Mixed lineage kinase 3 gene mutations in mismatch repair deficient gastrointestinal tumours

Sérgia Velho[1], Carla Oliveira[1,2], Joana Paredes[1], Sónia Sousa[1], Marina Leite[1], Paulo Matos[3], Fernanda Milanezi[1], Ana Sofia Ribeiro[1], Nuno Mendes[1], Danilo Licastro[4], Auli Karhu[5], Maria José Oliveira[1,6], Marjolijn Ligtenberg[7], Richard Hamelin[8], Fátima Carneiro[1,2,9], Annika Lindblom[10], Paivi Peltomaki[11], Sérgio Castedo[1,2], Simó Schwartz Jr[12], Peter Jordan[3], Lauri A. Aaltonen[5], Robert M.W. Hofstra[13], Gianpaolo Suriano[1,2], Elia Stupka[14,15], Arsenio M. Fialho[16] and Raquel Seruca[1,2,*]

[1]IPATIMUP—Institute of Molecular Pathology and Immunology of the University of Porto, 4200-465 Porto, Portugal, [2]Medical Faculty of the University of Porto, 4200-319 Porto, Portugal, [3]Centre of Human Genetics, National Health Institute Dr Ricardo Jorge, 1649-016 Lisbon, Portugal, [4]CBM S.c.r.l., Area Science Park, Basovizza - SS 14, Km. 163,5, 34012 Trieste, Italy, [5]Department of Medical Genetics, Haartman Institute, University of Helsinki, 00014 Helsinki, Finland, [6]NewTherapies Group, INEB—Institute for Biomedical Engineering, Porto, Portugal, [7]Department of Human Genetics, UMC Nijmegen, 6500 HB Nijmegen, The Netherlands, [8]INSERM U434 CEPH, 75010 Paris, France, [9]Hospital de S. João, 4200-319 Porto, Portugal, [10]Department of Molecular Medicine and Surgery, Karolinska Institutet, S 171 76 Stockholm, Sweden, [11]Department of Medical Genetics, Biomedicum Helsinki, University of Helsinki, 00014 Helsinki, Finland, [12]Centre d'Investigacions en Bioquimica i Biologia Molecular (CIBBIM), Hospital Universitari Vall d'Hebron, Barcelona 08035, Spain, [13]Department of Medical Genetics, University Medical Center Groningen, 9700RB Groningen, The Netherlands, [14]UCL Cancer Institute, Paul O'Gorman Building, University College London, Gower Street, London WC1E 6BT, UK, [15]The Blizard Institute, Barts and The London School of Medicine and Dentistry, 4 Newark Street, London E1 2AT, UK and [16]Institute for Biotechnology and BioEngineering (IBB), Center for Biological and Chemical Engineering, Instituto Superior Tecnico, 1049-001 Lisbon, Portugal

Mixed lineage kinase 3 (MLK3) is a serine/threonine kinase, regulating MAPkinase signalling, in which cancer-associated mutations have never been reported. In this study, 174 primary gastrointestinal cancers (48 hereditary and 126 sporadic forms) and 7 colorectal cancer cell lines were screened for MLK3 mutations. MLK3 mutations were significantly associated with MSI phenotype in primary tumours ($P = 0.0005$), occurring in 21% of the MSI carcinomas. Most MLK3 somatic mutations identified were of the missense type (62.5%) and more than 80% of them affected evolutionarily conserved residues. A predictive 3D model points to the functional relevance of MLK3 missense mutations, which cluster in the kinase domain. Further, the model shows that most of the altered residues in the kinase domain probably affect MLK3 scaffold properties, instead of its kinase activity. MLK3 missense mutations showed transforming capacity in vitro and cells expressing the mutant gene were able to develop locally invasive tumours, when subcutaneously injected in nude mice. Interestingly, in primary tumours, MLK3 mutations occurred in KRAS and/or BRAF wild-type carcinomas, although not being mutually exclusive genetic events. In conclusion, we have demonstrated for the first time the presence of MLK3 mutations in cancer and its association to mismatch repair deficiency. Further, we demonstrated that MLK3 missense mutations found in MSI gastrointestinal carcinomas are functionally relevant.

*To whom correspondence should be addressed. Tel: +35 1225570700; Fax: +35 1225570799; Email: rseruca@ipatimup.pt

INTRODUCTION

In gastrointestinal carcinomas, ~15% of sporadic and 90% of hereditary non-polyposis colon carcinoma (HNPCC or Lynch tumours) show microsatellite instability (MSI) in multiple repeat sequences, due to defects in mismatch repair genes (MMR). Despite showing distinct molecular mechanisms for MMR silencing (1,2), MSI sporadic gastric (GC) and colorectal carcinomas (CRC), and hereditary MSI carcinomas encompass very similar molecular mutation profiles in non-coding and coding repetitive tracts. Those genes that accumulate mutations in their coding sequences are called target genes (2). Moreover, MMR deficiency is also known to boost the frequency of point mutations, namely in proto-oncogenes. Sporadic MSI CRC harbour activating missense mutations in genes coding for protein members of the RAS-RAF-MAP kinase pathway, a crucial pathway in tumorigenesis. The BRAFV600E hotspot mutation has been described in ~40% of sporadic MSI CRC and point mutations in KRAS are present in ~20% of the cases (1). Mutations in both genes have been rarely described in MSI sporadic CRC (1). In MSI GC, only KRAS mutations have been observed in 20% of the cases (3). Similarly, 40% of HNPCC cases display KRAS mutations and BRAF mutations were never found (4,5). Still, many HNPCC and sporadic gastrointestinal carcinomas lack KRAS or BRAF activating mutations, and alterations in other genes involved in the KRAS/BRAF pathway may represent further genetic events, thereby leading to the deregulation of this pathway. Recently, MLK3 has been appointed as a pivotal protein involved in the regulation of the mitogen-protein (MAP) kinase pathway.

MLK3 contains an N-terminal Src-homology 3 (SH3) domain, a kinase domain, a leucine zipper, a Cdc42/Rac inter-active binding (CRIB) motif and a COOH-terminal proline–serine–threonine rich domain. The SH3 domain of MLK3 can bind a proline residue in a region between the leucine zipper and the CRIB motif resulting in auto-inhibition (6). MLK3 is a serine/threonine protein kinase that regulates the MAPKinase pathway activating ERK, p38 and JNK, in response to extracellular signals (7,8). Further, MLK3 has been demonstrated to function as a scaffolding protein, involved in the formation of a multiprotein complex containing MLK3/BRAF/RAF1 (7,9,10). The formation of this complex was shown to be important for the activation of wild-type BRAF and, consequently, to the activation of ERK signalling (7,9,10). Furthermore, MLK3 was reported to be important for the proliferation of tumour cells, bearing either oncogenic KRAS or neurofibromatosis-1 (NF1) or NF2 inactivating mutations (10).

Frequently, the overexpression of wild-type proto-oncogenes leads to cell transformation (11). Likewise, over-expression of the wild-type MLK3 was shown to induce transformation of NIH3T3 fibroblasts (12). Overall, these data suggest that MLK3 is likely to be involved in cancer and alterations of this gene could harbour transforming ability; nevertheless, no MLK3 gene alterations have ever been described so far in cancer.

In this study, we aimed at determining whether MLK3 gene is a target of point mutations in gastrointestinal carcinoma and, if so, what is the distribution and type of such mutations among different settings of gastrointestinal cancer, and whether MLK3 missense mutations encompass transforming and tumorigenic potential *in vitro* and *in vivo*. Moreover, we assessed the co-existence of mutations in MLK3 and in KRAS and BRAF genes, which are frequently mutated in this type of neoplasia.

