\ siRNA}/CI_{GAPDH\ siRNA}) \times 100$. *ALK* and *GAPDH* knockdown were determined by comparing the relative *ALK* expression: % knockdown $= (1 - ALK_{ALK\ siRNA}/ALK_{Control\ siRNA}) \times 100$. The average percentage knockdown of *ALK* across all cell lines was 60% (range 21%–86%). The average percentage knockdown of *GAPDH* was 75% (range 61%–95%).

ALK protein and phosphoprotein detection. Neuroblastoma cell lines were grown in T75 flasks under standard cell culturing conditions. For KELLY and SKNDZ lysates collected from the siRNA knockdown experiments, cells were plated in T25 flasks, transfected with 10 nM siRNA (as before) and collected at 24, 48 and 72 h after transfection. At 60%–80% confluency (or the appropriate time point), the cells were collected, pelleted and washed twice with ice cold PBS. Whole cell lysates were extracted with 100 µl cell extraction buffer (FNN011, Invitrogen) containing protease inhibitors (P-2714, Sigma) and phenylmethyl sulphonyl fluoride, briefly sonicated and rotated for 1 h at 4 °C. After a 30 min centrifugation at 4 °C, the supernatant was removed and protein quantification was performed using the Bradford method. Lysates (50 µg for siRNA experiment and 100 µg for native cell lines) were separated on 4%–12% Bis–Tris gradient gels and transferred to PVDF membranes. Membranes were then washed and incubated with anti-ALK (1:1,000; Cell Signaling, 3333), anti-phospho-ALK (1:1,000; Cell Signaling, 3341) and anti-actin (1:5,000; Santa Cruz, sc-2352) antibodies according to the cell signalling western protocol.