RESULTS

MLK3 heterozygous somatic mutations cluster in MSI gastrointestinal cancers

All MLK3 exons and intron–exon boundaries were screened for mutations in a series of 174 primary gastrointestinal cancers [114 MSI cases: 48 hereditary carcinomas (38 Lynch tumours with characterized germline mismatch repair gene mutations and 10 HNPCC patients that fulfilled the clinical criteria), 36 sporadic CRC and 30 sporadic GC; and 60 MSS sporadic CRC cases] and 7 CRC cell lines: 4 MSI and 3 MSS (Table 1). Whenever a sequence variant was found, matched constitutional DNA was analysed to exclude germline origin and 100 normal chromosomes were analysed to exclude polymorphic origin.

Overall, we found that 25/174 (14.4%) primary gastrointestinal carcinomas displayed heterozygous somatic and tumour-specific MLK3 mutations. Twenty-four of the 25 gastrointestinal mutant cases were MSI and only a single case was MSS. The frequency calculated only among MSI cases was 21.1% (24/114) (Table 1). These results demonstrated that MLK3 mutations occur preferentially in gastrointestinal cancers displaying an MSI phenotype ($P = 0.0005$). In terms of different MSI gastrointestinal cancer settings, we found MLK3 somatic mutations in 25% (12/48) of hereditary (Lynch/HNPCC) carcinomas; in 19.4% (7/36) of sporadic CRC and in 16.7% (5/30) of sporadic GC cases. The frequency of MLK3 mutations was not significantly different among the three settings of MSI carcinomas ($P = 0.6527$) (Table 1). Within the MLK3 mutant cases, two MSI carcinomas (case 1 and 17) harboured two different mutations (Table 2).

MLK3 mutations were also found in MSI CRC cell lines: 2/4 (RKO and CO115), mimicking the results found in primary tumours. The three MSS colorectal cancer cell lines (HT29, Caco2 and SW480) analysed did not show MLK3 mutations (Table 1).

From the 24 MLK3 mutated MSI cases, 15 (62.5%) displayed somatic missense mutations (Table 1), which are depicted in Figure 1. Seventy percent of these missense mutations occurred in previously described functional domains (Fig. 1) (6,13); and more than 80% affected MLK3 residues which have been evolutionarily conserved as far in evolution as *Drosophila melanogaster* (Fig. 1).

Frameshift mutations occurred in 33.3% (8/24) of MLK3 mutated MSI cases, preferentially at repetitive sequences (six out of eight) of about four to six mononucleotides. This type of mutations was preferentially clustered in the last MLK3 domain (P/S/T-rich). The frameshift mutation 2466delG was the only recurrent mutation and was found in two different cases (Table 2). A single nonsense mutation was also found among MLK3 mutated MSI cases (Table 2).

Table 1. Frequency of cases with MLK3 mutations found in MSI and MSS gastrointestinal tumours and cell lines

	Mixed lineage kinase 3 (MLK3) Mutational status		Type of mutation	
	Wild-type, n (%)	Mutant, n (%)	Missense, n (% total mut)	Frameshift/nonsense, n (% total mut)
Lynch/HNPCC ($n = 48$)	36 (75.0)	12 (25.0)	7 (58.3)	5 (41.7)
MSI sporadic CRC ($n = 36$)	29 (80.6)	7 (19.4)	5 (71.4)	2 (28.6)
MSI sporadic GC ($n = 30$)	25 (83.3)	5 (16.7)	3 (60.0)	2 (40)
Total number of MSI gastrointestinal tumours ($n = 114$)	90 (78.9)	24 (21.1)	15 (62.5)	9 (8/1) (37.5)
MSS sporadic CRC ($n = 60$)	59 (98.3)	1 (1.7)	1 (100%)	0
Total number of MSI and MSS gastrointestinal tumours ($n = 174$)	149 (85.6)	25 (14.4)	16 (64.0)	9 (36.0)
MSI CRC cell lines ($n = 4$)	2 (50)	2 (50)	1 (50)	1 (50)
MSS CRC cell lines ($n = 3$)	3 (100)	0	0	0

The frequency of MLK3 mutations found in primary MSI gastrointestinal tumours (24/114) is significantly different from the frequency of MLK3 mutations in primary MSS tumours (1/60) ($P = 0.0005$). P-value was calculated using Fisher's exact test and $P < 0.05$ was taken as statistically significant.
MSS, microsatellite stable; MSI, microsatellite unstable; CRC, colorectal cancer; GC, gastric cancer.

Table 2. Type of MLK3 mutations and association with other genetic alterations in hereditary and sporadic MSI gastrointestinal tumours and colorectal cancer cell lines

Tumour type	Mut Case	MLK3 alterations Missense mutations Nucleotide change	Amino acid change	Frameshift and nonsense mutations	Exon	Other alterations KRAS	BRAF	Altered MMR gene
Lynch/HNPCC	1	c.296 A>G	p.Y99C	—	1	G13D	—	nd
		c.886 G>A	p.A296T		2			
	2	c.493 G>T	p.A165S	—	1	—	—	MLH1
	3	c.718 C>T	p.R240C	—	1	—	—	MSH2
	4	c.1399 G>A	p.E467K	—	5	G13D	—	MSH6
	5	c.1976 G>A	p.R659H	—	9	G12D	—	MLH1
	6	c.2395 C>T	p.R799C	—	10	—	—	nd
	7	c.2514 G>T	p.Q838H	—	10	G13D	—	MSH2
	8	—	—	c.139delA	1	—	—	MSH2
	9	—	—	c.2032delC	9	G13D	—	MLH1
	10	—	—	c.2114delC	9	—	—	MLH1
	11	—	—	c.2466delG	10	G12D	—	MLH1
	12	—	—	c.2466delG	10	—	—	nd
Sporadic CRC	13	c.1054 G>A	p.A352R	—	3	G12D	—	
	14	c.1067 C>T	p.A356V	—	3	—	V600E	
	15	c.1418 G>A	p.R473H	—	5	—	V600E	
	16	c.2051 C>T	p.A684V	—	9	G13D	—	
	17	c.2011 C>T	p.R671C	—	9	—	—	
		c.2117 C>T	p.P706L					
	18	—	—	c.1241delA	4	—	—	
	19	—	—	c.1928_1929insG	9	—	V600E	
Sporadic GC	20		p.E53D	—	1	—	—	
	21	c.1534 C>T	p.R512W	—	6	—	—	
	22	c.2519 C>T	p.P840L	—	10	—	—	
	23	—	—	c.498_501delCTCT	1	—	—	
	24	c.1591 C>T	—	p.R531X	6	G12D	—	
CRC lines	25	c.755 C>A	p.P252H	—	2	—	V600E	
	26	—	—	c.1816delC	8	—	V600E	

c, coding DNA; p, protein; X, stop codon; >, substitution; del, deletion; ins, insertion; nd, not determined—patients fulfilled the clinical criteria.

Frameshift mutations affecting coding repetitive sequences are among the most common mutations occurring in MSI cancers, as a consequence of DNA polymerase slippage, frequently generating truncated proteins (2). The functional relevance of these frameshift mutations depends both on the target gene affected and on the tissue where they occur (14). In order to get insight on the functional role of MLK3 frameshift mutations, we compared their frequency in coding repetitive sequences with the frequency of frameshift mutations occurring in 11 non-coding intergenic repeats (15) of the same type and size (C/G6) in 30 MSI CRCs. The frequency of frameshift mutations in MLK3 repeats was higher (25%—6/24) when compared with the frequency found in intergenic repeats (0–3%, Supplementary Material, Table S3). These results suggest that frameshift mutations occurring in MLK3 coding repeat sequences are selected for
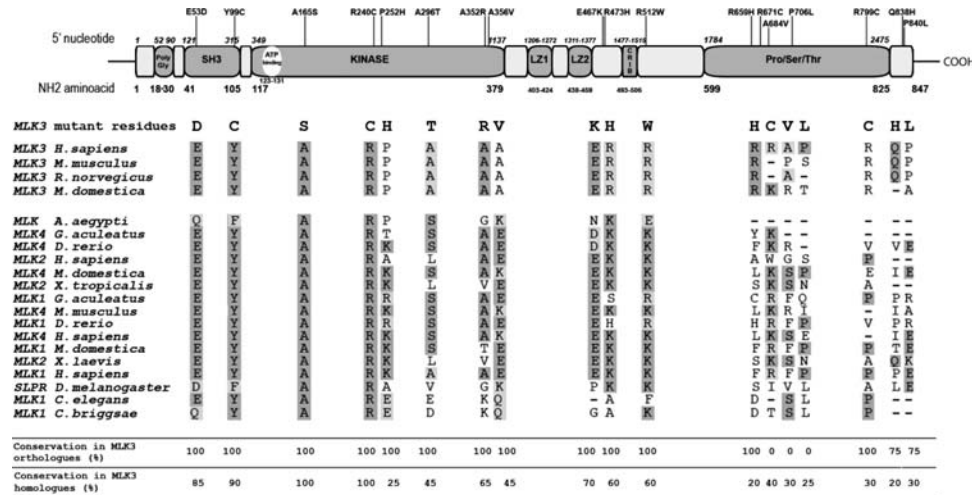
Figure 1. Summary of the localization of missense mutations in MLK3 gene found in MSI primary gastrointestinal carcinomas and cell lines. The upper diagram depicts the protein sequence domain architecture, indicating the affected residue and its amino acid change (top), nucleotide start-end coordinates (top) and amino acid start-end coordinates (bottom) for each domain. Below, the wild-type amino acid residue is shown for 20 eukaryotic species, and the conservation for each affected residue is summarized in the bottom two rows, one indicating the conservation in mammalian orthologues ($n = 4$) and the next in all eukaryotic homologues ($n = 16$).

during the process of tumour development of MSI gastrointestinal cancers. In contrast, missense mutations occur less frequently in MSI cancers, but tend to be oncogenic, as is the case of mutations affecting KRAS, BRAF and PI3KCA in MSI CRC (16). Moreover, missense mutations in MLK3 are far more frequent than frameshift mutations in our series of MSI gastrointestinal cancers. For these reasons, in the present study, we decided to address the potential impact of MLK3 missense mutations only.

Missense mutations in the kinase domain of MLK3 are predicted to impact protein function

Before addressing, *in vitro* and *in vivo*, the potential impact of MLK3 missense mutations in CRC cells, we decided to run a predictive analysis by modelling the 3D structure of MLK3 kinase domain, using the crystal structure of the same domain of MLK1 as template (17). There were three reasons underlying this choice: (i) a crystal structure of MLK3 is not available; (ii) MLK1 shows great amino acid homology with MLK3 (Pdb: 3dtc; 70% amino acid identity with MLK3), allowing an accurate modelling and (iii) even for MLK1, only the crystal structure of the kinase domain is available.

Within the MLK3 kinase domain, we mapped six single-amino acid changes (A165S, R240C, P252H, A296T, A352R and A356V). These mutations were spread along the kinase domain and did not occur at specific and highly conserved functional amino acid positions of the kinase catalytic core. Based on our predictive model, all these missense mutations affect surface-exposed residues, causing overall destabilization of the molecule (Fig. 2A and B). Mutants A165S, P252H and A352R are likely to induce the formation of new intramolecular hydrogen bonds with neighbouring residues, producing conformational changes in the protein backbone (Fig. 2a1, a3, a5). R240C mutation occurs at a coil region that runs parallel to an alpha helix in the C-terminal lobe of the MLK3 kinase domain (Fig. 2a2). The replacement of a larger residue (R) by a smaller

one (C) creates higher dissociation energy, resulting in a decrease of the steric hindrance. The A at position 296 is located at an exposed loop of the MLK3 kinase domain. The substitution of a hydrophobic A residue by and hydrophilic T residue (A296T) is predicted to destabilize this area of the protein, impacting most probably the scaffold properties of MLK3 (Fig. 2a4). The A residue at position 356 is part of an Alanine-rich helix of MLK3. The substitution of A residue by a V (A356V) does not alter the polarity at this site, as both are non-polar aliphatic; nevertheless, a Valine occupies more space than an Alanine and, therefore, the helix is likely to assume a different shape/deformation and/or suffer a destabilizing effect (Fig. 2a6).

Overall, our predictive MLK3 kinase domain 3D model shows that mutation at the six abovementioned residues, in the kinase domain of MLK3, are more likely to disturb the scaffold properties of the protein rather than its kinase activity. These *in silico* observations generated enough evidence to pursue the potential *in vitro* and *in vivo* analysis of MLK3 kinase domain missense mutations.

Cells expressing missense mutations of MLK3 have transforming potential *in vitro*

In order to address, *in vitro*, the transforming potential of MLK3 missense mutations, we selected five different mutations, Y99C, A165S, P252H, R799C, P840L occurring in distinct domains of the protein. A165S and P252H missense mutations were chosen among the six mutations occurring at the MLK3 kinase domain and the other three mutations occurred in three different domains of the protein (Table 2, Fig. 1).

After the generation of vectors expressing the five selected MLK3 mutations, as well as K144R, a kinase dead form of MLK3 to be used as a negative control (10), all these constructs were transiently transfected in NIH3T3 cells, as well as the wild-type form of MLK3 and the empty vector (Mock). A classical fibroblast focus formation assay was used to assess transforming ability of manipulated cell lines.
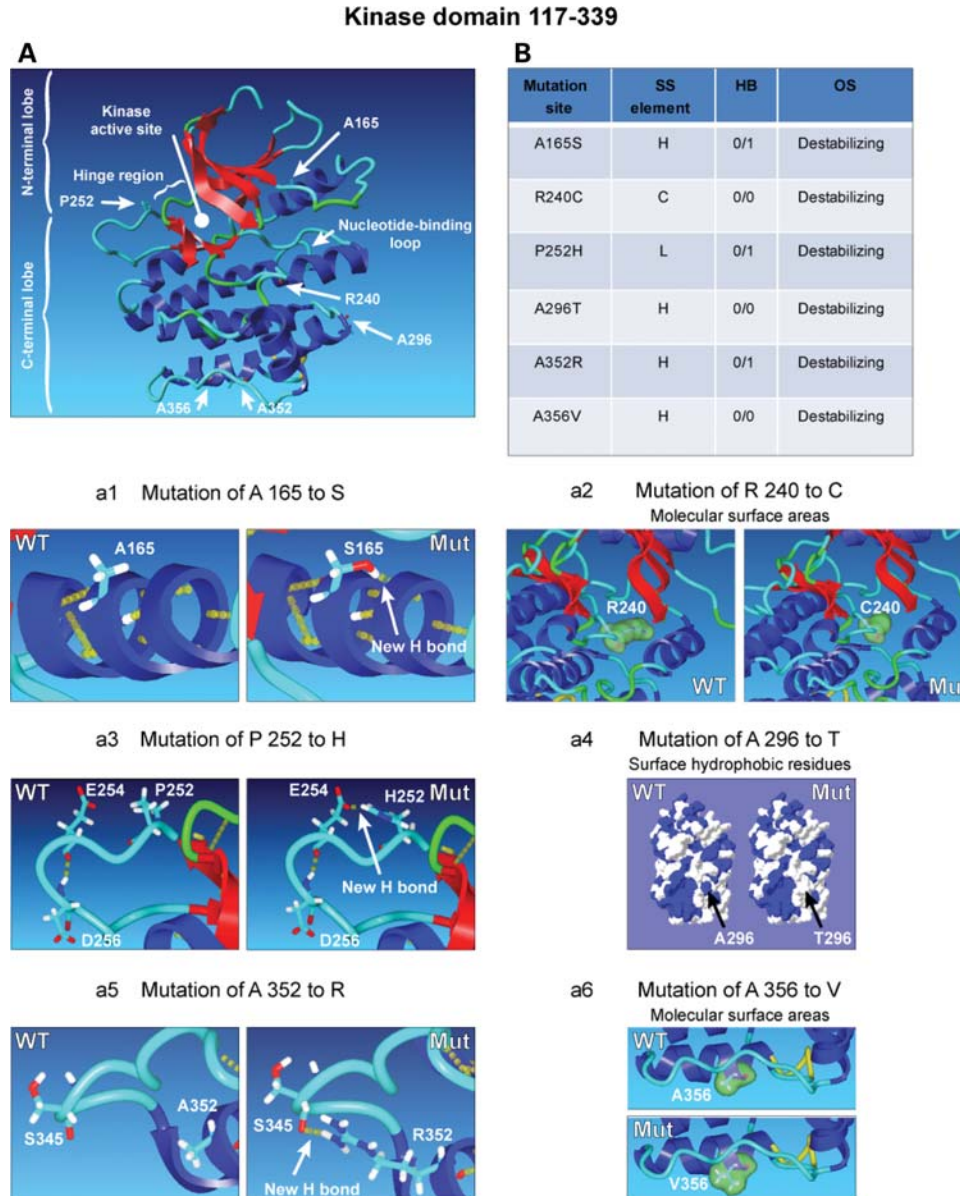
Kinase domain 117-339

Figure 2. 3D model to predict how MLK3 mutations could interfere with the MLK3 protein function. (**A**) Predictive 3D model of the MLK3 kinase domain, constructed using the crystal structure from MLK1 (Pdb: 3dtc; 70% amino acid identity with MLK3) as a template (17). Kinase active site, nucleotide binding loop, hinge region, the N- and C-termini lobes and the six missense mutations identified in this study are highlighted. (a1–a6) Magnified views of the different MLK3 mutations. Prediction of the impact of the six amino acid substitutions is described in the Discussion section. Dotted yellow lines represent potential H-bonds. (**B**) Prediction of mutations effects on the MLK3 kinase domain stability, based on the site directed mutator (SDM) program (37). SS element, secondary structure element (H, helix, C, coil, L, loop); HB, hydrogen bond; OS, overall stability.

All missense mutants analysed showed higher transforming capacity *in vitro* than the wild-type MLK3 and the kinase dead MLK3 ($P < 0.05$; Fig. 3).

Cells expressing missense mutations of MLK3 have tumorigenic potential *in vivo*

To further evaluate the tumorigenic potential of MLK3 missense mutations *in vivo*, NIH3T3 cells expressing P252H and R799C mutants, which were the most transforming mutations *in vitro* and found in two distinct key functional domains (the kinase and the P/S/T-rich domains, respectively)

were inoculated subcutaneously in nude mice. Cells expressing wild-type MLK3, the empty vector (Mock) and HRASV12 were used as controls. Six weeks after inoculation, neither mice inoculated with MLK3 wild-type expressing cells, nor those inoculated with the empty vector, generated tumours. As expected, HRASV12-expressing cells generated subcutaneous tumours within 4 weeks after inoculation. Nude mice inoculated with both MLK3 mutants developed subcutaneous tumours within 5 weeks after inoculation. One of these mice even presented a fast growing tumour 3 weeks after inoculation (Fig. 4A). The *in vivo* experiment was completed 6 weeks after inoculation and tumours were surgically
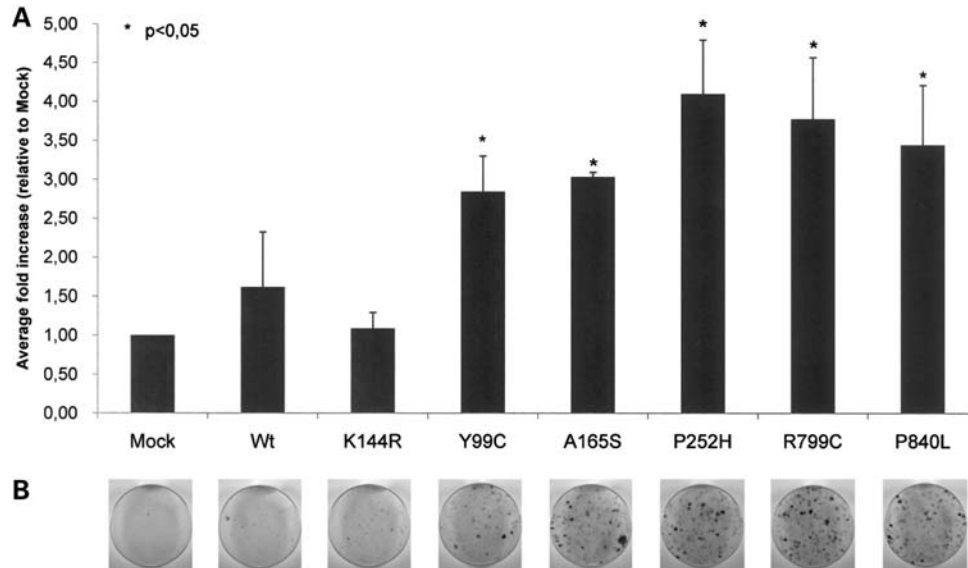
Figure 3. *In vitro* transforming capacity of mutant MLK3. NIH3T3 cells were transfected with the indicated constructs and then submitted to a Focus Formation Assay, in order to ascertain the transforming potential of mutant MLK3. Experiments were performed in triplicate. (**A**) Values in the graph represent the average fold increase of the number of foci observed when compared with the number of foci formed by NIH3T3 Mock cells. (**B**) Photographs of representative dishes; mutants were compared with wild-type for statistical analyses, using Student's *t*-test, and $P < 0.05$ was taken as statistically significant. Wt, wild-type MLK3.

removed and histopathologically analysed. Mutant MLK3 expressing cells grew into malignant tumours, with a high number of mitotic figures and an infiltrative pattern of growth. The number of mitotic figures, per observation field, in MLK3 mutant tumours was similar to those observed in HRASV12 tumours. Further, and in contrast to the positive control HRASV12, both MLK3 mutant tumours showed the ability to locally infiltrate the surrounding adipose tissue and muscle (Fig. 4B). Xenografted tumours expressing P252H showed the most aggressive behaviour, with invasion of backbone, bone marrow and spinal cord (Fig. 4C).

Our *in vivo* data clearly demonstrate the tumorigenic potential of at least two mutant forms of MLK3, suggesting a role for this gene in cell transformation and acquisition of a malignant behaviour, namely local invasion. To further support a role for MLK3 in cell invasion, we quantified *in vitro* the invasion potential of NIH3T3 cells expressing P252H and R799C, and the same cell line transfected with the empty vector, using the matrigel invasion assay. Cells expressing MLK3 mutants P252H and R799C showed a 3-fold increase ability to invade *in vitro* when compared with wild-type MLK3 cells (data not shown).

MLK3 mutations also occur in MSI gastrointestinal tumours with wild-type KRAS and BRAF

KRAS and BRAF, two key molecules of the MAPKinase pathway, are frequently mutated in sporadic MSI gastrointestinal cancer, and KRAS mutations occur frequently in HNPCCs (3,5). Nevertheless, oncogenic mutations at these genes only account for 30–50% of the cases (1,3–5,18,19). Overall, it is described that ∼50% of all MSI sporadic CRC, 40% of HNPCC and 30% of sporadic GC harbour oncogenic mutations in these genes (1,3–5,18,19). As our analysis

revealed that mutations at MLK3 are most probably functionally relevant, occurring almost exclusively in MSI carcinomas, and that MLK3 is described as a component of the multiprotein BRAF/RAF1 complex (7,9,10,20), we decided to analyse how mutations in MLK3 (all types) correlated with mutations in KRAS and/or BRAF in our series (Table 2 and Fig. 5).

In hereditary (HNPCC and Lynch) MSI tumours, 12.5% (6/48) harboured only MLK3 mutations; 25% (12/48) harboured only KRAS mutations and 12.5% (6/48) harboured mutations in both genes. In sporadic MSI CRCs, 5.6% (2/36) harboured only MLK3 mutations, 16.6% (6/36) harboured only KRAS mutations, 16.6% (6/36) harboured only BRAF mutations, 5.5% (2/36) harboured mutations in MLK3 and KRAS and 8.3% (3/36) had both MLK3 and BRAF mutations. In MSI sporadic GC, 13.3% (4/30) harboured only MLK3 mutations, 3.3% (1/30) harboured mutations in both MLK3 and KRAS genes, and none of the GC tumours harboured only mutations in KRAS.

Overall, 12 out of 24 (50%) of the MLK3 mutations occur in MSI gastrointestinal carcinomas bearing wild-type KRAS and BRAF (Table 2 and Fig. 5). Seven of these are of the missense type (Tables 1 and 2) and three of them showed transforming potential *in vitro* and/or tumorigenic potential *in vivo*.

DISCUSSION

In the last years, our group and others contributed to define the frequency and type of activating oncogenic mutations in MSI and MSS sporadic gastrointestinal tumours, namely occurring in KRAS and BRAF genes, which are members of the MAPKinase pathway (1,3–5,16,18,19,21–24). However fairly frequent, mutations at MAPK-related genes, like KRAS and BRAF, are not found in all gastrointestinal
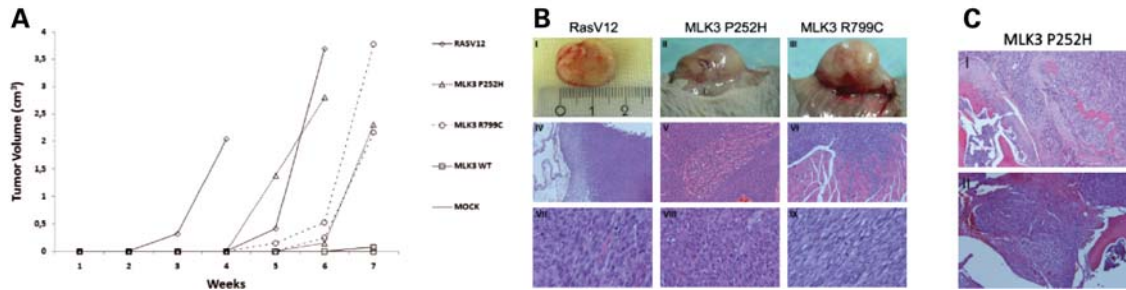
Figure 4. *In vivo* tumorigenic capacities of MLK3 mutations. (**A**) Tumour volume was monitored over time and plotted for each mouse. NIH3T3 cells over-expressing the P252H and R799C MLK3 mutations were able to generate tumours when subcutaneously inoculated in nude mice, as happens with HRASV12. Cells over-expressing wild-type MLK3 or the empty vector (Mock) were not able to generate tumours at 7 weeks after inoculation. At the end of the seventh week, one of the two mice inoculated with wild-type MLK3 expressing cells, presented a small tumour, according to the previously shown wild-type MLK3 *in vitro* transforming potential. (**B** and **C**) Histopathological analysis of tumours over-expressing MLK3 P252H and R799C mutations. (**B**) Effect of RASV12 (I), MLK3 P252H (II) and MLK3 R799C (III) transfection on the growth of NIH3T3 cells in nude mice. Tumour cells harbouring MLK3 mutations (V, VI) showed an infiltrative pattern of growth, with invasion of the surrounding adipose tissue and muscle, in contrast to RASV12 (IV). All the mutations induced high-grade fibrosarcomas, with a high number of mitotic figures (VII; VIII, IX). (**C**) MLK3 P252H-overexpressing tumours showed the most aggressive behaviour, with invasion of backbone, bone marrow and spinal cord.

tumours (1,3–5,18,19), although MAPK activation is reported to occur in over 70% of overall CRC cases (22).

MLK3 is known to be required for mitogen activation of BRAF and ERK (7,9,10,25). So, taken all these observations together, it is plausible to believe that activation of other MAPK upstream targets may occur in CRC and, eventually, in other gastrointestinal cancers. Therefore, we tested whether MLK3 was a target for mutations in this type of cancers.

Evidence pointing to the transforming potential of MLK3 has been described in the literature. However, MLK3 gene structural alterations have never been reported in cancer. In this study, we found, for the first time, MLK3 gene mutations in hereditary and sporadic primary gastrointestinal tumours. Moreover, we verified that these mutations did not occur randomly, but were rather clustered in the group of MSI cases. Further, MLK3 mutations were only found in CRC derived cell lines with MSI phenotype. This finding is interesting since, in CRC, it mimics what happens with BRAF gene, in which mutations also occur only in MSI cases (1,3,18,19) and, in GC, it mimics what happens with KRAS and PI3KCA genes, also mutated only in cancer cases with MSI phenotype (16,21,24).

As BRAF mutations have been described in 60% of melanoma (26), we tested whether MLK3 mutations would also occur in a small series of six melanoma cell lines (M14, UACC62, SKMEL2, SKMEL5, SKMEL28 and MEWO). Noteworthy, two MLK3 missense mutations (E121D and T680M) were found in two of these cell lines (SKMEL2 and UACC62) (data not shown), raising the hypothesis that MLK3 mutations might also occur in other types of malignant cancer associated to the activation of the MAPKinase pathway.

Further supporting the potential relevance of MLK3 mutations is the fact that most of them were missense (64%), spread throughout the MLK3 gene, and affecting amino acid residues that are evolutionary conserved, as it has been previously described for MLK4 in CRC, another MLK family member (27). Changes in less conserved amino acids were mainly found in the P/S/T-rich domain. This result may be related to the fact that this is the less conserved domain among MLK family and considered to confer function specificity to each MLK member (8,28).

A 3D predictive modelling of some of the MLK3 missense mutations revealed that, although located at the kinase domain of MLK3, the protein phosphorylation was not envisaged to be modified. Supporting this prediction, HEK293 cell lines transiently transfected with A165S and P252H mutations did not show increased levels of phosphorylated MLK3, using western blotting and an antibody for phospho-MLK3 (data not shown). Rather, the 3D model suggests that MLK3 mutants may interfere with its scaffold properties, and thus are likely to impact MAPkinase signalling. Further studies are needed to prove this hypothesis.

The transforming potential of several MLK3 mutations showed that the introduction of a mutant MLK3, in cells with an MLK3 wild-type background, was sufficient to confer significant higher *in vitro* transforming potential to these cells in comparison to cells over-expressing wild-type MLK3. These results demonstrate that mutant MLK3 proteins are likely to be functionally relevant, at least *in vitro*.

We additionally demonstrated the relevance of MLK3 mutations *in vivo*, since nude mice injected with cells expressing mutant MLK3 developed subcutaneous malignant tumours. Most importantly, we verified that these tumours had an infiltrative pattern of growth, a neoplastic behaviour that was not even observed in tumours induced by RASV12 oncogene. The local invasive ability of tumour cells expressing mutant MLK3, *in vivo*, was further supported by increased number of invading cells in matrigel invasion assay. The mechanism for invasion mediated by mutant MLK3 remains to be determined; nevertheless, a role for MLK3 in cell migration has been previously demonstrated (29). Interestingly, it was verified that MLK3-depleted cells showed defects in cell migration, with an increase in the thickness and number of stress fibres, as well as enlarged focal adhesions and absence of lamellipodia protrusions. These results showed that MLK3 is pivotal for cells to move. Further, it was proved that these non-migrating cells presented high levels of activated Rho, and the mechanism of MLK3-mediated Rho inhibition was independent of MLK3 kinase activity; instead, it was dependent of its scaffolding properties, as we also verified in our predictive 3D model for the MLK3 mutants found in this gastrointestinal tumours (29).
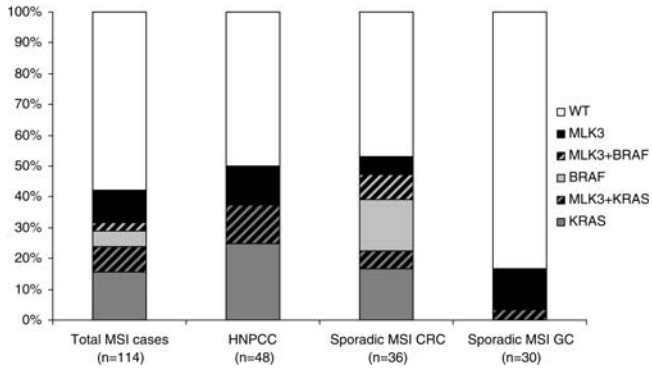
Figure 5. Graphic representation of the frequency of MLK3, KRAS and BRAF mutations in MSI primary carcinomas. From the left to the right, columns represent the frequency of cases with KRAS alone or in combination with MLK3, BRAF alone or in combination with MLK3, and MLK3 alone in the overall series of primary carcinomas, in the hereditary setting and in sporadic forms of colorectal and gastric carcinoma.

Finally, we found that, in MSI primary gastrointestinal tumours, half of the MLK3 mutations (12/24) were found in cases harbouring wild-type KRAS or BRAF genes. Nowadays, and imperative in clinical terms, the presence of mutations in genes belonging to KRAS/BRAF/MAPkinase pathway raises important therapeutic implications, namely in the selection of patients that can benefit from the treatment with monoclonal antibodies against EGFR (23,30). Most importantly, CRCs frequently harbour concomitant alterations in several genes, that lie up and downstream of KRAS or BRAF signalling cascade, which may represent an important factor for the acquisition of resistance to current therapies (16,31). This scenario is particularly relevant in the case of CRC patients, since the presence of KRAS mutations constitute exclusion criteria to treat metastatic colorectal patients with anti-EGFR therapies (30). Since it is known that MLK3 plays a role in controlling the signal transduction of KRAS and BRAF (7,9,10), it is tempting to speculate that MLK3 mutations are likely to constitute, in the near future, a biomarker for therapeutic selection of CRC patients for anti-EGFR therapies.

In conclusion, we describe for the first time MLK3 as a new cancer-related mutated kinase, associated to MSI gastrointestinal tumours and, eventually, to melanoma. Moreover, we show that missense MLK3 mutations, localized in the different domains of the protein, harbour transforming and tumorigenic potential, *in vitro* and *in vivo*. Although further studies are needed to unravel the timing of occurrence of MLK3 mutations, as well as its associated cellular effects and signalling pathways, our data put forward a new candidate biomarker and/or an attractive potential target for future therapy in MSI gastrointestinal tumours, namely in the hereditary setting.

MATERIALS AND METHODS

Tumour samples

A total of 174 primary gastrointestinal tumours and seven colorectal cancer cell lines—HCT15, HCT116, RKO, CO115, HT29, SW480 and Caco2, were analysed for MLK3 mutations. Tumour DNA samples were obtained from: University Hospital of Groningen and University Medical

Centre of Nijmegen (The Netherlands), Hospital of S. João and GDPN from Porto (Portugal), Saint-Antoine Hospital from Paris (France), University of Helsinki (Finland), Karolinska University Hospital from Stockholm (Sweden) and Hospital Universitari Vall d'Hebron from Barcelona (Spain). Genomic DNA was isolated from macro-dissected frozen or paraffin-embedded tumour tissues using standard methods. DNA from normal controls was used for the analyses of 100 chromosomes to exclude polymorphisms. The study protocol was reviewed and approved by the appropriate Ethics Committees, and control, constitutional and tumour samples were obtained with informed consent and in compliance with Helsinki declaration (http://www.wma.net/e/policy/b3.htm). All cases analysed were characterized for microsatellite status. Thirty-eight of 48 HNPCC families were characterized for MSH2, MLH1 or MSH6 germline mutations. KRAS and BRAF mutations were screened using current protocols. Additionally, DNA from six melanoma cell lines (M14, UACC62, SKMEL2, SKMEL5, SKMEL28 and MEWO) was also analysed for MLK3 mutations.

MLK3 mutation screening

All exons and intron–exon boundaries of MLK3 (ENSG00000173327) were screened for mutations. Exon 1 and exon 9 were subdivided in order to allow amplification by PCR. Primer sequences are described in Supplementary Material, Table S1. Except for exon 9, a multiplex PCR approach was used to amplify MLK3 sequence, using the QuantiTect Multiplex PCR Kit (Qiagen) and following the manufacturer's instructions. Exon 9 was amplified using a PCR Enhancer Solution (Invitrogen) and $MgSO_4$. The purified PCR product was directly sequenced. Sequence alterations were validated with a second independent PCR.

Homology modelling

The three-dimensional modelling was performed using the Swiss-PdbViewer v8.05 program package (32). The 3D model of the MLK3 kinase domain was constructed using the crystal structure of the MLK1 as a template (33). The initial model was subsequently subjected to energy minimization by using the GROMOS force field (34), as implemented in Swiss-PdbViewer. The quality of the structure model was validated using the PROCHECK suite of programs (35). Consequently, this structure served as the model structure for mapping MLK3 somatic mutations. Structures were manipulated using the Swiss-PDB viewer and were rendered using Yasara View program (36).

cDNA constructs and mutagenesis

Wild-type MLK3 and mutant sequences were cloned into pLENTID6/V5 directional TOPO (Invitrogen). Mutant MLK3 sequences (Y99C, K144R, A165S, P252H, R799C and P840L) were generated by site-directed mutagenesis, using the MLK3 wild-type sequence, also cloned in pLENTID6/V5, as a template. Primer sets used to produce MLK3 mutant sequences are described in Supplementary Material, Table S2. pLENTID6/V5 empty vector (Mock)

was obtained by the insertion of a small fragment of cDNA, in order to circularize the plasmid. The pRK5 vector expressing Myc-HRASV12 was used as a control in *in vivo* assays.

Cell lines

Mouse NIH3T3 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Gibco, Invitrogen), supplemented with 10% NCS and 1% penicillin–streptomycin (Gibco, Invitrogen). Human HEK293 were maintained in DMEM (Gibco, Invitrogen) supplemented with 10% FBS and 1% penicillin–streptomycin (Gibco, Invitrogen). Both cell lines were grown in a humidified incubator with 5% CO_2 at $37°C$.

Transfections

Transient transfections of NIH3T3 and HEK293 cells were carried out using Lipofectamine 2000 (Invitrogen), according with technical information provided by the manufacturer. For HEK293 stable transfections, ViraPower Lentiviral Expression kit (Invitrogen) was used for the transduction of the MLK3 wild-type and mutant P252H sequences, as well as the empty vector. Lentiviral transduction was performed following the manufacturer's instructions. Transduced cells were selected by antibiotic resistance to blasticidin (12 μg/ml) (Gibco, Invitrogen). The expression levels of MLK3 in the different clones selected was measured by western blot.

Focus formation assay

Low passage NIH3T3 cells, seeded on 35 mm dishes at 60–80% confluence, were co-transfected with MLK3 vectors and a vector expressing GFP protein (pCMV-GFP). Equivalent amounts of ectopic protein expression between different transfections were achieved, adjusting the concentration of each individual plasmid that was used. When required, the total amount of transfected DNA was adjusted with an empty vector. Twenty four hours later, cells were trypsinized, split into two 100 mm dishes and maintained in DMEM supplemented with 5% (v/v) new born calf serum (CS; Invitrogen). The medium was changed every 3 days thereafter. Twenty-nine days after, cells were fixed with methanol, and the GFP fluorescence of the foci was confirmed under an inverted fluorescence microscope (Leica DM 2000). Cells were then stained with 0.4% crystal violet in methanol, in order to count the foci and photograph the dishes. The experiments were done in triplicate and the levels of ectopic protein expression were monitored by western blot after each transfection.

In vivo assays

Female N:NIH(s)II:nu/nu nude mice are reproduced, maintained and housed at IPATIMUP Animal House, sited at the Medical Faculty of the University of Porto, in a pathogen-free environment, under controlled conditions of light and humidity. Animal experiments were carried out in accordance with the Guidelines for the Care and Use of Laboratory Animals, directive 86/609/EEC. Nude mice, aged 6–7 weeks, were used for *in vivo* experiments. Mice were subcutaneously injected in the dorsal flanks, using a 25-gauge needle, with 2×10^6 of NIH3T3 cells transfected with MLK3 wild-type sequence, or with the mutants MLK3-P252H or MLK3-R799C. Cells transfected with the empty vector or with the mutant HRASV12 were also injected, to be used as negative and positive controls, respectively. Mice were weighted, and tumour width and length were measured with calipers every week. Each mouse was euthanized 2–3 weeks after tumour development in order to minimize suffering of the mice. The *in vivo* experiment was stopped at the seventh week after inoculation, when the last mice that were inoculated with mutant cells (MLK3-P252H and MLK3-R799C) had already developed large and infiltrative tumours. At this time, mice that were subcutaneously inoculated with cells transfected with wild-type MLK3 or with the empty vector (Mock) were also euthanized, in order to allow comparison with mutants, and tumour development was evaluated. Histopathology of the tumours was evaluated using 5 μm sections and conventional Hematoxilin and Eosin (H&E) staining.

Matrigel invasion assay

Prior to each experiment, 24-well matrigel-coated invasion inserts of 8 μm pore size filters (Becton and Dickinson) were introduced into 24-well plates. For re-hydration, the inner and outer compartments of the system were filled with DMEM medium, and incubated for 2 h at $37°C$. After rehydration, 7×10^4 cells were added to the upper part of the insert, re-suspended in DMEM supplemented with 1% FBS, whereas the lower part was filled with DMEM supplemented with 10% FBS, and incubated for 24 h at $37°C$. Non-invasive cells were removed and filters were washed in PBS, fixed in methanol and mounted in Vectashield with 4',6-diamidino-2-phenylindole (DAPI, Vector Laboratories, Burlingame, CA). Invasive cells were counted ($\times 20$ objective), corresponding to the DAPI-counterstained nuclei, which represent cells that passed through the pores of the filter.

Statistical analysis

The statistical analysis was performed using Student's *t*-test or Fisher's exact test, when appropriated. Differences were taken to be statistically significant when $P < 0.05$.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

REFERENCES

1. Domingo, E., Espin, E., Armengol, M., Oliveira, C., Pinto, M., Duval, A., Brennetot, C., Seruca, R., Hamelin, R., Yamamoto, H. *et al.* (2004) Activated BRAF targets proximal colon tumors with mismatch repair deficiency and MLH1 inactivation. *Genes Chromosomes Cancer*, **39**, 138–142.

2. Duval, A. and Hamelin, R. (2002) Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res.*, **62**, 2447–2454.

3. Oliveira, C., Pinto, M., Duval, A., Brennetot, C., Domingo, E., Espin, E., Armengol, M., Yamamoto, H., Hamelin, R., Seruca, R. *et al.* (2003) BRAF mutations characterize colon but not gastric cancer with mismatch repair deficiency. *Oncogene*, **22**, 9192–9196.

4. Domingo, E., Niessen, R.C., Oliveira, C., Alhopuro, P., Moutinho, C., Espin, E., Armengol, M., Sijmons, R.H., Kleibeuker, J.H., Seruca, R. *et al.* (2005) BRAF-V600E is not involved in the colorectal tumorigenesis of HNPCC in patients with functional MLH1 and MSH2 genes. *Oncogene*, **24**, 3995–3998.

5. Oliveira, C., Westra, J., Arango, D., Ollikainen, M., Domingo, E., Ferreira, A., Velho, S., Niessen, R., Lagerstedt, K., Alhopuro, P. *et al.* (2004) Distinct patterns of KRAS mutations in colorectal carcinomas according to germline mismatch repair defects and hMLH1 methylation status. *Hum. Mol. Genet.*, **13**, 2303–2311.

6. Bock, B.C., Vacratsis, P.O., Qamirani, E. and Gallo, K.A. (2000) Cdc42-induced activation of the mixed-lineage kinase SPRK *in vivo*. Requirement of the Cdc42/Rac interactive binding motif and changes in phosphorylation. *J. Biol. Chem.*, **275**, 14231–14241.

7. Chadee, D. and Kyriakis, J. (2004) A novel role for mixed lineage kinase 3 (MLK3) in B-Raf activation and cell proliferation. *Cell Cycle*, **3**, 1227–1229.

8. Gallo, K.A. and Johnson, G.L. (2002) Mixed-lineage kinase control of JNK and p38 MAPK pathways. *Nat. Rev. Mol. Cell. Biol.*, **3**, 663–672.

9. Chadee, D.N. and Kyriakis, J.M. (2004) MLK3 is required for mitogen activation of B-Raf, ERK and cell proliferation. *Nat. Cell Biol.*, **6**, 770–776.

10. Chadee, D.N., Xu, D., Hung, G., Andalibi, A., Lim, D.J., Luo, Z., Gutmann, D.H. and Kyriakis, J.M. (2006) Mixed-lineage kinase 3 regulates B-Raf through maintenance of the B-Raf/Raf-1 complex and inhibition by the NF2 tumor suppressor protein. *Proc. Natl Acad. Sci. USA*, **103**, 4463–4468.

11. Bishop, J.M. (1987) The molecular genetics of cancer (genetic damage in neoplastic cells). *Science*, **v235**, 305(7).

12. Hartkamp, J., Troppmair, J. and Rapp, U.R. (1999) The JNK/SAPK activator mixed lineage kinase 3 (MLK3) transforms NIH 3T3 cells in a MEK-dependent fashion. *Cancer Res.*, **59**, 2195–2202.

13. Gallo, K.A., Mark, M.R., Scadden, D.T., Wang, Z., Gu, Q. and Godowski, P.J. (1994) Identification and characterization of SPRK, a novel src-homology 3 domain-containing proline-rich kinase with serine/threonine kinase activity. *J. Biol. Chem.*, **269**, 15092–15100.

14. Duval, A., Reperant, M., Compoint, A., Seruca, R., Ranzani, G.N., Iacopetta, B. and Hamelin, R. (2002) Target gene mutation profile differs between gastrointestinal and endometrial tumors with mismatch repair deficiency. *Cancer Res.*, **62**, 1609–1612.

15. Sammalkorpi, H., Alhopuro, P., Lehtonen, R., Tuimala, J., Mecklin, J.-P., Jarvinen, H.J., Jiricny, J., Karhu, A. and Aaltonen, L.A. (2007) Background mutation frequency in microsatellite-unstable colorectal cancer. *Cancer Res.*, **67**, 5691–5698.

16. Velho, S., Oliveira, C., Ferreira, A., Ferreira, A.C., Suriano, G., Schwartz, J.S., Duval, A., Carneiro, F. and Machado, J.C. (2005) The prevalence of PIK3CA mutations in gastric and colon cancer. *Eur. J. Cancer*, **41**, 1649–1654.

17. Hudkins, R.L., Diebold, J.L., Tao, M., Josef, K.A., Park, C.H., Angeles, T.S., Aimone, L.D., Husten, J., Ator, M.A., Meyer, S.L. *et al.* (2008) Mixed-lineage kinase 1 and mixed-lineage kinase 3 subtype-selective dihydronaphthyl[3,4-a]pyrrolo[3,4-c]carbazole-5-ones: optimization, mixed-lineage kinase 1 crystallography, and oral in vivo activity in 1-methyl-4-phenyltetrahydropyridine models. *J. Med. Chem.*, **51**, 5680–5689.

18. Oliveira, C., Velho, S., Moutinho, C., Ferreira, A., Preto, A., Domingo, E., Capelinha, A.F., Duval, A., Hamelin, R., Machado, J.C. *et al.* (2006) KRAS and BRAF oncogenic mutations in MSS colorectal carcinoma progression. *Oncogene*, **26**, 158–163.

19. Rajagopalan, H., Bardelli, A., Lengauer, C., Kinzler, K.W., Vogelstein, B. and Velculescu, V.E. (2002) Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. *Nature*, **418**, 934–934.

20. Zhang, H. and Gallo, K.A. (2001) Autoinhibition of mixed lineage kinase 3 through its Src homology 3 domain. *J. Biol. Chem.*, **276**, 45598–45603.

21. Lee, S.H., Lee, J.W., Soung, Y.H., Kim, H.S., Park, W.S., Kim, S.Y., Lee, J.H., Park, J.Y., Cho, Y.G., Kim, C.J. *et al.* (2003) BRAF and KRAS mutations in stomach cancer. *Oncogene*, **22**, 6942–6945.

22. Scartozzi, M., Bearzi, I., Berardi, R., Mandolesi, A., Pierantoni, C. and Cascinu, S. (2007) Epidermal growth factor receptor (EGFR) downstream signalling pathway in primary colorectal tumours and related metastatic sites: optimising EGFR-targeted treatment options. *Br. J. Cancer*, **97**, 92–97.

23. Seruca, R., Velho, S., Oliveira, C., Leite, M., Matos, P. and Jordan, P. (2009) Unmasking the role of KRAS and BRAF pathways in MSI colorectal tumors. *Expert Rev. Gastroenterol. Hepatol.*, **3**, 5–9.

24. Zhao, W., Chan, T.L., Chu, K.M., Chan, A.S., Stratton, M.R., Yuen, S.T. and Leung, S.Y. (2004) Mutations of BRAF and KRAS in gastric cancer and their association with microsatellite instability. *Int. J. Cancer*, **108**, 167–169.

25. Kyriakis, J.M. (2007) The integration of signaling by multiprotein complexes containing Raf kinases. *Biochim. Biophys. Acta (BBA) Mol. Cell Res.*, **1773**, 1238–1247.

26. Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W. *et al.* (2002) Mutations of the BRAF gene in human cancer. *Nature*, **417**, 949–954.

27. Bardelli, A., Parsons, D.W., Silliman, N., Ptak, J., Szabo, S., Saha, S., Markowitz, S., Willson, J.K.V., Parmigiani, G., Kinzler, K.W. *et al.* (2003) Mutational analysis of the tyrosine kinome in colorectal cancers. *Science*, **300**, 949.

28. Handley, M.E., Rasaiyaah, J., Chain, B.M. and Katz, D.R. (2007) Mixed lineage kinases (MLKs): a role in dendritic cells, inflammation and immunity? *Int. J. Exp. Pathol.*, **88**, 111–126.

29. Swenson-Fields, K.I., Sandquist, J.C., Rossol-Allison, J., Blat, I.C., Wennerberg, K., Burridge, K. and Means, A.R. (2008) MLK3 limits activated G[alpha]q signaling to Rho by binding to p63RhoGEF. *Mol. Cell.*, **32**, 43–56.

30. Velho, S., Oliveira, C. and Seruca, R. (2009) KRAS mutations and anti-epidermal growth factor receptor therapy in colorectal cancer with lymph node metastases. *J. Clin. Oncol.*, **27**, 158–159.

31. Oliveira, C., Velho, S., Domingo, E., Preto, A., Hofstra, R.M.W., Hamelin, R., Yamamoto, H., Seruca, R. and Schwartz, S. Jr (2005) Concomitant RASSF1A hypermethylation and KRAS//BRAF mutations occur preferentially in MSI sporadic colorectal cancer. *Oncogene*, **24**, 7630–7634.

32. Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, **22**, 195–201.

33. Brown, K., Vial, S., Dedi, N., Long, J., Dunster, N. and Cheetham, G. (2005) Structural basis for the interaction of TAK1 kinase with its activating protein TAB1. *J. Mol. Biol.*, **354**, 1013–1020.

34. van Gunsteren, W., Billeter, S., Eising, A., Hünenberger, P., Krüger, P., Mark, A., Scott, W. and Tironi, I. (1996) *Biomolecular Simulation: the GROMOS96 Manual and User Guide*. Hochschulverlag an der ETH Zürich/ Biomos, Zürich/Groningen.

35. Laskowski, R.A., Chistyakov, V.V. and Thornton, J.M. (2005) PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res.*, **33**, D266–D268.

36. Sanner, M. (1999) Python: a programming language for software integration and development. *J. Mol. Graph. Model.*, **17**, 57–61.

37. Topham, C.M., Srinivasan, N. and Blundell, T.L. (1997) Prediction of the stability of protein mutants based on structural environment-dependent amino acid substitution and propensity tables. *Protein Eng.*, **10**, 7–21.